W&M ScholarWorks

Dissertations, Theses, and Masters Projects

Theses, Dissertations, & Master Projects

1998

Comparative evolution of molecular markers: An analysis of genetic variation within the blue marlin (Makaira nigricans)

Vincent Patrick Buonaccorsi
College of William and Mary - Virginia Institute of Marine Science

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Fresh Water Studies Commons, Genetics Commons, Molecular Biology Commons, Oceanography Commons, and the Zoology Commons

Recommended Citation

Buonaccorsi, Vincent Patrick, "Comparative evolution of molecular markers: An analysis of genetic variation within the blue marlin (Makaira nigricans)" (1998). *Dissertations, Theses, and Masters Projects.* Paper 1539616587.
https://dx.doi.org/doi:10.25773/v5-e9nx-f157

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

COMPARATIVE EVOLUTION OF MOLECULAR MARKERS: AN ANALYSIS OF GENETIC VARIATION WITHIN THE BLUE MARLIN (*MAKAIRA NIGRICANS*).

A Dissertation

Presented to

The Faculty of the School of Marine Science

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

---

by

Vincent Patrick Buonaccorsi

1998

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 9903259

UMI Microform 9903259
Copyright 1998, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Vincent Patrick Buonaccorsi

Approved, August 1998

John E. Graves, Ph.D.
Committee Chairman/Advisor

Kimberly S. Reece, Ph.D.

J. Emmett Duffy, Ph.D.

Peter A. Van Veld, Ph.D.

John R. Gold, Ph.D.
Department of Wildlife and Fisheries Sciences
Texas A & M University

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

	Page
ACKNOWLDGEMENTS	vi.
LIST OF TABLES	vii.
LIST OF FIGURES	ix.
ABSTRACT	xii.
CHAPTER 1. GENERAL INTRODUCTION	1.
Evolutionary forces	2.
Molecular markers	5.
Marlin fisheries and life history characteristics	12.
Dissertation outline	14.
LITERATURE CITED	16.
CHAPTER 2. COMPARATIVE INTRASPECIFIC EVOLUTION OF MOLECULAR MARKERS IN THE BLUE MARLIN (*MAKAIRA NIGRICANS*): A HIERARCHICAL ANALYSIS OF VARIANCE REVEALED BY ALLOZYME, SINGLE-COPY NUCLEAR DNA, AND MITOCHONDRIAL DNA MARKERS	23.
INTRODUCTION	24.
MATERIALS AND METHODS	28.
Blue marlin tissue samples	28.
Allozymes	28.
Single-copy nuclear DNA loci	28.
Statistical analyses	30.
RESULTS	34.
Allozymes	34.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ScnDNA 34.
MtDNA 37.
DISCUSSION 39.
Allozymes 39.
Single-copy nuclear DNA 39.
Population genetic implications of scnDNA loci 40.
Concordance of scnDNA and allozyme markers 41.
Discordance of nuclear and mitochondrial markers 42.
Evolutionary forces 43.
LITERATURE CITED 46.

CHAPTER 3. ANALYSIS OF BLUE MARLIN POPULATION STRUCTURE AND PHYLOGEOGRAPHIC HISTORY WITH MTDNA AND MICROSATELLITE DNA MOLECULAR MARKERS. 70.
INTRODUCTION 71.
MATERIALS AND METHODS 73.
Blue marlin tissue samples 73.
Microsatellite library construction 73.
Population screening 76.
Microsatellite statistical analyses 77.
MtDNA 82.
RESULTS 83.
Microsatellites
Mini-library 83.
Hardy-Weinberg equilibrium 83.
Diversity 84.
Inter-annual and intra-ocean divergence 87.
Inter-ocean divergence 89.
Pairwise population comparisons 90.
Comparison of phenetic trees 91.
MtDNA

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Diversity 92.
AMOVA 92.
Inter-annual and Intra-ocean divergence 93.
Inter-ocean divergence 94.
Correspondence between mitochondrial clades and nuclear allele frequencies 94.
Male-mediated dispersal 95.
DISCUSSION
Microsatellites 96.
MtDNA 99.
Phylogeography 100.
Sex-biased dispersal 104.
Unidirectional gene flow 105.
Clade tests 105.
Statistical considerations and mutation, migration and drift equilibrium 106.
LITERATURE CITED 110.

CHAPTER 4. COMPARATIVE EVOLUTION OF ALLOZYME, SCNDNA, MTDNA, AND MICROSATELLITE DNA MOLECULAR MARKERS 170.
Mutation rates 171.
Hardy-Weinberg equilibrium 172.
Marker comparison: inter-annual and intra-ocean divergence 173.
Marker comparison: inter-ocean divergence 174.
Evolutionary processes 177.
Conclusions 181.
LITERATURE CITED 184.
VITA 187.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGMENTS

I gratefully acknowledge the help of my dissertation committee. This project would not have been possible without the generosity of my major advisor, Dr. John Graves. I sincerely appreciate the time, thought, guidance, financial support, patience, and encouragement provided through every aspect of this study. Dr. Kim Reece provided critical reviews and invaluable molecular expertise, greatly facilitating the development of my molecular markers. The editorial and theoretical contributions of Dr. John Gold and Dr. J. Emmett Duffy have greatly improved this dissertation. Thanks to Dr. Peter Van Veld, for encouragement, critical review, and helpful suggestions.

Thanks to those who assisted in sample collection: J. Graves, J. McDowell, E. Prince, J. Pepperrell, E. Evertt, C. Heberer, and G. Harvey. Thanks also to the officers and crew of the R/V *Townsend Cromwell.* This project was funded by cooperative agreement #NA37RJ0199 between the Joint Institute of Marine and Atmospheric Research (University of Hawaii) and the National Oceanic and Atmospheric Administration. Additional funding was obtained from two VIMS minor research grants, and an assistantship from the VMRC.

I owe a special thanks to my family and friends, for their love and support. Thank you Kristin for putting up with me and encouraging my every step. You are my intensity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Table Page

CHAPTER 2

1 Blue marlin samples surveyed for each marker 52.

2 ScnDNA restriction endonuclease survey 53.

3 ScnDNA allele frequencies, confidence intervals, and sample size per locus 54.

4 Nuclear marker comparison: diversity and inter-ocean divergence 56.

5. ScnDNA intra-ocean hierarchical analysis of variance 57.

6 AMOVA inter-ocean divergence estimates from mtDNA RFLP data of Graves and McDowell, 1995 58.

7 Test of independence between nuclear and mitochondrial allele frequencies within the Atlantic 59.

CHAPTER 3

1 Blue marlin samples used for mtDNA, microsatellite, and scnDNA analyses 116.

2 Microsatellite clone sequences 118.

3 *f* Values 119.

4 Microsatellite diversity by ocean and clade 120.

5 Microsatellite diversity by year and location 122.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6A	Hierarchical AMOVA: mtDNA, microsatellite, scnDNA markers	123.
6B	Detailed AMOVA: mtDNA, microsatellites	124.
6C	Inter-ocean divergence and heterozygosity values: allozymes, scnDNA, microsatellite, mtDNA	126.
7	Pairwise results summarized by ocean	127.
8	Phenetic trees summary	128.
9	Test of nuclear and mitochondrial allele frequency independence	129.
10	Mitochondrial sex frequency data	130.
11	Mitochondrial divergence by sex	132.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Figure		Page
Chapter 2		
1.	Dispersal capability of the blue marlin	60.
2.	*f* values for each sample at four scnDNA loci and four allozyme loci	62.
3.	ScnDNA hierarchical analysis of variance over all populations	64.
4.	ScnDNA neighbor-joining phenetic trees	66.
5.	Results of inter-ocean pairwise tests for population differentiation	68.
	A. Among individual samples	
	B. Among pooled samples	
Chapter 3		
1.	*f* Distribution: microsatellite, allozyme, scnDNA	133.
2.	Frequency distributions	135.
	A. GATA01	
	B. GATA08	
	C. GATA10	
	D. GATA60	
	E. GATA90	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

F. MtDNA trimmed haplotype tree with frequencies

3. Mismatch distributions overall populations for microsatellite and mtDNA data 142.

4. Mismatch distributions: Atlantic vs Pacific oceans 144.

A. GATA01

B. GATA08

C. GATA10

D. GATA60

E. GATA90

F. MtDNA

5. Mismatch distribution: special cases 149.

A. GATA08: Atlantic vs ubiquitous modes

B. MtDNA: Atlantic vs ubiquitous clades

6. Mismatch distribution: Atlantic vs Pacific over all microsatellite loci 152.

7. Overall AMOVA: microsatellite, scnDNA, mtDNA 154.

8. Inter-ocean difference in allele frequency 156.

A. GATA01

B. GATA08

C. GATA10

D. GATA60

E. GATA90

F. GATA60: Inter-annual pairwise comparisons of Hawaiian samples

9. Neighbor joining phenetic trees 161.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A. Microsatellite linearized *R*

B. Microsatellite linearized θ

C. Microsatellite *Dmu*

D. Overall scnDNA linearized θ, overall microsatellite linearized *R*,

Overall mtDNA linearized Φ

10. Results of pairwise comparisons 166.

A. Microsatellite exact *R*

B. Microsatellite exact probability

C. MtDNA exact Φ and exact probability

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

The blue marlin (*Makaira nigricans*) was chosen as a model organism to perform a comparative analysis of intra-specific molecular diversity, with three main goals: (i) to resolve the population genetic structure of the blue marlin to the extent possible given sampling limitations; (ii) to determine the sensitivity of a variety of molecular markers to population structure; and (iii) to evaluate the potential contributions of various evolutionary forces to the patterns of molecular diversity obtained among populations for each marker class.

A hierarchical analysis of molecular variance revealed that a majority of genetic variation was maintained within populations, with only a minor fraction attributable to variation among temporal replicates and between locations within oceans. In contrast, inter-ocean divergence was highly significant for a majority of loci within each marker class.

Previous studies of mitochondrial DNA (mtDNA; n = 104) genetic variation within blue marlin revealed two distinct clades of haplotypes, one of which was present only in the Atlantic (the 'Atlantic clade') at a frequency of 40%. To examine whether such pronounced inter-ocean divergence would be evident at the nuclear genome, single-copy nuclear DNA (scnDNA) and microsatellite DNA molecular markers were developed, and blue marlin samples screened for diversity. In addition, a previous allozyme data set and a greatly expanded mitochondrial data set (Graves and McDowell, unpublished data) were analyzed. ScnDNA and allozyme markers exhibited lower levels of diversity and inter-ocean divergence than mtDNA analyses. Enhanced genetic drift among populations, due to the four-fold lower effective population size of mtDNA, could be responsible for greater inter-ocean divergence. However, the low mutation rate of these markers, and possible homogenizing influence of male-mediated dispersal could also have contributed to the differences detected. To further explore the role of mutation rate, patterns of diversity were surveyed among oceans for hyper-variable microsatellite markers.

Microsatellite loci displayed a wide range of divergence estimates, a number of which were larger than those detected in previous nuclear markers. Among the patterns detected, a nuclear 'Atlantic clade' of alleles was found at one locus, indicating that the forces that generated the mitochondrial Atlantic clade also strongly influenced the nuclear genome. Although some microsatellite loci were much more sensitive to inter-ocean population structure than scnDNA and allozyme markers, on average, these differences were not significant, due to the wide range of microsatellite patterns detected. Inter-ocean divergence remained small on average for nuclear (F = 0.10) than mitochondrial DNA (F = 0.39) loci. The means and variances of inter-ocean divergence estimates (F) of the different nuclear marker classes were not significantly different, nor were they different from the theoretical expectations of the F distribution. These results indicated that selection did not significantly influence the diversity of loci within and among nuclear marker classes, and was consistent with the predominant influence of genetic drift. Correlations between diversity and divergence within and among marker classes were non-significant, indicating that difference in mutation rate can not explain the lower nuclear inter-ocean divergence. The patterns of diversity obtained within and among marker classes were consistent with expected values under migration-drift equilibrium.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

COMPARATIVE EVOLUTION OF MOLECULAR MARKERS: AN ANALYSIS OF GENETIC VARIATION WITHIN THE BLUE MARLIN (*MAKAIRA NIGRICANS*).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1

GENERAL INTRODUCTION

Rapid technical advances in the field of molecular biology have greatly facilitated development of assays for molecular characters that reveal genetic variation within species. However, the rapid increase in assays of molecular markers has outpaced our basic understanding of how markers with different modes and rates of evolution describe common demographic population parameters. Comparative studies of molecular markers provide detailed information concerning a demographic parameter of interest and enhance our basic understanding of how population level processes affect marker evolution.

In this dissertation, a comparative approach is used to study natural history and molecular marker evolution in blue marlin. The comparative molecular approach is used to: 1) estimate population structure and gene flow throughout the geographic range of blue marlin; 2) determine the sensitivity of a variety of molecular markers to population structure; and 3) investigate how evolutionary forces affect patterns of molecular diversity of several commonly utilized molecular markers.

Evolutionary Forces

The evolutionary forces of mutation, genetic drift, migration, and natural selection have been identified as important factors influencing the distribution of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

molecular diversity. Mutation is defined as a heritable change in the genetic material (Lincoln et al., 1982), and is the ultimate source of genetic diversity. Mutation tends to identify isolated populations through accumulation of random character changes. These changes typically accumulate very slowly. Production of genetic diversity by mutation is balanced by loss of diversity due to genetic drift and selection. This loss of diversity is inversely proportional to effective population size of a population ($4N_eu$, where u = mutation rate).

Genetic drift refers to random changes in allele frequencies due to sampling error in populations of finite size. The magnitude of genetic drift is inversely proportional to effective population size. Effective population size is defined as the average number of individuals in a population that are assumed to contribute genes equally to the succeeding generation (Lincoln et al., 1982). Factors that decrease effective population size include unequal sex ratios, fluctuations in population size, and unequal reproductive success among spawning adults (Avise, 1994). Genetic isolation eventually results in the establishment of fixed genetic differences between groups, and the rate to fixation of alleles is accelerated in small populations.

Migration is defined as the successful transfer of genes from one geographic region to another (Lincoln et al., 1982). Migration tends to homogenize the distribution of alleles among populations, counteracting effects of genetic drift. Eventually, an equilibrium level of divergence is established between two populations, that is inversely proportional to the absolute number of migrants (N_em) between populations (Allendorf and Phelps, 1981). Migration rate is indicated by m. The parameter N_em is usually estimated from the relationship

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$F_{ST} = 1/(4N_e m + 1)$ (Wright, 1978).

This relationship has a number of implicit assumptions (reviewed in Waples, 1998). These include the assumption that gene flow is equally likely among all populations surveyed (island model), that mutation rates are much lower than migration rates, and that equilibrium has been reached between drift and migration. For a particular allele, or character state at a locus, the parameter F_{ST} is defined by

$(F_{ST} = V_p/(p_{(AVG)})(1- p_{(AVG)})$ (Wright, 1978),

where V_p represents the variance in allele frequency (p) among populations. F_{ST} measures the proportion of genetic variance among populations, relative to total variance.

Natural selection is defined as non-random differential reproduction of different genotypes in a population (Lincoln et al., 1982). Most molecular markers, or genetic loci, are assumed to be selectively neutral, although controversy exists on this issue (Ashley and Warren, 1995; Karl and Avise, 1992; Pogson et al., 1995; Simmons et al., 1989; Whittam et al., 1986; William et al., 1995; Zouros et al., 1992). Migration and drift tend to affect all loci simultaneously, whereas natural selection is assumed to work independently on unlinked loci.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Molecular Markers

The extent to which different molecular markers are influenced by an evolutionary force depends on the mode and rate of evolution of each marker. This dissertation focuses on analyses of whole molecule mitochondrial DNA, and three classes of nuclear markers: allozymes, single-copy nuclear DNA (scnDNA), and microsatellite DNA. Each marker is discussed briefly below.

Allozymes. Allozymes are allelic variants of a particular protein representing different products of the same nuclear-encoded gene locus (Hillis and Moritz, 1990). During starch gel electrophoresis, non-denatured proteins migrate through a gel matrix to which a current has been applied (Avise, 1994). Allelic variants (allozymes) with different charges or shapes are separated as they migrate at different rates through the gel matrix. Specific histochemical stains allow visualization of a particular enzyme or non-enzymatic protein. Variation in protein charge or shape between individuals is due to amino acid substitutions that result from nucleotide substitutions. An allozyme mutation rate of 5.1×10^{-6} mutations per locus per generation has been estimated for *Drosophila* (Voelker et al., 1980). Allozymes are considered the most efficient and least expensive method that one can use to survey genetic variation among individuals (Avise, 1994). A tremendous amount of population genetic information has been obtained over the past thirty years by allozyme electrophoresis. However, there are several limitations to allozyme analysis: (i) tissues must be of excellent quality to maintain enzymatic activity; (ii) selection constrains the evolutionary rate of certain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

allozymes due to their functional significance; and (iii) allozymes frequently exhibit low levels of variation, in part due to the redundancy of the genetic code.

Mitochondrial DNA. Over the past 15 years, mtDNA has been the focus of numerous population genetic studies (Avise et al., 1987). The mitochondrial genome of vertebrates is haploid and clonally (maternally) inherited. Vertebrate mtDNA is a non-recombining molecule that lacks introns and pseudogenes, and has a relatively constant size of 16,000-18,000 bp. Because of the haploid condition and maternal inheritance, mtDNA has a smaller effective population size than loci in the nuclear genome, enhancing the effects of genetic drift (Ovenden, 1990). Except in cases of sex-biased dispersal, evolutionary histories of the mitochondrial molecule are assumed to represent that of both sexes (Palumbi and Baker, 1994). Base substitutions are the predominant mutational mechanism for mtDNA, detectable through direct sequencing or restriction enzyme digestion. Restriction enzymes digest DNA at specific 4-6 base pair sequences. Point mutations in a DNA sequence may cause a gain or loss of a restriction site, resulting in restriction fragment length polymorphisms (RFLPs).

The mutation rate of mtDNA is thought to be high relative to that of single-copy nuclear DNA or that inferred from allozymes, increasing the probability of detecting intraspecific variation.. An average mtDNA evolutionary rate of 2×10^{-8} substitutions per nucleotide site per year was calculated for primates (Brown et al., 1979). The high rate of nucleotide substitutions may be due to the presence of many free radicals within the mitochondria or to poor DNA repair mechanisms (Brown et al., 1979). A major disadvantage of basing a population genetic study entirely on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mtDNA data is that due to a lack of recombination, all sites within the mtDNA molecule are linked, making it a single genetic locus. In addition, because of maternal inheritance, mtDNA provides information only on female gene flow.

Single Copy Nuclear DNA. Single copy nuclear DNA (scnDNA) loci are non-repetitive regions of the nuclear genome. The use of scnDNA markers is relatively uncommon, although these markers are suitably polymorphic for use in population genetic studies. As with mtDNA and allozymes, base substitutions are the predominant form of mutations. ScnDNA loci may consist of protein-coding and/or non-coding regions, where variation may or may not have functional significance. Intervening sequences of genes (introns) and other non-coding regions are more likely to show variation than coding regions due to decreased selective constraints. Non-coding sequences account for 50% of yeast and 97% of human nuclear DNA (Gall, 1981; Lewin, 1975). As with other vertebrates, fish nuclear DNA likely comprises a high percentage of non-coding sequences. Consequently, there is a high probability that anonymously chosen regions will be variable. Based on whole genome hybridization studies of several primates, it was estimated initially that scnDNA mutates five to ten times more slowly than mtDNA (Brown et. al., 1979). However, a subsequent study of echinoids demonstrated roughly equal mutation rates for mtDNA and scnDNA, estimated at about one to two percent sequence divergence per million years (Vawter & Brown, 1986).

Analysis of scnDNA in population genetic studies has become practical with development of the polymerase chain reaction (PCR). One method of assessing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

scnDNA variation is PCR-based analysis of anonymously chosen scnDNA (Karl and Avise, 1993). After initial screening and development, single-copy sequences of DNA from independent regions throughout the nuclear genome are amplified and surveyed for polymorphism. This approach rapidly produces multiple nuclear markers for analysis. The identities of scnDNA sequences derived from this method generally are unknown. Because anonymously chosen regions could represent areas of the genome responding to various selective constraints, concordance of patterns of variation at several loci generally allows confidence in parameter estimation. Using an RFLP approach to screen samples for genetic variation, large numbers of individuals and nuclear loci can be surveyed rapidly.

Microsatellite DNA. Microsatellite loci form another class of polymorphic, PCR-amplifiable, nuclear markers. Microsatellite loci consist of tandemly repeated short (2-6 bp) DNA motifs (for example, GATA) commonly varying in repeat number (Tautz et al., 1986). They are extremely abundant, with estimates of up to 10^5 copies dispersed evenly throughout the nuclear genome, approximately every 6 to 10 kb (Beckmann and Weber, 1992; Brooker et al., 1994; Weber, 1990). There have been several reviews of microsatellite mutation rates and mutational processes (Goldstein and Pollock, 1997; Jame and Lagoda, 1996; O'Connell and Wright, 1997). There is extreme variation in the rate and mode of evolution among microsatellite loci, preventing generalizations, and merits of microsatellite loci are best evaluated on a locus-by-locus basis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Based on observed allele frequency distributions, the most likely model for microsatellite mutation involves several processes, including slipped-strand mispairing, unequal crossing over, and point substitutions. Slipped-strand mispairing of the two strands of the DNA double helix during replication represents the most common source of variation. This process generates alleles one or two repeats larger or smaller than the original. Unequal crossing over is presumed to occur less frequently, resulting in larger differences in allele size (Di Rienzo et al., 1994; Levinson and Gutman, 1987). Point mutations also occur occasionally, generating imperfect, or compound repeats (Levinson and Gutman, 1987; Weber, 1990). Several pedigree analyses of microsatellite mutation in humans and yeast indicated that over 90% of mutations were single step (Amos and Rubinstzein, 1996; Henderson and Petes, 1992; Weber and Wong, 1993). It has been suggested that the probability of multi-step mutations increases with total repeat range, implicating a larger role for unequal crossing over in these instances (Amos et al., 1996). Correlation between repeat variance and the probability of multi-step mutations also has been hypothesized (Di Rienzo et. al., 1994).

Recent studies have demonstrated that relationships among microsatellite alleles can not be determined unambiguously by size information alone. Even with a strict adherence to the stepwise mutation model, mutations to a previously existing state (homoplasy) occur frequently (Di Rienzo et al., 1994; Levinson and Gutman, 1987). The complex evolution of microsatellite loci has been documented in sequencing studies at both the population and species level, indicating that changes in microsatellite repeat and flanking regions are common. Sequencing studies have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shown that the structure of microsatellite loci is somewhat more conserved among populations within species than among species (Angers and Bernatchez, 1997; Estoup et al., 1995). However, mutations that include flanking-region insertions or deletions that alter the apparent repeat size within species also occur within populations (Grimaldi and Crouau-Roy, 1997). Orti et al. (1997) mapped the origin of microsatellite allele size classes within the horseshoe crab from a phylogeny generated from the flanking region of each repeat area. The authors concluded that the correlation between allele size and allele relationships was poor.

Microsatellite mutation rate estimates vary substantially among loci within the same species, between species at the same locus (McConnell et al., 1995), and even among alleles of the same locus (Jin et al., 1996). In humans, mutation rate estimates vary greatly, with reported averages of 10^{-2} (Mahtani and Willard, 1993), 10^{-3} (Weber and Wong, 1993; Weissenbach et al., 1992), and 10^{-4} to 10^{-5} mutations per locus per generation (Edwards, 1992). Pedigree analyses of yeast yielded mutation rate estimates of 10^{-3} (Henderson and Petes, 1992). Pedigree analysis of inbred strains of mice yielded mutation rates estimates of 10^{-5} (Dietrich et al., 1992) to 10^{-6} (Dallas, 1992). Observed heterozygosity at microsatellite loci can vary from 0% to 95% within a single species (Brooker et al., 1994; Tautz, 1989; Weber, 1990).

Structural properties of microsatellite loci may contribute to variation in mutation rates. Several studies have shown that loci with larger numbers of tandem repeats are more variable (Edwards, 1992; Goldstein and Clark, 1995; Hudson et al., 1992; Stephan and Cho, 1994; Weber, 1990). This implies an increased probability of slipped strand mispairing at loci with greater numbers of repeats. In contrast, based on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

an analysis of 108 loci, Valdes et al. (1993) detected no significant correlation between fragment size and mutation rate. However, allele size was not corrected for flanking region size in the study. In general, teleost fish have larger average numbers of microsatellite repeats than mammals (Brooker et al., 1994). Whether this is due to a difference in mutation rate has not been established.

Microsatellite loci without interruptions in the microsatellite repeats have shown increased levels of polymorphism (Brooker et al., 1994; FitzSimmons et al., 1995; Goldstein and Clark, 1995; Weber, 1990). The same locus amplified across species may demonstrate different levels of polymorphism, possibly due to varying levels of disruption of the microsatellite locus among species (FitzSimmons et al., 1995). A significant positive correlation also has been detected between the maximum size of a microsatellite locus and its mutation rate (Goldstein and Clark, 1995). If structure does influence the mutation rate of the molecule, it follows that even alternate alleles of the same locus could have different mutation rates. This has been demonstrated within a human population by Jin et al (1996), who superimposed the phylogeny of flanking region sequences onto closely linked microsatellite alleles. It was hypothesized that mutation rate was inhibited for alleles interrupted by point mutations.

The effect of microsatellite homoplasy and mutation rate variation on the reliability of demographic parameter estimation requires further investigation. Identity by descent is of fundamental importance when evaluating relatedness of populations through statistical estimates that incorporate allele relatedness. Mutation rate variation confounds comparisons of diversity across taxa or loci as more quickly evolving

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

markers generate diversity independent of the population-level demographic parameter of interest. In this dissertation, reliability of microsatellite demographic estimates within the blue marlin is assessed through comparison to estimates from markers without such mutational ambiguity.

Marlin Fisheries & Life History Characteristics

The organism chosen for this study was the blue marlin, *Makaira nigricans*. Substantial recreational and commercial fisheries exist for this circumtropical species (ICCAT, 1994; Nakamura, 1985). More than 90% of the world catch of billfish occurs as bycatch in tuna and swordfish longline fisheries. Blue marlin catch-per-unit-effort (CPUE) levels have declined precipitously from a peak in the early 1960s and 1970s (Beardsley, 1989), and the current stock biomass of Atlantic blue marlin is less than 25% of that necessary to achieve maximum sustainable yield (B_{msy}) (ICCAT, 1994). However, it is difficult to obtain reliable assessments of CPUE in species predominantly harvested as bycatch (Nakamura, 1985; Ueyanagi et al., 1989).

The blue marlin is a large, oceanic, epipelagic species that spends most of its life above the thermocline in tropical waters. Batch spawning takes place in tropical waters, and both eggs and larvae are pelagic (Nakamura, 1985). Blue marlin have a short larval duration and exhibit extremely rapid growth (Cyr et al., 1990; Prince et al., 1990). Adults are capable of long distance migration, although tagged individuals are frequently recaptured in the general vicinity of their release location even after years at liberty (ICCAT, 1994; Scott et al., 1990; Squire and Suzuki, 1990). Tagging results indicate that billfish migrate into subtropical or temperate waters to forage during the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

warm months, then return to tropical waters for spawning (Mather et al., 1972; Nakamura, 1985). The potential for inter-ocean billfish exchange around the Cape of Good Hope exists, as austral summer temperatures extend the billfish distribution to the south (Penrith and Cram, 1974; Talbot and Penrith, 1962). Tag and recapture studies also have demonstrated blue marlin migrations from Atlantic to Indian (Anonymous, 1994), and Pacific to Indian oceans (Julian Pepperrell, personal communication, 1997).

Genetic studies have shown that the degree of intra-specific population structure varies among highly migratory fishes. Whole-molecule, mtDNA RFLP studies revealed no significant differences between inter-ocean samples of skipjack, yellowfin, or albacore tunas, although sample sizes were small (Graves and Dizon, 1989; Graves et al., 1984; Scoles and Graves, 1993). In contrast, global population structuring was detected for a number of highly migratory pelagic species by whole molecule mtDNA analysis, including several istiophorid billfishes (Graves and McDowell, 1995). Using mitochondrial control region sequences, inter-ocean structure was also detected for swordfish and bigeye tuna (Alvarado Bremer et al., 1997; Rosel and Block, 1996). Whereas gene exchange among tunas may occur frequently between oceans, gene exchange among billfish may be deterred more effectively by cold waters frequently occurring around the southern tip of Africa or by behavioral differences (Nakamura, 1985). Mitochondrial studies have detected significant population subdivision within-oceans for striped marlin (Graves and McDowell, 1994) and swordfish (Alvarado Bremer et al., 1996). Although a variety of molecular techniques and sample sizes characterized these studies, the difference in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

levels of population structure among the species suggests that variations in life history may be more pronounced among the species than is commonly perceived. Possible life history characteristics influencing the diversity of population structure among species include temperature tolerances, migratory capabilities, spawning site fidelity, population size, and geographic range. For the circumtropical blue marlin, highly divergent mtDNA clades were detected, one of which (the Atlantic clade) was present in the Atlantic at a frequency of approximately 40% (Finnerty and Block, 1992; Graves and McDowell, 1995). The presence of divergent clades suggested that blue marlin in the Atlantic and Pacific oceans may have been isolated during the Pleistocene (Graves and McDowell, 1995).

This dissertation will introduce a variety of nuclear molecular markers to the analysis of population structure and gene flow for the blue marlin within and between ocean basins. Nuclear markers were chosen to cover a wide range of mutation rates in order to provide a sensitive analysis of population structure at each level of divergence (temporal, intra-ocean, inter-ocean). The markers also will be evaluated for evidence of the Pleistocene isolation. In addition, the comparison of markers will allow an evaluation of the effects of mutation, drift, selection, and migration on the distribution of diversity for each marker class.

DISSERTATION OUTLINE

The second chapter introduces a re-analysis of previously reported allozyme (Morgan, 1992) and newly developed scnDNA markers to questions of blue marlin gene flow. The role of population level processes in determining the distribution of nuclear

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and mitochondrial patterns of diversity also is discussed. In the third chapter, hypervariable microsatellite markers are used to analyze population structure and gene flow, along with analysis of an unpublished, greatly expanded version of the whole-molecule, RFLP mtDNA data of Graves and McDowell (1995). The chapter also includes discussion of historical phylogeographic insights from mtDNA and microsatellite data. In the fourth chapter, markers are compared with respect to mutation rates and sensitivity to population structuring at the intra-ocean and inter-ocean levels. Finally, the role of evolutionary forces in affecting the distribution of molecular diversity among oceans for each marker class is evaluated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LITERATURE CITED

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38:1507-1514.

Alvarado Bremer, J. R., J. Mejuto, T. W. Greig, and B. Ely. 1996. Global population structure of the swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial DNA control region. J. Exper. Mar. Biol. Ecol. 197:295-310.

Alvarado Bremer, J. R., B. Stequert, N. W. Robertson, and B. Ely. 1997. Genetic evidence for inter-ocean subdivision of bigeye tuna (*Thunnus obesus* Lowe) populations. Submitted. :In Press.

Amos, W., and D. C. Rubinstzein. 1996. Microsatellites are subject to directional evolution. Nat. Genet. 12:13-14.

Amos, W., S. J. Sawcer, R. W. Feakes, and D. C. Rubinsztein. 1996. Microsatellites show mutational bias and heterozygote instability. Nat. Genet. 13:390-1.

Angers, B., and L. Bernatchez. 1997. Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. Mol. Biol. Evol. 14:230-238.

Ashley, C. T., and S. T. Warren. 1995. Trinucleotide repeat expansion and human disease. Annu. Rev. Genet. 29:703-728.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall Inc., New York.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. Saunders. 1987. Intraspecific Phylogeography: The mitochondrial bridge between population genetics and systematics. Annual Review of Ecology and Systematics. 18:489-522.

Beardsley, G. L. 1989. Trends in Fisheries for Billfishes in the Atlantic Ocean. Pp. 17-30 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part I. Fishery and stock synopses, data needs and management. National Coalition for Marine Conservation, Inc., Savannah, GA.

Beckmann, J. S., and J. L. Weber. 1992. Survey of human and rat microsatellites. Genomics. 12:627-631.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Brooker, A. L., D. Cook, P. Bentzen, J. M. Wright, and R. W. Doyle. 1994. Organization of microsatellites differs between mammals and cold-water teleost fishes. Can. J. Fish. Aquat. Sci. 51:1959-1966.

Brown, W. M., M. George Jr., and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences, USA. 76:1967-1971.

Cyr, E. C., J. M. Dean, I. Jehangeer, and M. Nallee. 1990. Age, growth, and reproduction of blue marlin and black marlin from the Indian ocean. Pp. 309-316 *i n* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, GA.

Dallas, J. F. 1992. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. Mammal. Genome. 5:32-38.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M. Valdes, A. M. Slatkin, and N. B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences, USA. 91:3166-3170.

Dietrich, W., H. Katz, S. E. Lincoln, H. Shin, J. Friedman, N. C. Dracopoli, and E. S. Lander. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. Genetics. 131:423-447.

Edwards. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics. 12:241-253.

Estoup, A., C. Tailliez, J. M. Cornuet, and M. Solignac. 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). Mol. Biol. Evol. 12:1074-84.

Finnerty, J. R., and B. A. Block. 1992. Direct sequencing of mitochondrial DNA detects highly divergent haplotypes in blue marlin *(Makaira nigricans).* Mol. Mar. Biol. Biotechnol. 1:206-214.

FitzSimmons, N. N., C. Moritz, and S. S. Moore. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Mol. Biol. Evol. 12:432-440.

Gall, J. G. 1981. Chromosome structure and the C-value paradox. J. Cell Biol. 91:3S-14S.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Goldstein, D. B., and A. G. Clark. 1995. Microsatellite variation in North-American populations of Drosophila melanogaster. Nucleic Acids. Res. 23:3882-3886.

Goldstein, D. B., and D. D. Pollock. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic interference. J. Hered. 88:335-42.

Graves, J. E., and A. E. Dizon. 1989. Mitochondrial DNA sequence similarity of Atlantic and Pacific albacore tuna. Can. J. Fish. Aquat. Sci. 46:870-873.

Graves, J. E., S. D. Ferris, and A. E. Dizon. 1984. Close genetic similarity of Atlantic and Pacific skipjack tuna demonstrated with restriction endonuclease analysis of mitochondrial DNA. Marine Biology. 79:315-319.

Graves, J. E., and J. R. McDowell. 1994. Genetic analysis of striped marlin (*Tetrapturus audax)* population structure in the Pacific Ocean. Can. J. Fish. Aquat. Sci. 51:1762-1768.

Graves, J. E., and J. R. McDowell. 1995. Inter-ocean genetic divergence of Istiophorid billfishes. Marine Biology. 122:193-203.

Grimaldi, M.-C., and B. Crouau-Roy. 1997. Microsatellite homoplasy due to variable flanking sequences. J. Mol. Evol. 44:336-340.

Henderson, S. T., and T. D. Petes. 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 12:2749-2757.

Hillis, D. M., and C. Moritz. 1990. Molecular Systematics. Sinauer Associates, Inc., Sunderland, MA.

Hudson, T. J., M. Engelstein, M. K. Lee, E. C. Ho, M. J. Rubinfeld, C. P. Adams, D. E. Housman, and N. C. Dracopoli. 1992. Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. Genomics. 13:622-629.

ICCAT. 1994. Report for biennial period, 1994-1995. Pp. 13-28. , Madrid, Spain.

Jarne, P., and P. J. L. Lagoda. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. & Evol. 11:424-429.

Jin, L., C. Macaubas, J. Hallmayer, A. Kimura, and E. Mignot. 1996. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. Proceedings of the National Academy of Sciences, USA. 93:15285-15288.

Karl, S. A., and J. C. Avise. 1992. Balancing selection at allozyme loci in oysters: Implications from nuclear RFLP's. Science. 256:100-102.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Karl, S. A., and J. C. Avise. 1993. PCR-based assays of Mendelian polymorphisms from anonymous single-copy nuclear DNA: Techniques and applications for population genetics. Mol. Biol. Evol. 10:342-361.

Levinson, G., and G. A. Gutman. 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4:203-221.

Lewin, T. 1975. Units of transcription and translation: Sequence components of heterogenous nuclear RNA and messenger RNA. Cell. 4:77-93.

Lincoln, R. J., G. A. Boxshall, and P. F. Clark. 1982. A dictionary of ecology, evolution and systematics. Cambridge University Press

Mahtani, M. M., and H. F. Willard. 1993. A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Human Mol. Genet. 2:431-437.

Mather III, F. J., A. C. Jones, and G. L. Beardsley Jr. 1972. Migration and distribution of white marlin and blue marlin in the Atlantic Ocean. U. S. National Marine Fisheries Service Fishery Bulletin. 70:283-298.

McConnell, S. K., P. H. O'Reilly, L., J. M. Wright, and P. Bentzen. 1995. Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North American and European populations. Can. J. Fish. Aquat. Sci. 53:1863-1872.

Morgan, L. W. 1992. Allozyme analysis of billfish population structure. *in* Fisheries. College of William and Mary

Nakamura, I. 1985. Vol 5: Billfishes of the world: An annotated and illustrated catalogue of marlins, sailfishes, spearfishes, and swordfishes known to date. FAO Fish. Synop.

O'Connell, M., and J. M. Wright. 1997. Microsatellite DNA in fishes. Rev. Fish Biol. Fish. 7:331-363.

Orti, G., D. E. Pearse, and J. C. Avise. 1997. Phylogenetic assessment of length variation at a microsatellite locus. Proc. Nat. Acad. Sci. U. S. A. 94:10745-10749.

Ovenden, J. R. 1990. Mitochondrial DNA and marine stock assessment: A review. Aust. J. Mar. Freshwater Res. 41:835-853.

Palumbi, S. R., and C. S. Baker. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of Humpback Whales. Mol. Biol. Evol. 11:426-435.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Penrith, M. J., and D. L. Cram. 1974. The Cape of Good Hope: A hidden barrier to billfishes. Pp. 175-187 *in* R.S. Shomura and F. Williams, eds. Proceedings of the International Billfish Symposium, Kailua-Kona, HI, 9-12 August 1972. Part 2. Review and contributed papers. U.S. Dep. Commer., NOAA Tech. Rep. NMFS SSRF-675, Seattle, WA.

Pogson, G. H., K. A. Mesa, and R. G. Boutilier. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics. 139:375-385.

Prince, E. D., D. W. Lee, J. R. Zweifel, and E. B. Brothers. 1990. Ageing young Atlantic blue marlin (*Makaira nigricans*) from otolith microstructure (Summary Paper). Pp. 295-305 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, GA.

Rosel, P. E., and B. A. Block. 1996. Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. Marine Biology. 125:11-22.

Scoles, D. R., and J. E. Graves. 1993. Genetic analysis of the population structure of yellowfin tuna, *Thunnus albacares*, from the Pacific Ocean. U. S. National Marine Fisheries Service Fishery Bulletin. 91:690-698.

Scott, E. L., E. D. Prince, and C. D. Goodyear. 1990. History of the cooperative game fish tagging program in the Atlantic Ocean, Gulf of Mexico, and Caribbean Sea, 1954-1987. Am. Fish. Soc. Symp. 7:841-853.

Simmons, G. M., M. E. Kreitman, W. F. Quattlebaum, and N. Miyashita. 1989. Molecular analysis of the alleles of alcohol dehydrogenase along a cline in *Drosophila melanogaster*. I. Maine, North Carolina, and Florida. Evolution. 43:393-409.

Squire, J. L., and Z. Suzuki. 1990. Migration trends of striped marlin (*Tetrapturus audax*) in the Pacific Ocean. Pp. 67-81 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, Georgia.

Stephan, W., and S. Cho. 1994. Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. Genetics. 136:333-341.

Talbot, F. H., and M. J. Penrith. 1962. Tunnies and marlins of South Africa. Nature. 193:558-559.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Reseach. 17:6463-6471.

Tautz, D., M. Trick, and G. A. Dover. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature. 322:652-656.

Ueyanagi, S., R. S. Shomura, Y. Watanabe, and J. Squire. 1989. Trends in the Fisheries for Billfishes in the Pacific. Pp. 31-45 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part I. Fishery and stock synopses, data needs and management. National Coalition for Marine Conservation Inc., Savannah, GA.

Valdes, A. M., M. Slatkin, and N. B. Freimer. 1993. Allele frequencies at microsatellite loci: The stepwise mutation model revisited. Genetics. 133:737-749.

Voelker, R. A., H. E. Schaffer, and T. Muai. 1980. Spontaneous allozyme mutation in Drosophila melanogaster: rate of occurrence and nature of the mutations. Genetics. 94:961-968.

Waples, R. S. 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. J. Hered. In Press

Weber, J., and C. Wong. 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2:1123-1128.

Weber, J. L. 1990. Informativeness of human (dC-DA)n · (dG-dT)n polymorphisms. Genomics. 7:524-530.

Weissenbach, J., G. Gyapay, C. Dib, A. Vignal, J. Morissette, P. Milasseau, G. Vaysseix, and M. Lathrop. 1992. A second-generation linkage map of the human genome. Nature. 359:794-801.

Whittam, T. S., A. G. Clark, M. Stoneking, R. L. Cann, and A. C. Wilson. 1986. Allelic variation in human mitochondrial genes based on patterns of restriction site polymorphism. Proceedings of the National Academy of Sciences, USA. 83:9611-9615.

William, J., O. Ballard, and M. Kreitman. 1995. Is mitochondrial DNA a strictly neutral marker? Trends Ecol. Evol. 10:485-488.

Wright, S. 1978. Evolution and the genetics of populations. Vol. 4., Variability in and among natural populations. Univ. Chicago Press, Chicago.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Zouros, E., G. H. Pogson, D. I. Cook, and M. J. Dadswell. 1992. Apparent selective neutrality of mitochondrial DNA size variation: A test in the deep-sea scallop *Placopecten magellanicus*. Evolution. 46:1466-1476.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2. COMPARATIVE INTRASPECIFIC EVOLUTION OF MOLECULAR MARKERS IN THE BLUE MARLIN (*MAKAIRA NIGRICANS*): A HIERARCHICAL ANALYSIS OF VARIANCE REVEALED BY ALLOZYME, SINGLE-COPY NUCLEAR DNA, AND MITOCHONDRIAL DNA MARKERS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

INTRODUCTION

The few comparative studies of the evolution of molecular markers have provided conflicting conclusions regarding the relative influences of natural selection, mutation rate, sex-biased migration, and effective population size on the distribution of genetic variation within species. Notably discordant distributions of nuclear and mitochondrial DNA variation were identified within the American oyster (*Crassostrea virginica*), Atlantic cod (*Gadus morhua*), and green sea turtle (*Chelonia mydas*). In the American oyster, the geographic partitioning of genetic variation along the U. S. mid-Atlantic coast was much less pronounced in allozymes (Buroker, 1983) relative to single-copy nuclear DNA (scnDNA; Karl and Avise, 1992) or mitochondrial DNA markers (mtDNA; Reeb and Avise, 1990), suggesting that balancing selection affected the allozyme polymorphisms (Karl and Avise, 1992). A subsequent scnDNA study, however, detected no difference in genetic differentiation between scnDNA and allozyme loci in oysters taken from a similar geographic range (McDonald et al., 1996). The discrepancy between the two scnDNA studies has led some authors to speculate that the large variance in scnDNA estimates resulted from linkage of scnDNA loci to regions influenced by selective forces (FitzSimmons et al., 1997b).

In Atlantic cod, highly variable nuclear restriction fragment length polymorphisms (RFLPs; Pogson et al., 1995), single locus minisatellite (Galvin et al., 1995), and microsatellite DNA markers (Bentzen et al., 1996) exhibited significant heterogeneity among populations across the North Atlantic and across the Scotian shelf. These findings were in contrast to less polymorphic allozyme (Mork et al., 1985) and whole molecule mtDNA RFLP studies (Arnason et al., 1992; but see Dahle, 1991) in the same species. Differences in the amount of population structure detected within and among marker classes in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Atlantic cod have been attributed alternatively to balancing selection on allozyme polymorphisms (Pogson et al., 1995), directional selection on allozyme polymorphisms (Mork et al., 1985), directional selection on sites linked to scnDNA loci (Pogson et al., 1995), and increased discriminatory power in markers with higher mutation rates (Bentzen et al., 1996; Galvin et al., 1995).

In green sea turtle, global analysis of scnDNA polymorphisms suggested moderate gene flow, while analysis of mtDNA indicated very restricted female gene flow. It was suggested that enhanced male-mediated gene flow accounted for observed differences between the two classes of markers (Karl et al., 1992). Regional analyses of gene flow among northern and southern Great Barrier Reef populations of green turtle similarly revealed highly significant heterogeneity in mtDNA RFLPs but non-significant divergence in nuclear DNA markers. However, direct testing of differential natal homing tendencies among sexes revealed that males and females were equally philopatric to their natal sites. Marker discrepancy was thus attributed to opportunistic mating along migrational pathways (Fitzsimmons et al., 1997a), although differences in nuclear and mitochondrial effective population sizes also may have contributed to the difference. Clearly, more empirical and theoretical studies are needed to determine the relative importance of selection, mutation, sex-biased dispersal, and genetic drift on patterns of nuclear and mitochondrial genetic diversity.

The blue marlin (*Makaira nigricans*) was chosen for this comparative study of molecular markers to investigate the degree to which a pronounced mitochondrial divergence previously reported between Atlantic and Pacific oceans would be corroborated by nuclear loci. The blue marlin is circumtropically distributed, occurring in pelagic waters with surface temperatures greater than 22°C (Nakamura, 1985; Rivas, 1975). Individuals may live for more than twenty years and the species displays a sexual dimorphism in maximum size, with females commonly attaining three to four times the maximum size of males (Cyr et al., 1990). Tagging studies have revealed long distance movements of individuals across

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the Atlantic Ocean (Anonymous, 1994; Scott et al., 1990), from Pacific to Indian oceans (J. Pepperrell pers. comm., 1997), and from Atlantic to Indian oceans (Anonymous, 1994) (see Fig. 1). While long-distance movements are notable, the majority of reported tag recaptures occurred in the general vicinity of release, even after years at liberty (Anonymous, 1996). Spawning appears to occur over extensive geographic regions, based upon the distribution of larvae and adults with mature gonads (Matsumoto and Kazama, 1974; Strasburg, 1969). Evidence for discrete spawning cycles near island chains also has been reported (Hopper, 1990).

Despite the capacity for extensive gene flow within blue marlin, genetic evidence suggests limited exchange between ocean basins. Finnerty and Block (1992) sequenced a 612 bp region of the mitochondrial cytochrome *b* gene of 26 blue marlin, detecting two divergent clades of haplotypes. These clades occurred at significantly different frequencies in Atlantic and Pacific collections. Similar findings were reported in an analysis of whole-molecule mtDNA RFLPs (Graves and McDowell, 1995). The RFLP data were characterized by high levels of within-ocean variation and highly significant divergence between ocean samples. Haplotypes comprised two divergent clades, one of which was present only in the Atlantic at a frequency of approximately forty percent.

Existing mitochondrial data provide one perspective of blue marlin evolutionary history. MtDNA is a single, maternally inherited, haploid molecule. Independent loci may respond differently to a species' history due to the stochastic processes of genetic drift and mutation. Analysis of multiple regions of nuclear DNA is necessary for a robust estimation of demographic parameters within a species. This study compares gene flow estimates from analyses of allozyme, scnDNA, and mitochondrial DNA within and among Atlantic and Pacific samples of blue marlin. Allozyme and scnDNA nuclear markers are biparentally inherited and diploid. As a class of protein encoding loci, allozyme polymorphisms may be susceptible to selective constraints. In contrast, anonymously chosen scnDNA loci may consist of protein-coding and/or non-coding regions, where variation may or may not have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

functional significance. As with other vertebrates, fish nuclear DNA likely comprises a high percentage of non-coding sequences (Gall, 1981; Lewin, 1975). Because selection tends to act independently on individual loci, the influence of selection on an estimate of scnDNA gene flow is minimized when estimators from a number of such loci are averaged. ScnDNA loci also are potentially more variable than allozymes due to the presumed lack of selective constraints on non-coding polymorphisms and increased sensitivity of polymorphism detection. Thus, the two classes of nuclear markers used to estimate gene flow in this study differ in mutation rate and susceptibility to selective pressures. These nuclear estimates were compared to mitochondrial values derived from collections at the same locations, including many of the same individuals. Using these analyses, maternal spawning site fidelity, selection on allozyme polymorphism, sensitivity of each marker class to population structure, and stock identification of the blue marlin were investigated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MATERIAL AND METHODS

Blue Marlin Tissue Samples.---Blue marlin eye, liver, and heart tissues were obtained over several years in the Atlantic (U.S. mid-Atlantic coast and Jamaica) and Pacific oceans (Hawaii, Mexico, Ecuador, and eastern Australia; Table 1). Samples were obtained from fish landed at sport fishing tournaments, artisinal fisheries, and research cruises. Tissues were dissected within eight hours of death, chilled on ice, and subsequently frozen at -20°C. Samples were transported to the laboratory frozen on dry ice or a commercial ice substitute, and maintained at -70°C until analysis.

Allozymes.--- Horizontal starch gel electrophoresis was conducted following the protocols of Murphy et al. (1990) and Shaklee et al. (1990a). After initial screening, 33 enzyme systems were run on one of seven buffer systems resulting in 44 presumptive loci. Enzyme nomenclature follows Shaklee et al. (1990b). The following loci were surveyed: AAT-1, AAT-2, AATm, ACOH-1, ADH, ADA-1, CBP-1, CBP-2, CBP-3, CK-A, CK-B, DDH, EST-1, EST-2, EST-D, ENO, FUM, GAPDH-1, GAPDH-2, G3PDH-1, GCDH, G6PDH-1, GPI-1, GPI-2, IDH, IDDH, LDH-1, LDH-2, aMAN, MDH-1, MDH-2, MDHP-1, MDHP-2, MPI, PGK, ALA-MET, GLY-LEU, GLY-LEU-LEU, LEU-PEP, LEU-TYR, PGDH, PGM, SOD, and XDH. Details on the buffer system used for each enzyme, and tissue type surveyed are available from Morgan (1992).

Single-copy Nuclear Loci.--- The protocols of Karl and Avise (1993) were used for selection of scnDNA loci amplified by the polymerase chain reaction (PCR) with the following modifications and details. Total genomic DNA was isolated following the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

protocols of Sambrook et al. (1989). Genomic DNA was digested with *Pst I* (Gibco BRL) and fragments in the 500 - 2000 base pair (bp) size range were isolated onto DE-81 paper by electrophoresis (Whatman International Ltd., Maidstone, England). The plasmid vector Bluescript KS+ (Stratagene, La Jolla, CA) was digested with *Pst I*, phosphorylated using alkaline phosphatase (CIAP, Stratagene, La Jolla, CA), and ligated to genomic fragments with T4 DNA ligase (Stratagene, La Jolla, CA). *Escherichia coli* DH5α competent cells were transformed with the ligation products (Sambrook et al., 1989). Recombinant plasmids were screened for insert size by digestion with *Pst I*. Copy number was determined through genomic Southern analysis (Southern, 1975). In this procedure, *Pst I*-digested genomic DNA was probed with plasmid vectors containing cloned DNA fragments that were labeled using the BioNick labeling system (Gibco BRL, Gaithersburg MD). Hybridization and detection reactions were performed following manufacturer's protocols for BluGene Nonradioactive Nucleic Acid Detection System kits (Gibco BRL, Gaithersburg MD). The number of bands that appeared on the membrane approximated the copy number of the cloned fragment (Southern, 1975). The flanking 200 - 400 bp of single and low copy clones were sequenced using the Sanger dideoxy chain termination method (Sanger et al., 1977), with Sequenase Version 2.0 kits (U.S. Biochemical, Cleveland, OH). Primers were designed with the assistance of the computer program PC/GENE (Bairoch, 1989), and were optimized for a length of approximately 22 bp, 45 - 60% GC content, no more than four single base repeats, a maximum of four self-complementary base pairs, and a maximum of three primer-primer complementary base pairs. GenBank was searched for matches between published sequences and nuclear sequences obtained in the study. Open reading frames were searched for with PC/GENE (Bairoch, 1989), using the method of Fickett (1982).

The polymerase chain reaction (PCR) was used to amplify target fragments from genomic DNA. Template nuclear DNA was obtained from the nuclear band of cesium

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

chloride/ethidium bromide density gradient mtDNA purifications (Lansman et al., 1981). After dialysis, PCR was performed using the BRL PCR Reagent Systems (Gibco BRL), with a final concentration of 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 1 uM ea. primer, approximately 50 ng template DNA and 1.125 units *Taq* polymerase in 50 ul total volume. Typical cycling conditions included an initial denaturation of 5 min at 95°C, followed by 40 cycles of 1 min at 55°C, 1 min at 68°C, and 1 min at 95°C. Final extension was carried out for 7 min at 72°C. Annealing temperature, primer concentration, and template concentration were the main variables adjusted to optimize PCR conditions.

Ten to 20 blue marlin from different collections were screened for polymorphism by digestion of five to ten microliters of amplified DNA with 20 to 50 restriction enzymes. Digested products were electrophoresed on a 2.5% agarose gel, stained with ethidium bromide, and visualized under ultraviolet light. DNA samples were subsequently digested with a single restriction enzyme at each locus to generate genotypic frequencies (Table 2). Genotypes were scored for presence or absence of a restriction site.

Statistical Analyses.--- Two, non-parametric, exact-significance testing methods (exact θ_P significance tests and exact probability tests) were used to evaluate conformance to linkage and Hardy-Weinberg genotypic equilibria, and homogeneity of spatial and temporal distributions of allele frequencies. Unbiased estimators of exact significance probabilities were calculated by using the Markov chain algorithm of Guo and Thompson (1993) for tests of linkage equilibria, Hardy-Weinberg equilibria and population divergence, using the computer program GENEPOP (Raymond and Rousset, 1995). Exact probability tests of linkage equilibria establish the statistical independence of genotypes at different polymorphic loci and evaluate the null hypothesis that genotypes are distributed randomly among loci (Raymond and Rousset, 1995). Exact probability tests of Hardy-Weinberg equilibria

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

evaluate the null hypothesis that alleles are randomly distributed among individuals at each locus. Conformance of genotypic proportions to Hardy-Weinberg equilibria also is consistent with Mendelian inheritance and non-preferential amplification of alternative alleles by PCR. For these analyses, the Markov chain was initially set to 1000 dememorization steps, with a chain length of 50,000 steps. For p values approaching significance, the Markov chain length was increased to 500,000 steps. Type I error was controlled using the sequential Bonferroni method (Rice, 1989).

The Weir and Cockerham (1984) method of generating unbiased, hierarchical F-statistics was chosen to analyze patterns of genetic diversity within and between populations. This method incorporates explicit corrections for sample size and number of populations surveyed, facilitating comparison of different marker classes. The method also accommodates many types of genetic data (Excoffier et al., 1992) and intuitively partitions variance into components representing each level of organization. A four-level, hierarchical analysis of variance was performed as described in Weir (1996), using the computer program Genetic Data Analysis (GDA; Lewis and Zaykin, 1997). The hierarchical levels are outlined in Table 1. Description of variance components (σ^2) and F-statistics (θ) followed notations of Weir (1996), where subscripts refer to the following effects: P, populations (oceans); S, subpopulations within populations (geographic locations within oceans), SS, subsubpopulations within subpopulations (years within geographic locations); I, individuals within subsubpopulations; G, alleles within individuals. F_{IS} or f, the local inbreeding coefficient (Weir and Cockerham, 1984), was calculated with the program GENEPOP. The contribution to the overall hierarchical analysis of years within each geographic location and geographic locations within each ocean was subsequently examined. 95% confidence intervals (CI) of θ values were estimated from 15,000 bootstrap replicates over loci, performed using the computer program FSTAT (Goudet, 1995). Exact probability tests and exact θ_P permutation procedures were used to determine significance of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

population subdivision within a hierarchical level. Exact probability tests for population differentiation evaluated the null hypothesis that alleles were distributed randomly among populations, performed as described in Raymond and Rousset (1995). The exact θ_P permutation method tested the null hypothesis that θ was not greater than zero, evaluated through 1000 random permutations of alleles among populations and performed as described in Goudet (1995). For p values approaching significance, probability tests were adjusted as described for tests of equilibrium and permutation methods were adjusted by increasing the number of randomizations to 15,000. Pairwise tests for population differentiation were performed using the exact probability test.

Estimates of the number of migrants (N_em) between oceans were calculated from

$$[\theta_P = 1/(4N_em + 1)]$$

assuming an island model of migration (Wright, 1969). Neighbor-joining dendrograms (Saitou and Nei, 1987) were constructed for each polymorphic locus and over all four polymorphic loci (0.05 criterion) using Nei's (1978) unbiased genetic distance from GDA. Nei's (1972) genetic distance (D) also was calculated for comparison to previous works. Allele-frequency CI values were estimated as binomial variables assuming a normal distribution of allelic counts, taking into consideration deviations from Hardy-Weinberg equilibrium (Weir, 1996).

To relate the mitochondrial variation reported in Graves and McDowell (1995) to the nuclear data from this study, an analysis of molecular variance (AMOVA) was performed (Excoffier et al., 1992), generating divergence statistics equivalent to the θ values of Weir and Cockerham (1984). The AMOVA was performed (i) with haplotype distances input as squared Euclidean distances, calculated as the number of restriction site differences between

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

haplotypes; (ii) without haplotype distance information considered; and (iii) with haplotypes defined as belonging to either Atlantic or ubiquitous clades. Data were analyzed in hierarchical fashion, with Atlantic and Pacific populations consisting of two subpopulations each, as shown in Table 1 (Graves and McDowell, 1995). Significance of mitochondrial θ values was evaluated with a permutation procedure similar to that used in FSTAT, where alleles were permuted 1000 times among populations and the probability of obtaining a θ value as great or greater than the observed was determined. Since the θ_S level of divergence for allozyme and mtDNA RFLP data represents a mix of geographic and temporal samples (Table 1), the focus of the analysis for these markers was on the inter-oceanic (θ_P) level of variance. Estimates of the number of female migrants $N_e m_f$ between oceans were calculated from

$$[\theta_P = 1/(2N_e m_f + 1)].$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RESULTS

Allozymes.--- A total of 107 individuals from one Pacific and three Atlantic samples was screened for allozyme variation (Table 1). Four of 44 (9%) allozyme loci were polymorphic at the 0.05 criterion. The resulting average heterozygosity for the four polymorphic loci was 0.30, and over all 44 loci, an average heterozygosity of 0.0285 was calculated. Significant deviations from Hardy-Weinberg equilibrium were not detected at any polymorphic allozyme locus ($p > 0.08$; Fig. 2). θ_P values for ACOH, ADH, G3PDH, and IDDH loci were 0.107, 0.146, 0.05, and 0.012, respectively, with a 95% bootstrap CI over these four loci ranging from 0.011 to 0.138. Significant allele frequency differences were detected between ocean samples for ADH ($p < 0.001$), ACOH ($p < 0.001$), and G3PDH ($p = 0.011$; Tables 3, 4). Based on the four polymorphic loci, an average genetic distance of 0.038 separated the Atlantic and Pacific populations (Table 4), and over all 44 loci an inter-ocean D of 0.003 was calculated.

ScnDNA.--- A total of 145 plasmid DNA clones with insert sizes ranging from 500 to 2000 bp was generated. Eighteen randomly selected clones were screened for copy number by using genomic Southern blot analysis. Of these, thirteen were estimated to be single-copy, three low-copy, and two high-copy clones. Six of thirteen single-copy clones were chosen for partial sequencing. Primers were developed from five of the sequenced clones, and those producing a single, reliable, polymorphic amplification product were selected for further analysis (Appendix A). In addition, three primer pairs (WM08, WM13, WM84) from a white marlin (*Tetrapturus albidus*) mini DNA library of plasmid clones also were tested on blue marlin DNA. Six of these eight primer pairs reliably produced a single PCR product in blue marlin (Table 2). Four of these loci proved sufficiently polymorphic (0.05

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

criterion) for population genetic analysis (BM47, BM81, BM32, WM08). A total of 457 individuals from eight Pacific and seven Atlantic samples was screened for polymorphism at these four loci. A single restriction site polymorphism (RSP) was surveyed at each locus. Blast (Altschul et al., 1990) searches of the GenBank database for a match of the partial scnDNA sequences to published sequences were non-significant (smallest sum probability $p > 0.084$). No open reading frame larger than 200 bp was detected.

One of four scnDNA markers (BM32), exhibited a significant deviation from Hardy-Weinberg equilibrium for three of the fifteen populations surveyed (Fig. 2). To test for preferential amplification of alleles at this locus, PCR conditions and primer composition were varied (Hare et al., 1996). Genotypic scoring did not vary over a broad (45-65 $^{\circ}$C) range of PCR annealing temperatures but did vary upon primer redesign. Amplifications were performed with a redesigned reverse primer BM32R-2 in combination with the previous BM32F primer. There was a substantial shift in deviation from Hardy-Weinberg equilibrium upon re-amplification of all samples, from a strong heterozygote deficiency (mean $f = 0.37$) towards equilibrium (mean $f = 0.01$; Fig. 2). After adjustment, significant deviations from the null hypothesis ($f = 0$) were not detected. A slight heterozygote excess over all scnDNA loci (a negative σ^2_I) was detected (Fig. 3), due in large part to the negative f at the WM08 locus. Deviations from linkage equilibrium were non-significant after corrections for multiple tests ($p > 0.002$).

Within-population variation (among genes), designated σ^2_G, represented the largest proportion of overall variation at each locus (Fig. 3). A negligible fraction of overall variation was attributed to temporal (σ^2_{SS}) and intra-ocean (σ^2_S) components of variance. Significant allele-frequency heterogeneity was not detected at any scnDNA locus among temporal samples from Hawaii, Mexico, Australia, Jamaica, or the continental USA (Tables 3, 5). Tests of temporal variation were most powerful for Jamaican and Hawaiian comparisons, where three to five years of samples were compared at each location and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sample sizes averaged 34 individuals per year (Table 1). Tests for population differences using the exact θ_P and exact probability tests, were in close agreement. A minor inconsistency between the two tests occurred at a borderline p value in the inter-annual comparisons of Jamaican samples at the WM08 locus (Table 5). The permutation test of θ_{SS} equals zero was significant ($p = 0.009$), whereas the probability test yielded a non-significant p value after correction for multiple tests ($p = 0.012$; initial $\alpha = 0.01$). The pairwise comparison between 1993 and 1994 samples from Jamaica at the WM08 locus was significant at the 0.01 level ($p = 0.005$). Inclusion of negative values in the 95% CI on all θ_{SS} values supported non-significance (Weir, 1996).

Significant intra-ocean heterogeneity was not detected among samples within Pacific or North Atlantic oceans. θ_S estimates were non-significant at all loci and their confidence intervals included negative values (Table 5). No pairwise comparisons were significant among populations within ocean basins following corrections for multiple tests (initial α = 0.00045). Neighbor-joining trees generated from each locus and over all scnDNA loci revealed a lack of structure within ocean basins, as temporal population replicates were intermingled among geographic locations (Fig. 4). A single, non-significant ($p = 0.14$) clustering of U.S. populations occurred at the BM32-2 locus (Fig. 4).

Inter-oceanic differentiation (σ^2_P) formed a substantial component (up to 12%) of overall variance (Fig. 3). Three scnDNA loci (BM81, BM32-2, and WM08) demonstrated significant allele frequency differences between samples from the Atlantic and Pacific oceans (Table 4). θ_P values between the two oceans for BM47, BM81, BM32-2, and WM08 were -0.003, 0.052, 0.102 and 0.123, respectively, with a 95% CI over all loci ranging from 0.038 to 0.11 (Table 4). Atlantic and Pacific populations were separated by an average D of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

0.055 (Table 4; Fig. 4). An average nuclear θ_P of 0.08 was calculated over all polymorphic allozyme and scnDNA markers, with 95% bootstrap CI values ranging from 0.042 to 0.111. Corresponding N_em values were 2 to 5.7 migrant individuals per generation (Table 4). Populations clustered into distinct Atlantic and Pacific clades for the two most divergent loci (BM32-2 and WM08) and over all loci (Fig. 4), illustrating the degree of relatedness within relative to between ocean basins. The BM81 locus was discriminatory to a lesser extent, with intermingling of three Atlantic and two Pacific populations. None of 56 inter-ocean pairwise comparisons of allele frequencies was significant for BM47, nine of 56 were significant at the 0.01 level for BM81, 24 of 56 were significant at the 0.01 level for BM32-2, and 25 of 56 were significant at the 0.01 level for WM08. Fifteen of 224 inter-ocean pairwise comparisons were significant after corrections for multiple tests (Fig. 5).

MtDNA.---Reanalysis of the blue marlin mtDNA RFLP data from Graves and McDowell (1995) included 114 individuals from two Pacific and two Atlantic samples. When haplotype distances were input as squared Euclidean distances, the mtDNA AMOVA θ_P was calculated as 0.25, which represents 1.5 inter-ocean female migrants per generation (N_em_f; Table 6). Without haplotype distance information considered, an AMOVA θ_P of 0.06 was calculated ($N_em_f = 7.8$), and when haplotypes were designated as belonging to either ubiquitous or Atlantic clades (Graves and McDowell, 1995), the θ_P rose to 0.39 ($N_em_f = 0.85$). All θ_P values were highly significant ($p < 0.001$).

To investigate whether individuals belonging to the two highly divergent mtDNA clades comprised separate nuclear gene pools, blue marlin were tested for scnDNA allele frequency divergence between Atlantic individuals belonging to the two mtDNA clades. A total of 226 individuals from the Atlantic Ocean was assigned to either Atlantic or ubiquitous clades based on whole molecule RFLP data (J. Graves, unpublished data). Allele frequency

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

differences were not detected at any of the four scnDNA loci in pairwise comparisons between Atlantic blue marlin belonging to the two mitochondrial clades (Table 7).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DISCUSSION

Allozymes.--- The level of allozyme variation and inter-populational divergence revealed by this investigation falls within the range of other studies of pelagic fishes. Shaklee et al. (1983) surveyed 35 allozyme loci in 95 blue marlin from Hawaii, detecting a level of variation somewhat higher than that exhibited in this study when considering heterozygosity over all loci ($H = 0.061$ vs $H = 0.029$) and percent polymorphic loci ($P =$ 20% vs $P = 9\%$). However, heterozygosity over polymorphic loci ($H = 0.29$ vs $H = 0.30$) was similar, and the current study included a different set of allozyme loci.

The magnitude of genetic divergence between Atlantic and Pacific blue marlin ($D =$ 0.055, $\theta_P = 0.085$) was consistent with levels of divergence reported among geographically distant samples of several other broadly distributed marine fishes. In an allozyme analysis Rosenblatt and Waples (1986) reported little divergence between widely dispersed populations of 12 marine fish species. Non-significant trans-Pacific and inter-oceanic *D*s ranged from less than 0.01 to 0.06. A trans-Pacific F_{ST} of 0.04 for milkfish (*Chanos chanos*) populations was revealed in another study (Winans, 1980). Studies of skipjack (*Katsuwonus pelamis*) and yellowfin (*Thunnus albacares*) tunas revealed little trans-Pacific or inter-oceanic population structure based on blood type, serum esterase, transferrin (Fujino, 1970; Suzuki, 1962) allozyme, or mtDNA polymorphisms (Ward et al., 1994).

Single-copy Nuclear DNA.--- The majority of randomly selected genomic DNA fragments were single-copy and demonstrated suitable levels of polymorphism for assessment of nuclear gene flow between distant geographic areas. The four polymorphic scnDNA loci used in population screening were shown to be statistically independent, non-coding DNA regions whose terminal sequences do not match any known sequences. After

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

redesign of BM32 primers, the distribution of genotypes in all populations conformed to expectations of Hardy-Weinberg equilibrium, consistent with Mendelian inheritance. In initial analysis of the BM32 locus, several populations demonstrated significant deviations from Hardy-Weinberg equilibrium. Although a number of population genetic factors could contribute to heterozygote deficiency, the isolated nature of this phenomenon suggested that PCR-related artifacts were responsible for the deviations (Hare et al., 1996). Relative to the other scnDNA loci, BM32 demonstrated a high level of nucleotide sequence diversity. It is likely that intraspecific variation within priming sites had resulted in preferential amplification of alleles.

Population Genetic Implications of scnDNA Loci.--- It is important to estimate the magnitude of allele frequency variance among temporally replicated samples at a location before assigning significance to a divergence estimate between locations. Significant temporal heterogeneity was not detected among samples at any scnDNA locus. Tests included five consecutive years of samples from Jamaica and three years of samples from Hawaii. Both are areas of intense spawning activity (De Sylva, 1974; Hopper, 1990). Each blue marlin sample likely comprised a number of year classes, since individuals reach maturity within two to three years and can survive for longer than twenty (Cyr et al., 1990; Hopper, 1990; Prince and Brown, 1994; Rivas, 1975). Over a period of three to five years, however, a population is likely to experience a substantial shift in the composition of year classes. Temporal stability of allele frequencies observed in this study is consistent with recruitment from a single gene pool. Temporal stability indicates that frequencies are not heavily skewed by random sampling associated with either the formation of gametes among individuals, or the physical sampling of a population by an investigator (Weir, 1996).

Analysis of scnDNA loci revealed significant inter-ocean but non-significant intra-ocean divergence. The hypothesis of a single nuclear gene pool within either the Pacific or Atlantic oceans was supported by non-significant θ_S values, non-significant intra-ocean

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pairwise comparisons, and mixing of temporally replicated samples among geographic locations upon phenetic clustering of populations by genetic distance. In contrast to within-ocean analysis, significant θ_P values at three of four scnDNA loci, significant inter-ocean pairwise comparisons, and discrete clustering of populations into oceans at BM32-2, WM08, and overall neighbor-joining trees indicate that the two oceans represent distinct populations. The distribution of significant pairwise comparisons is influenced both by genetic distance separating the populations, and sample sizes. The higher number of significant pairwise comparisons among pooled samples (11 of 24 significant at the 0.01 level) demonstrates the increased statistical power accompanying larger sample sizes. Although significant, the level of inter-ocean differentiation was low, suggesting some nuclear gene flow between areas. An average of 2.67 genetically effective migrants between oceans per generation was estimated (range = 2.0 to 6.3 migrants). These values are sufficiently greater than the single migrant sufficient to prevent accumulation of fixed genetic differentiation between populations (Allendorf and Phelps, 1981).

Concordance of scnDNA and allozyme markers.--- In general, scnDNA and allozyme markers revealed similar levels of genetic variation and population structure within and between Atlantic and Pacific populations of blue marlin. Polymorphic allozyme and scnDNA loci revealed similar levels of average heterozygosity (H = 0.30 and 0.39, respectively), although a higher percentage of scnDNA loci (67% versus 9%) were polymorphic. Significant differences in allele frequency were detected at three of the four polymorphic scnDNA and three of four polymorphic allozyme loci. Average θ_P values, based on four polymorphic allozyme loci, were similar to those of the four polymorphic scnDNA loci, with broadly overlapping 95% bootstrap CI values. When considered together, 8% of overall nuclear genetic variation was attributed to the inter-ocean component of variance, with 95% confidence intervals ranging between 4.2% and 11.1%.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The range in levels of inter-oceanic divergence exhibited by allozyme and scnDNA loci presumably illustrates the stochastic and independent process of genetic drift, and underscores the need for examination of a number of loci for estimating population parameters. Goudet (1995) has suggested that bootstrap CI values are most accurate with at least five informative loci. Although the combined nuclear estimate substantially increased the precision of the allozyme bootstrap CI values, the scnDNA estimate was only slightly improved. The difference in CI shift between the allozyme and scnDNA markers reflects the larger degree of variation in θ_P estimates from allozyme loci.

Other comparative studies have suggested widespread balancing selection on allozymes, based on substantially lower allozyme divergence levels relative to nDNA values (Pogson et al., 1995), or both nDNA and mtDNA markers (Karl and Avise, 1992). The broad overlap of divergence levels from allozyme and scnDNA nuclear markers in this study suggests that both sets of allele frequency distributions represent neutral gene flow in blue marlin. A similar concordance of polymorphic allozyme and scnDNA molecular markers was observed in one study of American oyster (*Crassostrea virginica*; McDonald et al., 1996).

Discordance of Nuclear and Mitochondrial Markers.--- The degree of variation and inter-ocean divergence revealed by both classes of nuclear markers was smaller than that revealed by mtDNA (Graves and McDowell, 1995). RFLP analysis of mtDNA (employing 11 enzymes) revealed 38 alleles in 114 individuals and a nucleon diversity (*h*) of 0.85. In contrast, only a single restriction site polymorphism was detected for three of four scnDNA loci when surveyed with 20 to 50 restriction enzymes. Phylogenetic analysis of the mtDNA RFLP data divided haplotypes into Atlantic and ubiquitous clades separated by 1.2% nucleotide sequence divergence. It has been suggested that clades were formed during the Pleistocene when tropical zones were compressed, restricting inter-ocean gene flow for a number of tropical species (Graves and McDowell, 1995). When converted into θ_P

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

estimates, mitochondrial divergence levels between oceans were highly significant ($p < 0.001$), and substantially higher (0.25) than those observed in the nuclear markers (0.08). With haplotype divergence removed as a distance measure between alleles, the mtDNA divergence estimate from AMOVA was reduced (0.06), due to the relatively small sample sizes for the number of alleles present.

Evolutionary Forces.--- Mutation rate, effective population size, natural selection, and sex-biased migration could contribute to the observed difference between mitochondrial and nuclear estimates of genetic divergence. The mitochondrial genome has an effective population size one-fourth that of the nuclear genome, effectively increasing expected divergence in mtDNA relative to nuclear loci. The mtDNA θ_P (0.25) is greater than the nuclear average (0.08) by a factor of three, suggesting that effective population size considerations alone are sufficient to explain the observed difference. When converted into numbers of migrants and corrected for mode of inheritance, a mitochondrial $N_e m_f$ of 1.50, and average nuclear $N_e m$ of 2.87 were obtained. These gene flow estimates are similar when one considers that the number of migrating females is expected to be half the total number of migrants. Contrary to expectations of genetic drift, greater nuclear than mitochondrial divergence was reported among populations of the broad whitefish (*Coregonus nasus*) in North Alaska (Patton et al., 1997), cod (*Gadus morhua*) in the North Atlantic (Arnason et al., 1991; Bentzen et al., 1996; Galvin et al., 1995; Pogson et al., 1995), and brook trout (*Salvelinus fontinalis*) in eastern Canada (Angers et al., 1995; Jones et al., 1996). These differences were attributed largely to the high mutation rate and sensitivity of nuclear markers used relative to those of the mitochondrial markers. Among stable populations at equilibrium one would expect mitochondrial markers, on average, to reveal greater levels of divergence than nuclear markers due to genetic drift.

Under an infinite alleles model of evolution, markers with higher mutation rates generate greater diversity, leading to more rapid divergence of isolated populations. The high

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mitochondrial mutation rate in the blue marlin generated distinct clades presumably during a historic condition of oceanic isolation. The AMOVA divergence estimate that incorporated mutational events between haplotypes through Euclidean distances (0.25) was substantially smaller than the estimate based on clade frequency difference alone (0.39). This suggests that mixing of clades between oceans, not degree of divergence between clades or more recent mutational events, contributes to the observed level of mitochondrial differentiation.

Selection also could prevent divergence of nuclear allele frequencies during a time of allopatry. However, it is difficult to imagine the same selective pressures acting on both the presumably independent, non-coding, randomly selected nuclear DNA loci and the protein encoded allozyme loci. A final factor contributing to lower divergence in nuclear markers in mtDNA could be sex-specific migration, i.e., preferential (and occasional) exchange of males between oceans. For blue marlin, sexual dimorphism in maximum size (females attain a maximum size three to four times larger than males; De Sylva, 1974) presents the possibility for sex-specific migrational activity. Evidence for sex-dependent migratory behavior includes predictable sex ratio shifts in spawning regions (Erdman, 1968; Hopper, 1990; Nakamura, 1949; Nakamura et al., 1953) and differences in sex composition between regions in the Pacific Ocean (Kume and Joseph, 1969). A paucity of tag returns and lack of sex identification in tagged blue marlin precludes inferences regarding greater male or female dispersal. It is worth noting, however, that many migratory blue marlin appear to be large females (E. Prince pers. comm., 1997).

In a number of studies, greater divergence in mtDNA markers (relative to nuclear DNA) has been explained by invoking complex behavioral patterns, including male-mediated dispersal. These include studies of humpback whales in the North Pacific (Palumbi and Baker, 1994) and North Atlantic (Larsen et al., 1996), three sympatric forms of brown trout (*Salmo trutta*; Ferguson et al., 1995), residential and anadromous forms of brown trout (Ferguson et al., 1995), global populations of green turtle (*Chelonia mydas*; Karl et al., 1992), and green turtle populations on a regional scale (Fitzsimmons et al., 1997a;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

FitzSimmons et al., 1997b). The hypothesis of male-mediated gene flow in these studies is supported by higher mitochondrial than nuclear divergence, and in some cases by behavioral observations. While greater mitochondrial divergence suggests the possibility of male-mediated gene flow, the role of genetic drift effectively confounds the ability to identify sexually dimorphic behavioral patterns with this data alone.

Sequence-level analysis of the quickly mutating BM32-2 or microsatellite markers may help reveal the history of nuclear gene flow during the Pleistocene. However, male-mediated dispersal and/or larger effective population size may have prevented fixed differences from accumulating in the nuclear genome during the presumed period of allopatry (Graves and McDowell, 1995). If estimates of inter-ocean divergence derived from quickly mutating nuclear loci are comparable to estimates based on mtDNA, the discordance observed in this study may be attributed to lower mutability of scnDNA and allozymes. If persistently lower divergence in all nuclear marker classes is observed, both drift and male-mediated gene flow remain plausible hypotheses to be tested.

The absence of haplotypes from the Atlantic mtDNA clade in the Pacific suggests a predominantly one-way migratory route of individuals from the Pacific into the Atlantic. If such one-way migration occurs, there could be a detectable difference in nuclear allele frequency within the Atlantic between individuals belonging to the two mitochondrial clades. However, tests at each of the four scnDNA loci failed to detect significant differences between these two groups, indicating that random mating between clades in the Atlantic has mixed nuclear genes of individuals belonging to the two mitochondrial clades.

Polymorphic allozyme and scnDNA markers yielded concordant estimates of average heterozygosity and population subdivision, consistent with the hypothesis that variation at both markers is selectively neutral. Genetic drift alone is sufficient to explain the lower amount of nuclear than mitochondrial inter-ocean divergence. In addition, the role of mutation rate is further examine in the next chapter, with the addition of hyper-variable nuclear loci.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LITERATURE CITED

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38:1507-1514.

Altschul, S. F., G. W. Miller, E. W. Myers and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Angers, B., L. Bernatchez, A. Angers, and L. Desgroseillers. 1995. Specific microsatellite loci for brook charr reveal strong population subdivision on a microgeographic scale. J. Fish Biology 47 (Supplement A):177-185.

Anonymous. 1994. Cooperative Game Fish Tagging Program Annual Newsletter: 1992. Southeast Fisheries Science Center, Miami, FL.

-----. 1996. 1996 Billfish Newsletter. Southwest Fisheries Science Center, La Jolla, CA.

Arnason, E., S. Palsson, and A. Arason. 1992. Gene flow and lack of population differentiation in Atlantic cod, *Gadus morhua* L., from Iceland, and comparison of cod from Norway and Newfoundland. J. Fish Biol. 40:751-770.

Bairoch, A. 1989. PC/GENE. Intelligenetics, Inc., University of Geneva, Switzerland.

Bartlett, M. R., and R. L. Haedrich. 1968. Newston nets and South Atlantic larval blue marlin (*Makaira nigricans*). Copeia 3:469-474.

Bayley, R. E., and E. D. Prince. 1994. A review of tag release and recapture files for Istiophoridae from the southeast fisheries center's cooperative game fish tagging program, 1954 to present. Pp. 527-548 *in* ICCAT, ed. Collective volume of scientific papers. Vol. XLI. Report of the second ICCAT billfish workshop. ICCAT, Madrid, Spain.

Bentzen, P., C. T. Taggart, D. E. Ruzzante, and D. Cook. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 53:2706-2721.

Buroker, N. E. 1983. Population genetic of the American oyster *Crassostrea virginica* along the Atlantic coast and the Gulf of Mexico. Mar. Biol. 75:99-112.

Cyr, E. C., J. M. Dean, I. Jehangeer, and M. Nallee. 1990. Age, growth, and reproduction of blue marlin and black marlin from the Indian ocean. Pp. 309-316 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, GA.

Dahle, G. 1991. Cod, *Gadus morhua* L., populations identified by mitochondrial DNA. J. Fish Biol. 38:295-303.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

De Sylva, D. P. 1974. Life history of the Atlantic blue marlin, *Makaira nigricans*, with special reference to Jamaican waters. P. 80. *in* (Shomura and Williams, 1974).

Erdman, D. S. 1968. Spawning cycle, sex ratio, and weights of blue marlin off Puerto Rico and the Virgin Islands. Trans. Am. Fish. Soc. 97:131-137.

Eschmeyer, W. N. 1968. Four advanced larval specimens of the blue marlin, *Makaira nigricans*, from the western Atlantic Ocean. Copeia 2:414-417.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Ferguson, A., J. B. Taggart, P. A. Prodohl, O. McMeel, C. Thompson, C. Stone, P. McGinnity, and R. A. Hynes. 1995. The application of molecular markers to the study and conservation of fish populations, with special reference to *Salmo*. J. Fish Biol. 47 (Supplement A):103-126.

Fickett, J. W. 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. 10:5303-5318.

Finnerty, J. R., and B. A. Block. 1992. Direct sequencing of mitochondrial DNA detects highly divergent haplotypes in blue marlin *(Makaira nigricans)*. Mol. Mar. Biol. Biotech. 1:206-214.

Fitzsimmons, N. N., C. L. Limpus, J. A. Norman, A. R. Goldizen, and J. D. Miller. 1997a. Philopatry of male marine turtles inferred from mitochondrial DNA markers. Proc. Nat. Acad. Sci., U.S.A. 94:8912-8917.

FitzSimmons, N. N., C. Moritz, C. J. Limpus, L. Pope, and R. Prince. 1997b. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. Genetics 147:1843-54.

Fujino, K. 1970. Immunological and biochemical genetics of tunas. Trans. Am. Fish. Soc. 1:152-178.

Gall, J. G. 1981. Chromosome structure and the C-value paradox. J. Cell Biol. 91:3S-14S.

Galvin, P., D. McGregor, and T. Cross. 1995. A single locus PCR amplified minisatellite region as hypervariable genetic marker in gadoid species. Aquaculture 137:31-40.

Goudet, J. 1995. FSTAT version 1.2: a computer program to calculate F-statistics. J. Hered. 86:485-486.

Graves, J. E., and J. R. McDowell. 1995. Inter-ocean genetic divergence of Istiophorid billfishes. Mar. Biol. 122:193-203.

Guo, S. W., and E. A. Thompson. 1993. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometics 48:361-372.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Hare, M. P., S. A. Karl, and J. C. Avise. 1996. Anonymous nuclear DNA markers in the American Oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. Mol. Biol. Evol. 13:334-345.

Hopper, C. N. 1990. Patterns of Pacific blue marlin reproduction in Hawaiian waters. Pp. 29-39 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, GA.

Howard, J. K., and S. Ueyanagi. 1965. Distribution and relative abundance of billfishes (Istiophoridae) of the Pacific Ocean. Stud. Trop. Oceanogr. 2:1-134.

Jones, M., W, D. Clay, and R. G. Danzmann. 1996. Consevation genetics of brook trout (*Salvelinus fontinalis*): population structuring in Fundy National Park, New Brunswick, and eastern Canada. Can. J. Fish. Aquat. Sci. 53:2776-2791.

Karl, S. A., and J. C. Avise. 1992. Balancing selection at allozyme loci in oysters: Implications from nuclear RFLP's. Science 256:100-102.

-----. 1993. PCR-based assays of Mendelian polymorphisms from anonymous single-copy nuclear DNA: Techniques and applications for population genetics. Mol. Biol. Evol. 10:342-361.

Karl, S. A., B. W. Bowen, and J. C. Avise. 1992. Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analysis of anonymous nuclear loci. Genetics 131:163-173.

Kume, S., Joseph, J. 1969. Size composition and sexual maturity of billfish caught by the Japanese longline fishery in the Pacific Ocean east of 120°W. Bull. Far Seas Fish. Res. Lab. (Shimizu) 2:115-162.

Lansman, R. A., R. O. Shade, J. F. Shapira, and J. C. Avise. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications. J. Mol. Evol. 17:214-226.

Larsen, A. H., J. Sigurjonsson, N. Oien, G. Vikingsson, and P. Palsboll. 1996. Populations genetic analysis of nuclear and mitochondrial loci in skin biopsies collected from central and northeastern North Atlantic humpback whales (*Megaptera novaeangliae*): population identity and migratory destinations. Proc. R. Soc. Lond. Ser. B Biol. Sci. 263:1611-1618.

Lewin, T. 1975. Units of transcription and translation: Sequence components of heterogenous nuclear RNA and messenger RNA. Cell. 4:77-93.

Lewis, P. O, and D. Zaykin. 1997. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0. Free distribution by authors over the internet from the GDA HOME PAGE at http://cheeunm.edu/gda/.

Matsumoto, W. M., and T. K. Kazama. 1974. Occurrence of young billfishes in the central Pacific Ocean. Pp. 238-251 *in* (Shomura and Williams, 1974).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

McDonald, J. H., B. C. Verrelli, and L. B. Geyer. 1996. Lack of geographic variation in anonymous nuclear polymorphisms in the American oyster, *Crassostrea virginica*. Mol. Biol. Evol. 13:1114-1118.

Morgan, L. W. 1992. Allozyme analysis of billfish population structure. Masters Thesis. College of William and Mary.

Mork, J., N. Ryman, G. Stahl, F. Utter, and G. Sunders. 1985. Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. Can. J. Fish. Aquat. Sci. 42:1580-1587.

Murphy, R. W., J. Sites J. W., D. G. Buth, and C. H. Haufler. 1990. Proteins: Isozyme Electrophoresis. Pp. 45-126. *in* D.M. Hillis and C. Moritz, eds. Molecular Systematics. Sinauer Associates Inc., Sunderland, MA.

Nakamura, H. 1949. The tunas and their fisheries. Pacific Oceanic Fishery Investigations, U.S. Fish and Wildlife Service.

Nakamura, H., Y. Tabuta, and S. Ueyanagi. 1953. Relation between the spawning season and the sex ratio of some fishes of the Family Istiophoridae. Contrib. Nankai Reg. Fish. Res. Lab. 1:1-8.

Nakamura, I. 1974. Some aspects of the systematics and distribution of billfishes. Pp. 45-53 *in* (Shomura and Williams, 1974).

-----. 1985. Vol 5: Billfishes of the world: An annotated and illustrated catalogue of marlins, sailfishes, spearfishes, and swordfishes known to date. FAO Fish. Synop.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.

-----. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nishikawa, Y.S., Ueyanagi, S., Kikawa, S. 1985. Average distribution of larvae of oceanic species of scombroid fishes, 1956 - 1981. Far Seas Fish. Res. Lab. Japan.

Palumbi, S. R., and C. S. Baker. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of Humpback Whales. Mol. Biol. Evol. 11:426-435.

Patton, J. C., B. J. Gallaway, R. G. Fechhelm, and M. A. Cronin. 1997. Genetic variation of microsatellite and mitochondrial DNA markers in broad whitefish (*Coregonus nasus*) in the Colville and Sagavanirktok rivers in northern Alaska. Can. J. Fish. Aquat. Sci. 54:1548-1556.

Pogson, G. H., K. A. Mesa, and R. G. Boutilier. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics 139:375-385.

Prince, E. D., and B. E. Brown. 1994. Historical development if the ICCAT enhanced research program for billfishes. Pp. 1-3 *in* ICCAT, ed. Collective volume of scientific papers. Vol. XLI. Report of the second ICCAT billfish workshop. ICCAT, Madrid, Spain.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. J. Hered. 86:248-249.

Reeb, C. A., and J. C. Avise. 1990. A genetic discontinuity in a continuously distributed species: Mitochondrial DNA in the American oyster, *Crassostrea virginica*. Genetics 124:397-406.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Rivas, L. R. 1975. Synopsis of biological data on blue marlin, *Makaira nigricans* (Lacepede), 1802. Pp. 1-27 *in* R.S. Shomura and F. Williams, eds. Proceedings of the International Billfish Symposium, Kailua-Kona, HI, 9-12 August 1972. Part 3. Species Synopses. U.S. Dep. Commer., NOAA Tech. Rep. NMFS SSRF-675, Seattle, WA.

Rosenblatt, R. H., and R. S. Waples. 1986. A genetic comparison of allopatric populations of shore fish species from the eastern and central Pacific ocean: dispersal or vicariance? Copeia :275-284.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstruction of phylogenetic trees. Mol. Biol. and Evol. 4:406-425.

Sambrook, J., Fritsch, E.F., Maniatus, T. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, USA.

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

Scott, E. L., E. D. Prince, and C. D. Goodyear. 1990. History of the cooperative game fish tagging program in the Atlantic Ocean, Gulf of Mexico, and Caribbean Sea, 1954-1987. Am. Fish. Soc. Symp. 7:841-853.

Shaklee, J. B., R. W. Brill, and R. Acerra. 1983. Biochemical genetics of Pacific blue marlin, *Makaira nigricans*, from Hawaiian waters. Fish. Bull., U.S. 81:85-90.

Shaklee, J. B., S. R. Phelps, and J. Salini. 1990a. Analysis of fish stock structure and mixed-stock fisheries by the electrophoretic characterization of allelic isozymes. Pp. 173-96 *in* D.H. Whitmore, ed. Electrophoretic and isoelectric focusing techniques in fisheries management. CRC Press, Boca Raton, FL.

Shaklee, J. B., F. R. Allendorf, D. C. Morizot, and G. S. Whitt. 1990b. Gene nomenclature for protein-coding loci in fish. Trans. Am. Fish. Soc. 119:2-15.

Shomura, R. S. and F. Williams, eds. 1974. Proceedings of the International Billfish Symposium, Kailua-Kona, HI, 9-12 August 1972. Part 2. Review and contributed papers. U.S. Dep. Commer., NOAA Tech. Rep. NMFS SSRF-675, Seattle, WA.

Southern, E. 1975. Detection of specific sequences among DNA fragmentsseparated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Strasburg, D.W. 1969. Billfishes of the central Pacific Ocean. U.S. Fish and Wildl. Serv. Circ. 311:11p.

Suzuki, A. 1962. On the blood types of yellowfin and bigeye tuna. Am. Nat. 96:239-246.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ueyanagi, S., S. Kikawa, M. Uto, and Y. Nishikawa. 1970. Distribution, spawning and relative abundance of billfishes in the Atlantic Ocean. Bull. Far Seas Fish. Res. Lab. (Shimizu) 3:15-55.

Ward, R. D., and P. M. Grewe. 1994. Appraisal of molecular genetic techniques in fisheries. Rev. Fish Biol. Fish. 4:300-325.

Ward, R. D., N. G. Elliott, P. M. Grewe, and A. J. Smolenski. 1994. Allozyme and mitochondrial DNA variation in yellowfin tuna (*Thunnus albacares*) from the Pacific Ocean. Mar. Biol. 118:531-539.

Weir, B. 1996. Genetic data analysis II. Methods for discrete population genetic data. Sinauer Associates, Inc., Sunderland, Massachusetts.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Winans, G. A. 1980. Geographic variation in the milkfish *Chanos chanos*. I Biochemical evidence. Evolution 34:558-574.

Wright, S. 1969 Evolution and the genetics of populations, Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1. Samples of blue marlin surveyed. For scnDNA, ocean, location, and year form components to a hierarchical analysis of molecular variance (AMOVA). 'N' refers to the number of individuals sampled.

Ocean	Location	Year	Abbreviation	ScnDNA N	Allozyme N	MtDNA[1] N
ATLANTIC						
	Puerto Rico					
		1990	PR90		12	
		1991	PR91		21	31
	Jamaica					
		1991	JM91	48	20	25
		1992	JM92	55		
		1993	JM93	42		
		1994	JM94	45		
		1995	JM95	24		
	Eastern USA					
	New Jersey	1994	NJ94	11		
	North Carolina	1992	NC92	12		
PACIFIC						
	Hawaii					
		1990	HI90		54	26
		1992	HI92	30		32
		1993	HI93	48		
		1994	HI94	25		
	Eastern Pacific					
	Mexico	1994	MX94	49		
		1995	MX95	24		
	Ecuador	1995	EC95	19		
	Australia					
		1991	AU91	10		
		1994	AU94	15		

[1] MtDNA samples from Graves and McDowell, 1995.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2. ScnDNA restriction endonuclease survey. Complete sequences were obtained for two blue marlin, a striped and white marlin at WM13. % Base pairs polymorphic was calculated as [100 (# sites polymorphic)/ (# base pairs surveyed)], assuming one base pair polymorphism per polymorphic restriction site. For BM32, the *Nci* I polymorphism was screened in population analyses.

Locus	Length (base pairs)	# Base Pairs Surveyed*	Polymorphic Enzymes	# Sites Polymorphic	% Base Pairs Polymorphic
BM47	1900	134	*Bcl* I	1	0.75
BM32	1900	140	*Msp* I/*Dra* I	8	5.7
			Dde I/*Hinc* II		
			Hinf I/*Nci* I		
BM81	1700	98	*Dra* I	1	1.02
WM08	1200	114	*Ban* II	1	0.88
WM13	380	380		0	0.00
WM84	900	74		0	0.00
Overall	7980	940		11	1.2

* Based on restriction digestion with a battery of 20 to 50 restriction enzymes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3. Allele frequencies, confidence intervals, and sample size per locus of Atlantic and Pacific blue marlin. Allele frequencies refer to frequency of the most common allele (A). All loci were diallelic. 'CI' refers to the confidence interval calculated following Weir (1995), not assuming Hardy-Weinberg equilibrium. '2N' refers to twice the number of individuals surveyed. Sample abbreviations follow Table 1.

ScnDNA	BM47			BM81			BM32-2			WM08		
Sample	A	CI	2N	A	CI	2N	A	CI	2N	A	CI	2N
HI92	0.91	+/- 0.07	(58)	0.54	+/- 0.16	(48)	0.67	+/- 0.12	(64)	0.45	+/- 0.13	(60)
HI93	0.94	+/- 0.05	(96)	0.61	+/- 0.11	(88)	0.57	+/- 0.10	(90)	0.50	+/- 0.10	(86)
HI94	0.82	+/- 0.13	(50)	0.68	+/- 0.13	(50)	0.48	+/- 0.11	(50)	0.56	+/- 0.15	(48)
Hawaii	0.90	+/- 0.04	(204)	0.61	+/- 0.08	(186)	0.58	+/- 0.07	(204)	0.50	+/- 0.07	(194)
MX94	0.95	+/- 0.04	(98)	0.60	+/- 0.08	(98)	0.58	+/- 0.10	(98)	0.42	+/- 0.09	(90)
MX95	0.92	+/- 0.08	(48)	0.72	+/- 0.10	(46)	0.60	+/- 0.14	(48)	0.52	+/- 0.16	(44)
EC95	0.92	+/- 0.09	(38)	0.50	+/- 0.17	(38)	0.53	+/- 0.18	(38)	0.50	+/- 0.16	(36)
E.Pac	0.94	+/- 0.04	(184)	0.61	+/- 0.06	(182)	0.58	+/- 0.07	(184)	0.47	+/- 0.07	(170)
AU91	1.00	+/- 0.00	(18)	0.71	+/- 0.20	(14)	0.50	+/- 0.24	(18)	0.42	+/- 0.16	(12)
AU94	0.97	+/- 0.07	(30)	0.57	+/- 0.24	(28)	0.59	+/- 0.22	(32)	0.43	+/- 0.17	(30)
Aus	0.98	+/- 0.04	(48)	0.62	+/- 0.17	(42)	0.56	+/- 0.15	(50)	0.43	+/- 0.12	(42)
JM91	0.91	+/- 0.06	(96)	0.81	+/- 0.10	(84)	0.35	+/- 0.10	(92)	0.79	+/- 0.09	(66)
JM92	0.94	+/- 0.04	(108)	0.81	+/- 0.07	(110)	0.39	+/- 0.10	(108)	0.67	+/- 0.09	(92)
JM93	0.89	+/- 0.07	(84)	0.72	+/- 0.10	(82)	0.38	+/- 0.10	(84)	0.83	+/- 0.08	(78)
JM94	0.96	+/- 0.04	(90)	0.67	+/- 0.10	(90)	0.32	+/- 0.10	(92)	0.62	+/- 0.11	(82)
JM95	0.85	+/- 0.10	(48)	0.81	+/- 0.10	(48)	0.31	+/- 0.13	(48)	0.79	+/- 0.10	(48)
Jamaica	0.92	+/- 0.03	(426)	0.76	+/- 0.04	(414)	0.35	+/- 0.05	(424)	0.73	+/- 0.05	(366)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3 (Continued). Allele frequencies, confidence intervals, and sample size per locus.

ScnDNA	BM47			BM81			BM32-2			WM08		
Sample	A	CI	2N	A	CI	2N	A	CI	2N	A	CI	2N
NC92	0.96	+/- 0.09	(24)	0.73	+/- 0.17	(22)	0.25	+/- 0.16	(24)	0.58	+/- 0.22	(24)
NJ94	0.91	+/- 0.13	(22)	0.73	+/- 0.22	(22)	0.23	+/- 0.22	(22)	0.80	+/- 0.17	(20)
U.S.	0.94	+/- 0.08	(46)	0.73	+/- 0.14	(44)	0.24	+/- 0.13	(46)	0.68	+/- 0.15	(44)

Allozymes:	ACOF			ADH			G3PD			IDDH		
Sample	A	CI	2N	A	CI	2N	A	CI	2N	A	CI	2N
Atlantic	0.89	+/- 0.06	(106)	0.85	+/- 0.08	(106)	0.77	+/- 0.09	(102)	0.86	+/- 0.07	(100)
Pacific	1.00	+/- 0.00	(108)	0.59	+/- 0.11	(102)	0.62	+/- 0.10	(102)	0.86	+/- 0.09	(86)

Table 4. Nuclear marker comparison: diversity and inter-ocean divergence. *F*-statistics (θ_P; Weir and Cockerham, 1984) and genetic distances (D_1, Nei, 1972; D_2 Nei, 1978) were based on polymorphic loci. *p* values represent the probability that θ_P not = 0, permuting alleles among populations. 95% CI refers to 95% confidence intervals calculated from bootstrap replicates over loci. *H* refers to Nei's (1978) unbiased heterogeneity estimate. N_em refers to the effective number of migrants inferred from θ_P values. 'inf' indicates infinite numbers of migrants.

Class	Locus	H	θ_P	(p)	N_em	D_1	D_2
ScnDNA:	BM47	0.14	-0.003	0.892	inf		
	BM81	0.43	0.052	< 0.001	4.6		
	BM32-2	0.47	0.102	< 0.001	2.2		
	WM08	0.45	0.123	< 0.001	1.8		
	Overall	0.37	0.086	< 0.001	2.7	0.068	0.054
	95% CI		(0.038 to 0.11)		(6.32 to 2)		
Allozyme:	ACOH	0.11	0.107	< 0.001	2.09		
	ADH	0.40	0.146	< 0.001	1.46		
	G3PDH	0.43	0.050	0.011	4.78		
	IDDH	0.24	-0.012	0.999	inf		
	Overall	0.30	0.077	< 0.001	3.00	0.038	0.034
	95% CI		(0.011to 0.138)		(22 to 1.56)		
Nuclear:	Overall	0.33	0.08	<0.001	2.875		
	95% CI		(0.042 to 0.111)		(5.7 to 2)		

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5. ScnDNA intra-ocean hierarchical analysis of variance. θ_{SS} refers to the effect of temporal replicates within populations. θ_S refers to the effect of populations within oceans, p_1 refers to the exact θ_{SS} test and p_2 refers to the exact probability of homogeneity in allele frequency between populations. * Indicates significance after corrections for multiple tests (initial $\alpha = 0.01$). 95% CI refers to confidence intervals calculated from bootstrap replicates over loci.

	BM47	BM81	BM32-2	WM08	Over All Loci	
Population	θ_{SS}	θ_{SS}	θ_{SS}	θ_{SS}	θ_{SS}	95% CI
Hawaii	0.023	-0.002	0.020	-0.005	0.006	(-0.003 to 0.021)
E. Pacific	-0.011	0.023	-0.013	-0.004	0.001	(-0.013 to 0.016)
Australia	-0.019	-0.025	-0.037	-0.041	-0.034	(-0.04 to -0.024)
Jamaica	0.006	0.013	-0.007	0.031	0.011	(-0.004 to 0.027)
p_1				*0.009**	*0.017*	
p_2				*0.012*	*0.021*	
U.S.	-0.023	-0.043	-0.043	0.062	-0.005	(-0.043 to 0.044)
	θ_S	θ_S	θ_S	θ_S	θ_S	95% CI
Atlantic	-0.011	-0.010	0.017	-0.006	0.000	(-0.010 to 0.012)
Pacific	0.007	-0.008	-0.007	-0.004	-0.005	(-0.008 to -0.001)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6. AMOVA inter-ocean divergence estimates from mtDNA RFLP data of Graves and McDowell (1995). θ_P refers to percent of genetic variation attributable to inter-ocean divergence. Allelic relatedness refers to the information on molecular haplotypes used in calculation of θ_P. p values indicate probability that θ_P differs significantly from 0. $N_e m_f$ refers to the effective number of female migrants per generation.

Composite Haplotypes	Nucleon Diversity	Allelic Relatedness	θ_P	p	$N_e m_f$
38	0.85	Restriction sites	0.25	0.001	1.50
		None	0.06	0.001	7.80
		Clade frequency	0.39	0.001	0.78

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

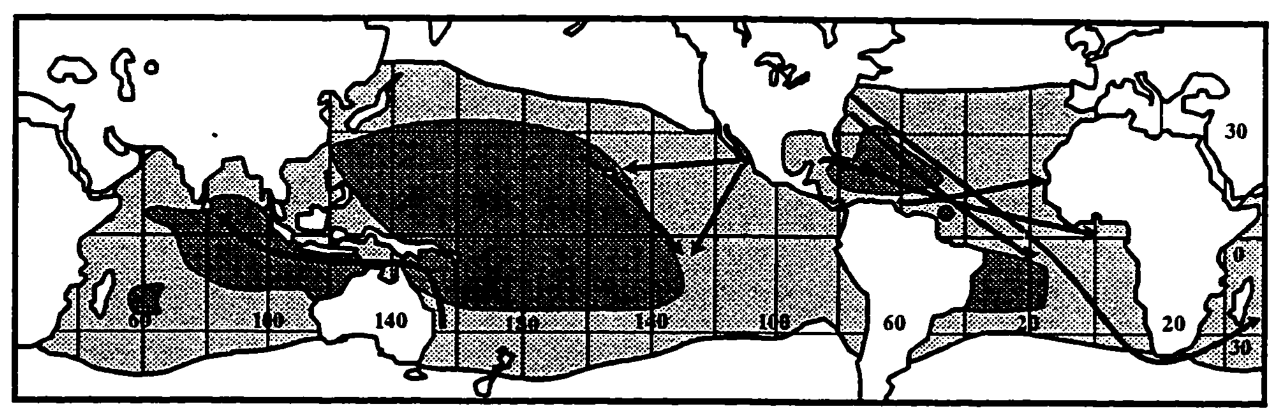
Table 7. Test of independence between nuclear and mtDNA allele frequencies within the Atlantic. p values refer to exact probabilities estimated with a Markov chain. 'A' refers to the frequency of the most common allele. '2N' refers to twice the number of individuals used in the test. Individual sample pairwise comparisons within the Atlantic were non-significant ($p > 0.015$) after correction for multiple tests (initial $\alpha = 0.00051$).

	BM47			BM81			BM32-2			WM08		
Atlantic Ocean	A	2N	*p*	A	2N	*p*	A	2N	*p*	A	2N	*p*
Atlantic Clade	0.94	192	0.23	0.72	176	0.12	0.35	180	0.77	0.73	154	1.0(
Ubiquitous Clade	0.90	260		0.78	262		0.33	268		0.73	242	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 1. Range and dispersal capability of blue marlin. Light shaded areas designate species' range (Nakamura 1974). Darker shaded areas designate larval distribution in the Atlantic (Bartlett and Haedrich 1968; Eschmeyer 1968; Ueyanagi et al. 1970) and Indo-Pacific (Howard and Ueyanagi 1965; Matsumoto and Kazama 1974; Nishikawa et al. 1985). Arrows indicate selected long-distance migration routes inferred from tag and recapture data in the Atlantic (Anonymous 1994; Bayley and Prince 1994; Scott et al. 1990), and Indo-Pacific oceans (Anonymous 1996; J. Pepperrell Pers. Comm. 1997).

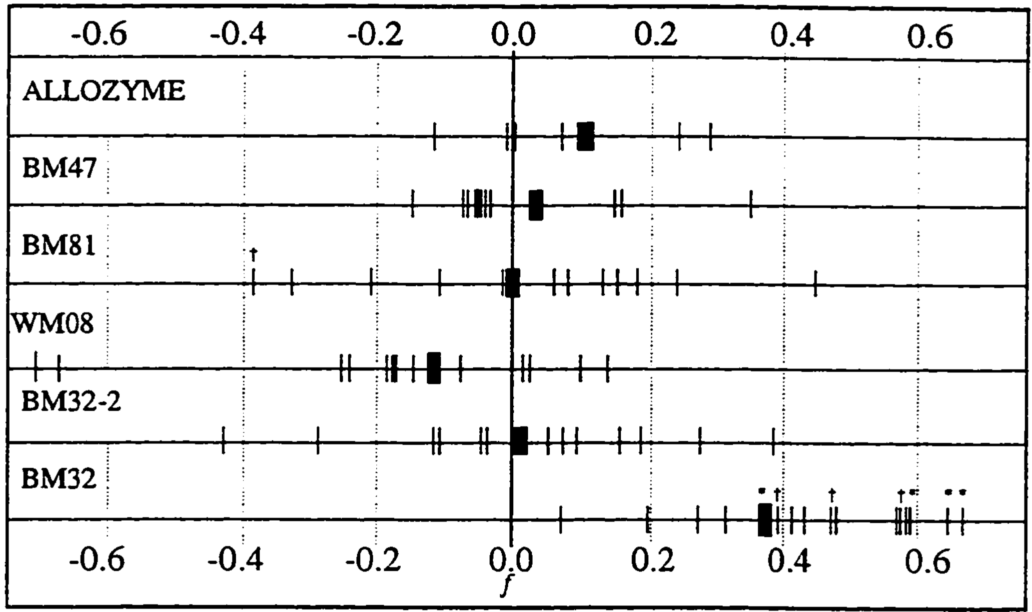
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 2. *f* values for each sample at four scnDNA loci (shown separately), and four allozyme loci (shown together). *f* values were calculated from [1 – (*Ho/He*)] (Weir and Cockerham, 1984), where *Ho* and *He* refer to observed and expected heterozygosities, respectively, measuring the deviation of each sample from conformance to Hardy-Weinberg expectations (f=0). Values non-significant unless indicated otherwise. "†" Indicates non-significant after correction for multiple tests (initial $\alpha = 0.0033$). "*" Indicates significant after corrections for multiple tests. Thicker line indicates mean *f* for each set of values.

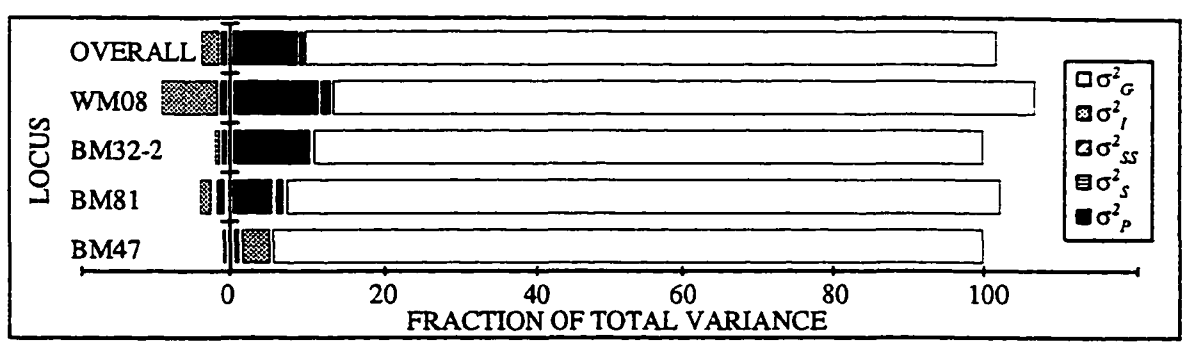
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 3. ScnDNA hierarchical analysis of variance over all populations. Source of each variance component: σ^2_G, genes within individuals; σ^2_I, individuals within subsubpopulations (years); σ^2_{SS}, subsubpopulations within subpopulations (geographic locations); σ^2_S, subpopulations within populations (oceans); and σ^2_P, between populations (oceans). Negative values result from bias corrections when the true percentage is near zero.

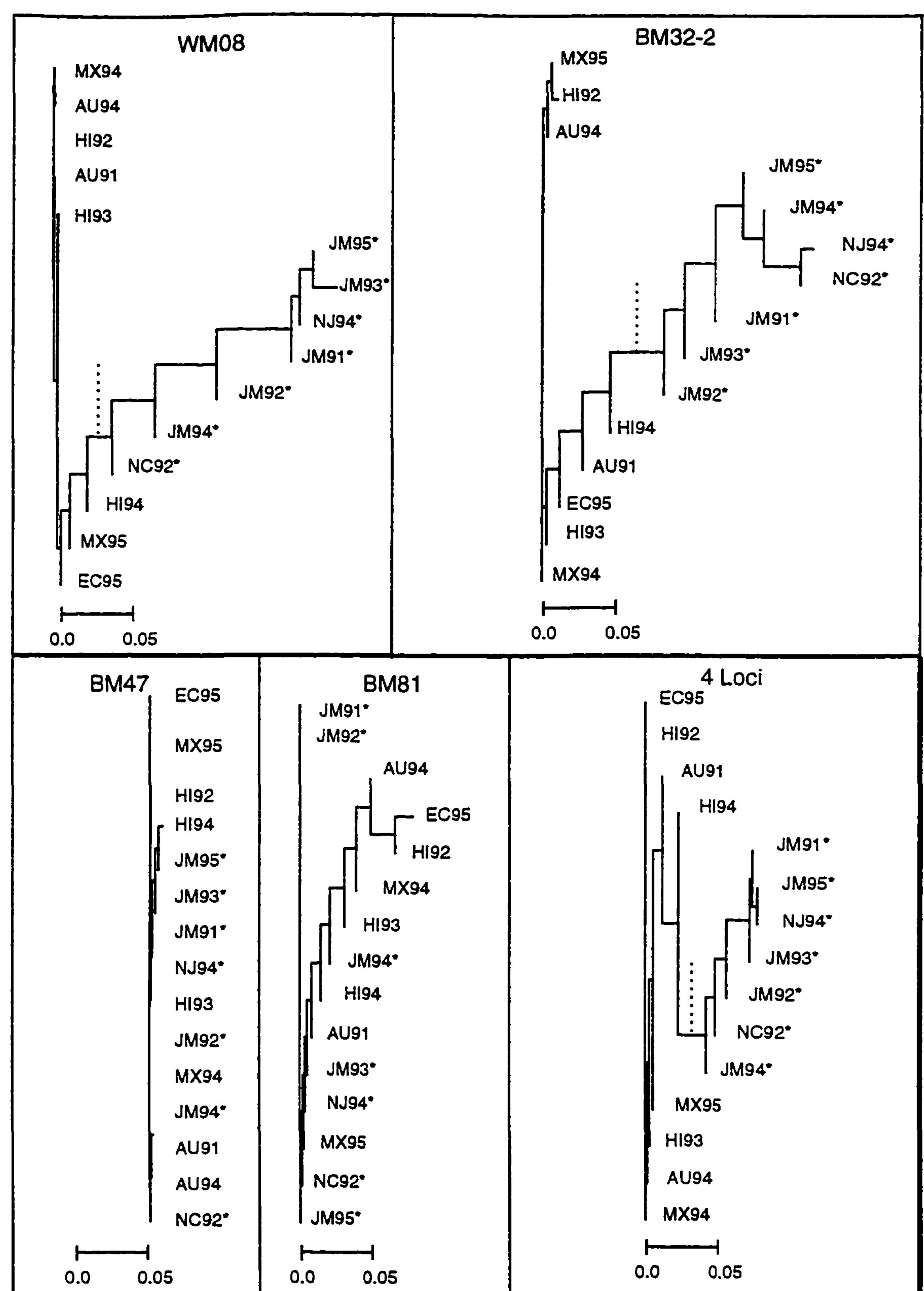
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 4. Neighbor-joining trees generated from four polymorphic scnDNA markers shown individually and over all loci. Nei's (1978) unbiased genetic distance was used as a measure of divergence. Populations are abbreviated as in Table 1. Dashed line indicates assumed midpoint root.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 5. Results of inter-ocean pairwise tests for population differentiation. Tests were performed among all populations (A), and among geographic locations pooled over years (B). For each population comparison, boxes indicate the significance level for BM47, BM81, BM32-2, and WM08, respectively, as indicated in the legend. Within-ocean probability levels were greater than 0.01 except at WM08 for JM93 vs JM94 (p =0.005; initial $\alpha = 0.00045$). Population abbreviations follow Table 1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A

	JM91	JM92	JM93	JM94	JM95	NC92	NJ94
HI92							
HI93							
HI94							
MX94							
MX95							
EC95							
AU91							
AU94							

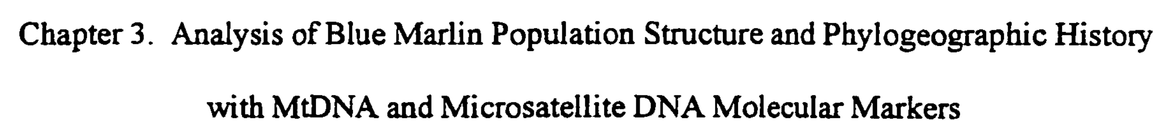
B

	JM	US
HI		
E-PAC		
AU		

P< 0.01 but non-significant after correction for multiple tests.

Significant after correction for multiple tests.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Chapter 3. Analysis of Blue Marlin Population Structure and Phylogeographic History with MtDNA and Microsatellite DNA Molecular Markers

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

INTRODUCTION

Previous genetic studies of mitochondrial and nuclear DNA markers within blue marlin revealed different levels of genetic diversity and inter-ocean divergence. In Chapter 2, analyses of eight, moderately polymorphic allozyme and single copy nuclear DNA (scnDNA) loci revealed relatively minor yet statistically significant allele-frequency differences between ocean samples. In contrast, sequencing of the mitochondrially encoded cytochrome *b* gene (Finnerty and Block, 1992) and whole-molecule mtDNA RFLP studies (Graves and McDowell, 1995) demonstrated two highly divergent clades of haplotypes, one of which was found almost exclusively in Atlantic samples. The greater divergence of mtDNA was explained by genetic drift resulting from the smaller effective population size of the mitochondrial genome, but the influence of the high mtDNA mutation rate could not be excluded.

In this chapter, the development and analysis of highly polymorphic microsatellite DNA markers for blue marlin is reported. Microsatellites consist of short (2-6 bp), tandemly repeated DNA motifs that commonly vary in repeat number (Tautz et al., 1986). Microsatellites have been of great utility for population genetic studies as they are typically highly polymorphic, abundant, widely distributed, and non-coding (Wright and Bentzen, 1994). Microsatellite size mutations occur at rates that are typically orders of magnitude greater than base pair substitutions (Weber and Wong, 1993). The high mutation rate of microsatellite loci has allowed detection of population structure in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

species where previous studies employing mitochondrial and/or nuclear markers were uninformative (Bentzen et al., 1996; McConnell et al., 1995; Paetkau et al., 1995). However, highly polymorphic markers also can hinder detection of genetic divergence due to both high within-population variance (Bowcock et al., 1994; Jin and Chakraborty, 1995) and homoplasy due to allele size constraints (Nauta and Weissing, 1996) or flanking region mutations (Angers and Bernatchez, 1997).

This chapter includes a hierarchical analysis of molecular variance within blue marlin based on restriction fragment length polymorphisms (RFLPs) of the mitochondrial genome (Graves and McDowell, unpublished data) and five highly polymorphic, tetranucleotide microsatellite loci. The AMOVA provided a flexible statistical tool for comparison of molecular markers by providing completely analogous statistical descriptors for diverse data types (Excoffier et al., 1992). Different methods of analyzing data from the same type of molecular marker were used to identify the relative influence of mutation and drift on patterns of allele frequency divergence among populations. The RFLP data extend the work of Graves and McDowell (1995) by the addition of temporal samples spanning a period of five years, and inclusion of a number of geographically distant collection locations throughout the Atlantic and Pacific oceans. Development of microsatellite markers adds a new level of sensitivity to the question of blue marlin population structure, and further, allows investigation into the role of mutation rate in the differences previously observed between nuclear and mitochondrial markers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MATERIALS AND METHODS

Blue Marlin Tissue Samples

Blue marlin heart tissue samples were obtained over several years from fish landed at sport fishing tournaments, artisinal fisheries, and research cruises. Sample locations within the Atlantic included the coasts of New Jersey (NJ), North Carolina (NC), and Jamaica (JM). Samples from the Pacific were obtained from Hawaii (HI), Mexico (MX), Ecuador (EC), and eastern Australia (AU; Table 1). Tissues were dissected within eight hours of death, chilled on ice, frozen at -20°C, transported to the laboratory on dry ice or a commercial ice substitute, and maintained at -70°C until analysis. Overall, 243 Atlantic and 221 Pacific individuals were available for genetic analyses.

Microsatellite Library Construction

Protocols of Kijas et al. (1994) and Waldbieser (1995) were modified as described below to generate a microsatellite-enriched mini-library of plasmid DNA clones. Briefly, the procedure involved the following steps: (i) blue marlin DNA fragments were isolated from genomic DNA and ligated to a plasmid vector to add universal priming sites; (ii) the construct was denatured (made single-stranded), hybridized to a microsatellite DNA probe, and exponentially amplified via PCR by using universal priming sites; and (iii) the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

amplified product was ligated to a plasmid vector and transformed into *E. coli* bacterial cells.

Total genomic DNA was isolated from frozen, blue marlin heart tissue following protocols of Sambrook et al. (1989). Genomic DNA was digested with *Mbo* I (Gibco BRL) and fragments in the 500 - 1000 base pair (bp) size range were isolated by electrophoresis onto DE-81 paper (Whatman International Ltd., Maidstone, England). Universal priming sites were added to selected DNA fragments by ligating genomic fragments to the plasmid vector Bluescript KS^+ (Stratagene, La Jolla, CA), using T4 DNA ligase (Stratagene, La Jolla, CA). The vector was prepared for ligation by digestion with *Bam*H I, then phosphorylated with alkaline phosphatase (CIAP; Stratagene, La Jolla, CA) and cleaned with phenol/chloroform extractions. To generate a collection of single-stranded DNA fragments, asymmetric PCR was performed on the clones by using a 10:1 ratio of universal T7 forward and M13 reverse primers (Genosys). Microsatellite-containing fragments were selected from single-stranded genomic DNA by using a series of hybridization reactions between genomic DNA and a probe containing a microsatellite motif. The probe consisted of streptavidin coated para-magnetic beads (Promega) bound to a biotin-labeled $(GATA)_5$. Procedures for probe preparation, hybridization of the probe complex to the genomic fragments, and exponential amplification of the resulting purified fragments followed Kijas et al. (1994). To prepare for a second series of ligation reactions, double stranded PCR products were digested with *Mbo* I. The plasmid vector was freshly digested, phosphorylated, and cleaned as above. Competent cells were transformed with ligation products following protocols of the TA cloning kit (InVitrogen). Bacterial colonies were grown overnight on solid agar media containing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ampicillin (Sigma, St. Louis, MO) for preferential growth of plasmid-containing cells, and X-gal (Sigma, St. Louis, MO) for blue/white selection of clones.

The resulting mini-library was screened for microsatellite-containing clones by using PCR methods. PCR was performed on putative recombinants by using a $(GATA)_5$ primer in combination with either universal T7 forward or M13 reverse primers (Genosys). To prepare template DNA for screening reactions, DNA was released from approximately 50 mg of bacteria through boiling in 500 uL water. PCR was performed using BRL PCR Reagent Systems (Gibco BRL) with a final concentration of 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.05 uM of each primer, approximately 5 ng (0.5 uL) template DNA, and 0.1125 units *Taq* polymerase in 5 uL total volume. Cycling conditions included an initial denaturation of 5 min at 95°C, 32 cycles of 1 min at 55°C, 1 min at 72°C, and 1 min at 95°C, and a final 7 min extension at 72°C. Since the $(GATA)_5$ primer annealed to many complementary sites at microsatellite-containing clones, smeared amplification products were considered positive indicators for the presence of microsatellites. Positive clones were cycle-sequenced with Thermo Sequenase sequencing kits (Amersham) by using labeled (IRD-800) *M13* universal primers. Reaction products were electophoresed and detected on an automated sequencer with flourometric detection (LiCor, Lincoln, Nebraska). Primers were designed from microsatellite flanking sequences with assistance of the computer program PC/GENE (Bairoch, 1989). Primer optimization criteria included a length of approximately 22 bp, 45 - 60% GC content, and a maximum of four single base repeats, four self-complementary base pairs, and three primer-primer complementary base pairs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Clones containing microsatellite repeat motifs were chosen for further analysis without regard to repeat size, degree of perfection of repeat area, level of variability, or patterns of divergence between populations. Ultimately, microsatellite loci were selected with the following criteria: DNA fragments had to hybridize to the $(GATA)_5$ probe, contain a sufficient amount of flanking region for primer construction, and amplify during asymmetric and exponential PCR under a variety of primer combinations (Kijas et al., 1994; Waldbieser, 1995).

Population Screening

The polymerase chain reaction (PCR) was used to amplify target microsatellite loci from genomic DNA. Template nuclear DNA was obtained from the nuclear band of cesium chloride/ethidium bromide density gradient mtDNA purifications after dialysis (Lansman et al., 1981). PCR was performed using the BRL PCR Reagent System (Gibco BRL), with a final concentration of 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.05 uM of each primer, approximately 5 ng template DNA and 0.1125 units *Taq* polymerase in 5 uL total volume. Typical cycling conditions included an initial denaturation of 5 min at 95°C, followed by 32 cycles of 1 min at 55°C, 1 min at 72°C, and 1 min at 95°C. Final extension was carried out for 7 min at 72°C. Annealing temperature and primer sequence were the main variables adjusted to optimize PCR conditions for reliable amplification. Either the forward or reverse primer was labeled with a flourescent dye (IRD-800) to allow for size detection of amplified products. PCR products were electrophoresed on 25 cm, 8% Long-Ranger denaturing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

polyacrylamide gels (FMC BioProducts), and detected using a LiCor scanner. Products were scored using RFLPSCAN software (Scanalytics, CSPI).

Microsatellite Statistical Analyses

Several parameters were estimated to characterize patterns of within-population diversity of microsatellite loci within individual and combined samples. Expected heterozygosities represent the probability of obtaining alleles of different size in a random draw from a sample. Repeat number mean, variance, and standard deviation were estimated to describe location and dispersion pattern of each allele frequency distribution. The distribution of pairwise differences between all alleles (mismatch distribution) was generated with Arlequin v. 1.1 (Schneider et al., 1997). The mean of this distribution represents the average difference in repeat number between two randomly chosen microsatellite alleles at any locus or sample under consideration. The y-intercept of the mismatch distribution is the expected homozygosity; the mismatch range extends to the total number of alleles minus one. Modes within the mismatch distribution represent comparisons of groups of alleles in the allele frequency distribution.

The local inbreeding coefficient, F_{IS} (Wright, 1978) or f (Weir and Cockerham, 1984) was calculated using the program GENEPOP (Raymond and Rousset, 1995). This statistic measures the deviation of each population from the genotypic proportions expected under Hardy-Weinberg equilibrium. For nuclear markers, conformance of genotypic proportions to Hardy-Weinberg equilibrium is consistent with Mendelian inheritance and non-preferential amplification of alternative alleles during PCR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

amplification. Once Mendelian inheritance was accepted, conformance of genotypic frequencies to Hardy-Weinberg equilibrium within individual populations was evaluated. Unbiased estimators of exact significance probabilities for tests of Hardy-Weinberg disequilibrium were used to evaluate the null hypothesis that alleles were randomly distributed among individuals at each locus. These tests were performed using the Markov chain algorithm described in Guo and Thompson, (1993), as implemented in GENEPOP. For these analyses, the Markov chain length was set to 500,000 steps to reduce the standard error of probability values to < 0.0002 for tests of significance. Type I error was controlled using the sequential Bonferroni method (Rice, 1989).

The Excoffier et al., (1992) extension of the Weir and Cockerham (1984) method for generating unbiased hierarchical *F*-statistics was chosen to analyze patterns of genetic diversity within and among populations. This method incorporates explicit corrections for sample sizes and number of populations surveyed, facilitating comparison of marker classes. The method also has the flexibility to incorporate information on relatedness of alleles as either restriction site differences or number of stepwise mutations. For purposes of clarity, *F* statistics based on allele frequency alone were denoted θ, and those incorporating the microsatellite stepwise genetic distances denoted by *R*. The term *F* was reserved as a general reference to *F* statistics, without regard to method of calculation.

Hierarchical analyses of variance were performed using the computer program Arlequin v. 1.1 (Schneider et al., 1997); overall genetic variance was partitioned into levels as outlined in Table 1. Description of variance components (σ^2) and *F* statistics followed notations of (Weir, 1996). Subscripts refer to the following: *P*, populations (samples grouped by oceans); *S*, subpopulations within populations (geographic locations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

within oceans); *SS*, subsubpopulations within subpopulations (temporal samples at a geographic locations); *I*, individuals within subsubpopulations; and *G*, alleles within individuals. For the *S* level, effects averaged over oceans and for each ocean were determined. For the *SS* level, effects averaged over all locations and for each location were determined. For microsatellite loci the AMOVA was performed (i) with allelic distances input as squared Euclidean distances calculated as the square of the repeat number (*R*); (ii) without allelic distance information considered (θ); and (iii) with alleles binned into modes. Modes within a distribution of microsatellite alleles were defined as occurring at a frequency greater than 10% and including more than a single allele.

Several methods were available to evaluate genetic distance between populations based on microsatellite data. θ is a function of the variance among population allele frequencies divided by the total variance in allele frequency. Variances were calculated separately for each microsatellite allele, and the numerator and denominator were summed over alleles to obtain a weighted average for each locus. This method did not incorporate any relatedness information between alleles. The *R* statistic is a function of the repeat number variance between population means divided by the total repeat variance. Both θ and *R* statistics were corrected for sample size and sample number bias, following Weir and Cockerham (1984) and Rousset (1996), respectively. Defined in this manner, the *R* statistic is equivalent to the *Rho* statistic proposed by Michalakis and Excoffier (1996) and Rousset (1996), and is related to the R_{ST} of Slatkin (1995) by a correction for sample number, as described in Rousset (1996). The *R* statistic is flawed because there is a large variance associated with the estimator. The Delta mu statistic (*Dmu*; Goldstein et al., 1995) was calculated as an additional measure of divergence

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

between ocean samples. *Dmu* also incorporates the stepwise mutation model, and is equivalent to the square of the difference in the mean repeat number of two populations. This measure allows for comparison to other measures of microsatellite population divergence that involve partitioning of variance, such as θ and R. *Dmu* has a lower variance as populations become more distant, and maintains linearity for a longer period of divergence time.

Exact probability tests and exact F tests were used to determine significance of population subdivision within each hierarchical level and for individual population comparisons. To assess significance of R values, an exact R test was performed as described by Excoffier et al. (1992) and implemented by the computer program Arlequin (Schneider et al., 1997). Exact probability tests calculated the probability of each randomly generated data assortment, under the null hypothesis that alleles are taken from a single population. This method does not incorporate allelic relatedness, and is a test of the significance of the θ statistic. Unbiased estimators of exact significance probabilities were obtained as described for tests of Hardy-Weinberg equilibrium.

Significance of population structure was evaluated at several levels by using both exact probability tests and exact R significance testing methods. Significance of inter-annual variation in allele frequencies was evaluated over all temporal samples at a location, and in pairwise comparisons between all possible temporal samples at a location. Significant variation in allele frequencies among locations within an ocean was evaluated among all temporally pooled samples and in pairwise comparisons with temporal samples considered individually and pooled. Significance of inter-ocean pairwise population comparisons was determined with temporal replicates considered

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

separately and pooled. Critical values for sample pairwise comparisons were highly conservative, given very low initial critical values when correcting for large numbers of tests per locus (initial $\alpha = 0.00047$ for 105 pairwise comparisons). For greater sensitivity, results are presented for comparisons that are significant after corrections for multiple tests, and for non-significant comparisons with nominal uncorrected p-values less than 0.01.

Phenetic dendrograms were constructed to determine how well each genetic locus clustered individual samples to their geographic location and ocean of origin. Phenetic dendrograms were constructed from Slatkin's linearized R, Slatkin's linearized θ, and *Dmu* distances. θ and R statistics were calculated as above, with a correction to make these distances linear with respect to time (corrected distance = -log (1-F); Reynolds et al., 1983). Neighbor-joining (Saitou and Nei, 1987) and UPGMA (Sneath and Sokal, 1973) algorithms were used to construct dendrograms for each locus and over all loci. The neighbor-joining method differs from UPGMA by removing the assumption that all lineages diverge at equal rates. Variability in tree topology among clustering methods is one measure of the relative stability of the topology obtained. Individual loci were evaluated for accuracy of tree reconstruction. Clustering of samples from the same ocean into the opposite oceans was considered an "error". Estimates of the number of migrants (N_em) between oceans were calculated from [$F_{ST} = 1/(4N_em + 1)$], assuming an island model of migration (Wright, 1978). F values were corrected for sample number bias.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MtDNA

Whole molecule restriction fragment length polymorphism (RFLP) data were available from a total of 422 individuals (Graves and McDowell, unpublished data). Data consisted of composite haplotypes generated from a series of eleven restriction enzymes. Sample sizes were comparable to those of the microsatellite data and included many of the same individuals (Table 1). Notable exceptions were the 1993 sample from Jamaica and the 1994 sample from Mexico, where tissues were not of sufficient quality for whole mitochondrial molecule RFLP analysis.

Statistical analyses of mtDNA data were performed as with the microsatellite data but with the following modifications. To describe relationships among mtDNA haplotypes, phenetic trees were generated from a matrix of nucleotide sequence divergence estimates between mtDNA haplotypes (Nei and Miller, 1990), generated with the computer program REAP (McElroy et al., 1991). AMOVA was performed (i) with haplotype distances input as squared Euclidean distances, calculated as the number of restriction site differences between haplotypes; (ii) without relatedness criteria considered; and (iii) with haplotypes binned into either ubiquitous or Atlantic clades. F statistics incorporating mtDNA restriction site differences were denoted by Φ. For phenogram construction, Slatkin's linearized Φ was used as a distance measure for the mtDNA data (Schneider et al., 1997). Estimates of the number of female migrants $N_e m_f$ between oceans were calculated from [$\Phi = 1/(2N_e m_f + 1)$] (Birky et al., 1983).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RESULTS

Microsatellite Loci

Mini-Library. The enrichment protocols produced more than 500 putatively recombinant (white) colonies from which to screen for microsatellite-containing DNA fragments. A total of 105 clones was screened with the PCR method, and amplification products were obtained from 64. Of these, 31 clones produced a smear pattern, indicating the presence of a (GATA) microsatellite repeat. Of the 33 that did not produce smears, five were sequenced, and of those, none contained microsatellite motifs. Of the 31 clones that produced smears, 22 were sequenced and all contained microsatellite motifs. Two sequences represented alternative alleles for the same locus (GATA08). The sequence and length distribution of 22 clones is shown in Table 2. The 22 clones averaged 28.2 repeats per locus. Of the 22 clones, eight sequences contained uninterrupted repeat motifs, four were interrupted by a single point substitution, five were interrupted by two to nine point substitutions, two were uninterrupted composites of GATA and GACA repeat motifs, and three were interrupted composites of GATA and GACA repeat motifs. In no case was a change in repeat motif size detected.

Hardy-Weinberg Equilibrium. Genotypic frequencies were in close conformance with expectations of Hardy-Weinberg equilibrium for four of five microsatellite loci (Fig. 1, Table 3). Significant deviations (heterozygote deficiencies) from Hardy-Weinberg

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expectations were detected at GATA10 for three populations (HI92, JM92, and MX95), and when tested over all populations (average $f = 0.11$). An average of 2.8 populations per locus (Fig. 1) deviated further than $f = 0.1$ from Hardy-Weinberg expectations, with nine of 14 values attributable to GATA10 (Fig. 1, Table 3). Deviations greater than $f = 0.1$ for each location (Table 3) revealed an even distribution of deviations among oceans (Mann Whitney U $p = 0.767$). Samples with the fewest number of individuals were most deviant (AU and US).

Diversity. The diversity of allele frequency distribution patterns among microsatellite loci is shown in Figures 2A-2E and Figure 3. All loci were hypervariable, displaying bell shaped allele frequency distributions (Fig. 2). A total of 161 microsatellite alleles was detected over all populations, with an average of 32.2 alleles per locus (Table 4). The average size of microsatellite alleles over all populations and loci was 24.9 repeats (standard deviation of 6.5 repeats), with a range of 1 to 58 repeats. Over all loci, there was an average of 6.4 repeat differences. Heterozygosity per locus averaged 0.93, with a range of 0.91 to 0.95. The mismatch distribution for each locus, and over all loci, is shown in Figure 3.

GATA01 and GATA10 demonstrated the lowest heterozygosities ($H = 0.93$ and $H = 0.91$, respectively) and allelic divergences of the five microsatellite loci (Table 4). GATA01 had 27 alleles and an average allelic divergence of 4.5, while GATA10 had 28 alleles and an average allelic divergence of 4.8 (Table 4). Both GATA01 and GATA10 exhibited a single major mode of alleles in both oceans (Figs. 2A, 2C). Levels of genetic diversity for both GATA01 and GATA10 were higher within Atlantic than Pacific

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

samples. Atlantic samples also exhibited higher variance in repeat number and in average number of pairwise differences in repeat number (Table 4; Fig. 4). For GATA10, variance in repeat number was more than twice as high in Atlantic samples, and the average number of pairwise differences in repeat number was 1.4 times greater in the Atlantic samples. The broader and less peaked mismatch distribution of Atlantic samples (Fig. 4C) further reflected the greater allelic divergence relative to Pacific samples.

GATA08 exhibited the greatest genetic diversity of the five microsatellite loci, with 47 alleles, a broad range in allele size (1 – 58 repeats), an average heterozygosity of 0.95 across samples, and an average allelic divergence of 9.7 repeats between alleles over populations. Alleles were divided into two modes (Fig. 2B). The first mode of alleles in Atlantic samples extended from 1 to 36 repeats (average = 20.3 repeats), contained several minor peaks, and exhibited an average allelic divergence of 6.0 repeats. Alleles in Pacific samples contained a mode with a similar average number of repeats (19.0) but were less diverse, with an average allelic divergence of 5.2 repeats. A second mode of GATA08 alleles spanned from repeats 40 to 58, and occurred at a frequency of 18.3% in the Atlantic and 0.5% in the Pacific (Fig. 2B). The second mode was less diverse than the first, with a smaller average allelic divergence (4.8 vs 5.7 repeats). Comparison of the mismatch distribution for the two modes within the Atlantic (Fig. 5A) demonstrates the larger diversity of the first mode. A wide mode in the overall Atlantic mismatch distribution is present from 18 to 50 repeat difference units, clearly inflating Atlantic divergence (Fig. 4B). The high frequency of both modes in the Atlantic resulted in an

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

increased variance in repeat number (144.4 vs 27 squared repeats), and divergence between alleles (12.5 vs 5.4 repeats) relative to Pacific samples.

Allele frequency distributions for GATA60 (Fig. 2D) and GATA90 (Fig. 2E) were both bimodal, with greater diversity in the Pacific. For GATA60, 30 alleles were detected over a range of 48 repeats. Variability was intermediate with respect to heterozygosity (0.93) and mean allelic divergence (5.7 repeats) relative to other microsatellite loci. GATA60 (Fig. 2D) exhibited two overlapping modes separated by an average of 8.4 repeats. Over all populations, the first mode (< 31 repeats) was less diverse than the second (31 to 53 repeats), with average allelic divergences of 2.4 and 3.6 repeats, respectively. The second mode was present at a frequency of 63% in the Pacific and 26% in the Atlantic, and was more variable in the Pacific than Atlantic, as measured by average divergence between alleles (3.9 vs 2.9 repeats). The stronger bimodality and higher diversity of the second mode in Pacific samples resulted in 1.26 times greater allelic divergence over all Pacific samples, as seen with its broader, flatter mismatch distribution (Fig. 4D).

A total of 33 alleles was observed at GATA90 (Fig. 2E). As with GATA60, intermediate levels of variability ($H = 0.945$) and allelic divergence (mean divergence = 7.5) were detected relative to other microsatellite loci. Two overlapping modes were evident, separated by an average of 14.5 repeats. The second mode, ranging from repeats 29 to 42, was present in Pacific samples at a frequency of 28% and Atlantic samples at 11%. This mode had a smaller range (14 vs 21 repeats) and divergence between alleles (2.6 vs 4.6 repeats). Because of stronger bimodality of the Pacific distribution, overall allelic divergence in Pacific samples was 1.3 times larger than in Atlantic samples. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mismatch distributions for both GATA60 and GATA90 were flatter for the combined Pacific samples. Averaged over all microsatellite loci, Atlantic samples were more diverse than Pacific samples with respect to allelic divergence (6.6 vs 5.4; Fig. 6). In the pooled analysis, most mismatch modes were not apparent, yet the second GATA08 mode was visible and contributed substantially to the greater overall microsatellite diversity in Atlantic samples (Fig. 6).

The largest proportion of microsatellite repeat variance was present within populations (89% over all loci; range among loci 58 to 92%; Fig. 7). The mean proportion of repeat variance attributable to oceanic divergence was 14.5% over all loci, and ranged from 1.2 to 27%. The proportion of variance attributable to division of locations within oceans and years within locations was negligible when considered over all loci, comprising -0.8% (combined).

Inter-annual and Intra-Ocean Divergence. No major differences in genetic diversity were detected among temporal replicates or locations within oceans (Table 5). In general, smaller samples (AU and US) demonstrated less consistency in estimates of diversity among replicate samples. The degree of inter-annual and intra-ocean divergence of microsatellite allele frequencies also was negligible. $R_{SS\text{-}OVERALL}$ and $\theta_{SS\text{-}OVERALL}$ values ranged from -0.0025 to 0.0023 across the five loci, with a value of 0.0005 for all loci combined (Tables 6A, B). p values for both R_{SS} and θ_{SS} were non-significant in nearly all cases after corrections for multiple tests ($p > 0.06$). Borderline significance was detected for $\theta_{SS\text{-}HI}$ at GATA60 ($p = 0.009$ +/- 0.001; initial $\alpha = 0.01$; Table 6B). Over all loci, only $\theta_{SS\text{-}HI}$ was close to significance ($p = 0.05$). The difference in allele

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequencies at GATA60 for each pairwise comparison among the three temporal samples from Hawaii is shown in Fig 8F. Rather than cohesive shifts in frequency among groups of alleles, variation included uniformly distributed inter-annual frequency shifts.

Exact probability and exact R_{SS} tests for pairwise population differentiation among temporal replicates within geographic locations were non-significant in all cases after corrections for multiple tests ($p > 0.018$; initial $\alpha = 0.00047$). For each locus and testing method, a total of 16 pairwise inter-annual comparisons was evaluated. Considering all tests, only four within-location pairwise probability tests had p values below 0.05; these involved comparisons of HI93 versus HI94 at both GATA01 and GATA60, and two JM92 comparisons at GATA01. Genetic divergence among samples from the same ocean, for individual loci as well as all five loci combined, was less than 0.005 (Fig. 7, Table 6). Probability values approaching significance (initial $\alpha = 0.025$) were detected for $\theta_{S\text{-}ATL}$ at GATA08 ($p = 0.054$), and for $\theta_{S\text{-}PAC}$ at GATA10 ($p = 0.047$). However, over all loci within each ocean, θ_S and R_S values were non-significant ($p > 0.17$). Exact probability and exact R_S tests for pairwise population differentiation among locations within oceans were non-significant in all cases when corrected for multiple tests ($p > 0.005$; initial $\alpha = 0.00047$). Three individual pairwise comparisons had p values less than 0.01; comparisons between HI93 vs AU91 and HI93 vs MX94 at GATA01 ($p > 0.005$), and HI92 vs MX95 at GATA60 ($p = 0.005$ +/- 0.0005; initial $\alpha = 0.00047$). When samples were pooled over temporal replicates, θ_S pairwise comparisons were non-significant ($p > 0.022$; initial $\alpha = 0.0125$), as were R_S comparisons ($p > 0.16$). A lack of intra-ocean population structure also was evident from analyses of neighbor-joining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clustering of samples by overall similarity of allele frequencies. For each locus, temporal replicates were dispersed among locations within an ocean (Figs 9A - 9C).

Inter-Ocean Divergence. Several microsatellite loci exhibited striking patterns of genetic divergence between ocean samples. The difference in frequency between Pacific and Atlantic samples for each locus is illustrated in Figures 8A through 8E. Significant heterogeneity of allele frequencies between oceans was detected at each microsatellite locus in both exact probability tests ($p < 0.0003$) and exact R_P tests ($p < 0.034$; Table 6A-C). For GATA01 and GATA10, allele frequency distributions broadly overlapped between oceans (Figs. 2A, 2C), with differences in mean repeat number of only 0.86 and 0.87, respectively. For GATA01 and GATA10, these distributions resulted in small *Dmu* divergences (0.75 and 0.76), significant R_P values of 0.02 ($p = 0.00079$ +/- 0.00027) and 0.01 ($p = 0.034$ +/- 0.002), and highly significant θ_P values of 0.001 ($p = 0.0003$ +/- 0.0002) and 0.014 ($p < 0.00001$). GATA08 displayed highly discordant frequency distributions between oceans, resulting in a mean inter-ocean difference of 6.3 repeats, a *Dmu* of 40.1, an R_P of 0.18 ($p < 0.00001$), and a θ_P of 0.019 ($p < 0.00001$). At GATA60, Atlantic and Pacific samples were separated by an average of 4.2 repeats, resulting in a *Dmu* of 17.2, R_P of 0.27 ($p < 0.00001$) and θ_P of 0.017 ($p < 0.00001$). For GATA90, alleles were larger in the Pacific by an average of 3.6 repeats, resulting in a *Dmu* of 12.7, R_P of 0.13 ($p < 0.00001$), and θ_P of 0.009 ($p < 0.00001$). Cohesive modal shifts of alleles for GATA08, GATA60, and GATA90 are illustrated in Figure 8. When alleles were binned into modes, inter-ocean divergence estimates remained high for bimodal loci;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

values of 0.164, 0.245 and 0.093 were obtained for GATA08, GATA60, and GATA90, respectively.

Pairwise Population Comparisons. Both exact probability tests and exact R_P tests reflected patterns of divergence similar to those detected with AMOVA. A large degree of variation was detected in the number of significant comparisons ($p < 0.01$) among loci, samples, and method of evaluation. The distribution of significant inter-ocean pairwise population comparisons for each locus is shown in Figure 10, for both exact R_P and exact probability testing methods. For exact R_P tests, a range of zero to 46 of 56 total inter-ocean comparisons was significant ($p < 0.01$), averaging 21.4 significant comparisons per locus (Fig. 10A; Table 7). For the exact probability test, a range of four to 33 inter-ocean comparisons was significant, with an average of 19.8 significant comparisons per locus (Fig. 10B). GATA01 and GATA10 exhibited the fewest significant results, with a maximum (over R and θ methods) of four and 14 out of 56 significant comparisons, respectively. GATA60 displayed the highest number of significant inter-ocean differences for any microsatellite locus.

Although the average number of significant inter-ocean comparisons over loci changed by less than one between exact R_P and exact probability methods, results for individual loci varied by an average of nine comparisons. GATA01 and GATA10 loci failed to detect a single significant difference with the exact R method, yet detected 4 and 14 differences with θ. In contrast, comparisons for the other three loci were more significant using the exact R method. Decreases of 1, 13, and 10 significant comparisons accompanied the shift from R to θ methods, for GATA08, GATA60, and GATA90,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

respectively. The relatively small shift at GATA08 resulted from both strong bimodality (*R*) and large individual allele frequency shifts (θ) differentiating inter-ocean comparisons.

The number of significant inter-ocean comparisons was affected by sample size. A range of 15 to 63% of pairwise comparisons per sample was significant over all testing methods. When sample size was increased by pooling, the significance of comparisons increased (Fig. 10 A-C).

Comparison of Phenetic Trees. Neighbor-joining and UPGMA phenograms were constructed for each microsatellite locus using Slatkin's linearized *R*, Slatkin's linearized θ, and *Dmu* distance measures (Figs. 9A – 9C). Of the five microsatellite loci, the two most divergent loci identified from the inter-ocean AMOVA, GATA60 and GATA08, separated samples into ocean basins in all analyses (Table 8). Across distance measures, *Dmu* outperformed linearized *R*, grouping all populations for GATA90 accurately. Over both *R* and *Dmu* distance measures, an average of 2.5 of five loci grouped populations without error, whereas four of five were grouped correctly using θ. The overall genetic difference between samples was summarized in the phenograms for each distance method. Examination of repeat means from Table 5, as well as neighbor-joining (Fig. 9) and UPGMA dendrograms (not shown) for *R* and *Dmu* distance methods, show that samples with the fewest number of individuals displayed the most variance in placement with respect to ocean affinity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Mitochondrial DNA

Diversity. RFLP analysis of mtDNA revealed considerable variation. A total of 127 haplotypes was encountered among 422 individuals, resulting in a haplotype diversity of 0.87 (Table 4). The overall mismatch distribution was strongly bimodal (Fig. 3), with an average of 4.5 restriction sites differences between alleles. UPGMA clustering divided mtDNA haplotypes into two distinct clades (Fig. 2F). One of the two clades (Atlantic) was almost completely restricted to Atlantic samples, comprising 40% of Atlantic individuals and only 0.5% of Pacific individuals (all from Australia). The presence of the Atlantic clade within the Atlantic Ocean contributed to the relatively large haplotypic diversity and divergence measures within those samples. The average haplotypic diversity in the Atlantic was 0.93 compared to 0.77 in the Pacific. The mismatch distribution was more divergent in Atlantic samples, with a mean allelic divergence of 6.1 sites in Atlantic samples and 1.8 in the Pacific samples (Fig. 4F). The Atlantic clade was much more diverse than the ubiquitous clade: mean divergence between alleles in the Atlantic clade was 5.2 restriction sites and nucleon diversity (h) was 0.97, as compared to a mean divergence of 1.7 restriction sites and an h of 0.79 in the ubiquitous clade (Fig. 5B).

AMOVA. The majority of mitochondrial variance occurred within populations, with negligible amounts of inter-annual and intra-ocean variance (Fig. 7, Table 6A). In contrast, approximately 21% of genetic variance was attributable to inter-ocean divergence when calculated based on restriction site information (Φ; Table 6A, 6B). Using haplotypic frequency data alone, only 3% of variance was attributed to inter-ocean

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

divergence (θ). When alleles were binned into the two clades, inter-ocean divergence accounted for 38.7% of the total variance.

Inter-annual and Intra-ocean Divergence. The degree of mitochondrial divergence was negligible in all comparisons involving years at a location. θ_{SS} and Φ_{SS} values averaged over all inter-annual tests were small ($\theta_{SS} = 0.005$, $\Phi_{SS} < 0.001$). Comparisons over all years within locations were non-significant after corrections for multiple tests (initial $\alpha = 0.01$), with the only p-value less than 0.05 detected for $\theta_{SS\text{-}HI}$ ($p = 0.027$). Pairwise comparisons among inter-annual samples were non-significant for θ ($p > 0.013$) and Φ ($p > 0.17$) when corrected for multiple tests (initial $\alpha = 0.00047$). The proportion of genetic variance due to heterogeneity among locations within oceans, was also small (Table 5). The closest value to significance was $\Phi_{S\text{-}PAC}$ (0.013; $p = 0.036$; initial $\alpha = 0.025$). Pairwise sample comparisons among locations within either ocean were non-significant using both exact probability ($p > 0.13$) and exact Φ_S ($p > 0.074$) testing methods. When temporal samples were pooled at a location, intra-ocean comparisons remained non-significant for exact probability tests ($p > 0.057$).

With exception of AU samples, there was little evidence for clustering of temporal replicates by location when using linearized Φ (Fig. 9D). This shows that the amount of variation detected among locations within oceans was generally no greater than the variance among replicate samples at a location. However, the phenogram does depict the two Australian samples as intermediate in overall similarity to those of either ocean. Accompanying these differences, three Atlantic clade haplotypes were present in the two Australian samples (identified as 'Pacific' in Fig. 2F).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Inter-Ocean Divergence. Mitochondrial RFLP data revealed a highly discordant distribution of haplotypes between samples from the Atlantic and Pacific. The asymmetric distribution of the Atlantic clade between oceans dominated the outcome of all inter-ocean comparisons. Inter-ocean divergence of haplotype frequencies was significant when measured as θ_P and Φ_P, contributing 3% ($p = 0.0003$ +/- 0.0002) and 21% ($p < 0.00001$) of overall genetic variance to the difference between ocean basins, respectively. A large proportion of inter-ocean pairwise sample comparisons (Φ_P) was significant at the 0.01 level (36 out of 56; Fig. 10C; Table 7). In addition, tests between MX94 and six of seven Atlantic populations were significant at the $p < 0.05$ level. Significant differences in allele frequency between ocean samples were barely detectable with exact probability tests (Fig. 10C). MtDNA haplotype frequencies allowed for unambiguous separation of samples into ocean of origin when neighbor-joining dendrograms were constructed using linearized Φ (Fig. 9D).

<u>Correspondence Between Mitochondrial Clades and Nuclear Allele Frequencies</u>

Individuals belonging to the two mitochondrial clades may represent distinct genetic units at the nuclear genome. Exact probability and exact R tests between individuals belonging to alternative mitochondrial clades were performed for each microsatellite locus. These tests were performed for Atlantic individuals only because of the low frequency of Atlantic clade haplotypes in the Pacific. Significant deviation from a random assortment of microsatellite alleles with respect to the two mitochondrial clades was not detected at any locus for either test ($R < 0.011$; $p > 0.22$; Table 9).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Male-Mediated Dispersal

It has been suggested that males may be more migratory than females in the blue marlin (Chapter 2). Tests were conducted to evaluate allele-frequency homogeneity between sexes within oceans, as well as the relative degree of inter-ocean population structure for each sex. To determine whether detectable differences in allele frequency existed between sexes within oceans, individuals were separated by sex and tested for homogeneity by using exact Φ and exact probability tests. The degree of inter-ocean population structure for each sex was evaluated by estimating Φ and θ statistics. The number of individuals for which sex was identified was limited, varying from 19 to 98 among groups (Table 1). Patterns of mitochondrial diversity among males and females within oceans are shown in Table 10. As with the overall analysis, haplotypes were broadly distributed among Atlantic and ubiquitous clades.

Significant differences between sexes in mtDNA haplotype frequency were not detected in either Pacific or Atlantic samples. Inter-ocean mitochondrial divergence (Φ) among females (0.32) was twice that among males (0.167; Table 11). Significance levels for these values were also higher for females both for exact Φ and exact probability tests ($p < 0.00001$) than males (exact $\Phi = 0.0006$; exact probability $p > 0.1$). To evaluate the significance of the difference in inter-ocean divergence between sexes, samples of Atlantic males were drawn of the same size as the observed Atlantic female sample (19), and male inter-ocean divergence recalculated. Seven out of 20 randomizations yielded Φ values of greater than 0.3 between males, indicating that the larger female divergence can be attributed to sampling error.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DISCUSSION

Microsatellites

The microsatellite enrichment protocols produced a large number of clones containing microsatellite repeats. The sequence of 20 microsatellite clones indicated that base pair substitutions within repeat motifs were common: 66% of the sequences were interrupted. At three compound loci, GGTA and GACA tetranucleotide motifs were detected and these motifs may provide additional sources of tetranucleotide microsatellite diversity in the blue marlin genome. The average repeat length of the 20 sequenced clones (28.2) and of the five tetranucleotide loci surveyed over all populations (24.9) was within range of that reported for dinucleotide repeats in teleost fishes, including the rainbow trout (20), *Oncorhynchus mykiss*, Atlantic salmon (24), *Salmo salar*, and Atlantic cod (20-46), *Gadus morhua* (O'Connell and Wright, 1997).

Genotypic distributions for four of five microsatellite loci conformed to expectations of Hardy-Weinberg equilibrium over all populations. GATA10 exhibited significant heterozygote deficiencies in three samples and over all samples combined. This may have resulted from occurrence of null alleles, where primer site mutations prevent alleles from amplifying with equal strength or altogether (Callen et al., 1993; Pemberton et al., 1995). The presence of null alleles has been correlated with tissue quality (O'Connell and Wright, 1997), but for this study, the distribution of Hardy-Weinberg deviations among many samples argues against sample-specific phenomenon.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Significant heterozygote deficiencies at a microsatellite locus were also reported for the Atlantic cod (Bentzen et al., 1996), and in an increasing number of studies with other organisms (Barker et al., 1997; Lehmann et al., 1996; Paetkau and Strobeck, 1995).

Each of the five microsatellite loci analyzed were hypervariable. Heterozygosity values (average = 0.93) were among the highest observed for any species. The average number of tetranucleotide alleles (32.2) was lower than those reported for dinucleotide loci in a number of species, including Atlantic salmon (37), Atlantic cod (46), and Pacific herring (33) (O'Connell and Wright, 1997).

Mutation rates were similar among microsatellite loci in blue marlin. Comparison of loci with the highest (GATA08) and lowest (GATA01) repeat variance yielded a 3.8-fold range in mutation rates. These estimates were generated by employing functions of repeat variance as described in Chakraborty et al. (1997) and assuming a stepwise mutation model, mutation-drift equilibrium, and no difference in mutation rate among populations. The range is lower than that detected in a freshwater snail (*Bulinus truncatus*) among tetranucleotide loci (7-fold; Viard et al., 1996) and among dinucleotide loci in the honeybee (10-fold; Estoup et al., 1995). These studies used the observed number of alleles to estimate $4N_e u$ (Chakraborty and Neel, 1989).

Alleles were highly divergent over all microsatellite loci and populations. Bimodal loci exhibited increased repeat variance over all samples and for individual samples where modes were shared. Among bimodal loci, distance between modes had the greatest effect on total repeat variance. This observation is consistent with Di Rienzo et al. (1994) who reported that large jumps in repeat number generate the most repeat variance. Bimodal allelic distributions have been reported for microsatellite loci in other

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

organisms (FitzSimmons et al. 1997; Edwards, 1992). However, bimodality as pronounced and consistent across loci as seen in blue marlin is rare.

Modal formation could occur under a strict stepwise mutation model (SSM) or a two-phase mutation model (TPM) that includes rare but large multi-step mutations (Di Rienzo et al., 1994). Shriver et al. (1993) showed that a number of modes are expected at a locus under the SSM model and simple mutation-drift equilibrium. However, in this study, criteria for modal definition were restrictive, in order to focus on groups of alleles with possible phylogeographic implication.

Analysis of Molecular Variance. The overall AMOVA demonstrated that the majority of microsatellite DNA genetic variance was present within populations; inter-annual and inter-location subdivision accounted for only a minor proportion of genetic variance. This pattern was consistent for both hierarchical R and θ statistics. In a few cases, p-values for pairwise comparisons of temporal samples at a location approached significance. These latter comparisons likely stemmed from stochastic sampling variance. Temporal stability of microsatellite allele frequencies has been observed in other fishes including Atlantic cod (Bentzen et al., 1996; Ruzzante et al., 1996).

All tests of microsatellite allele frequency homogeneity among samples within oceans were non-significant after corrections for multiple testing. However, the highly migratory nature of blue marlin suggests that much larger sample sizes would be necessary to detect significant intra-ocean population structure should differences exist (Waples, 1998). In contrast, microsatellite analysis of striped marlin detected highly

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

significant population structure among a number of geographic samples from the Pacific (Buonaccorsi and Graves, unpublished data).

Highly significant differences among samples from different oceans were detected at all microsatellite loci. These results strongly indicate that ocean populations of blue marlin represent distinct gene pools with little gene flow between oceans. Differences among loci were due largely to the degree that modes of alleles were shared between oceans. From the divergence (*R*) values, an average of 1.47 effective inter-ocean migrants ($N_e m$) per generation was estimated (standard deviation = 0.7 to 6.7). Values greater than one are large enough to prevent fixed differences from accumulating between oceans, but not enough to prevent significant population subdivision (Allendorf and Phelps, 1981).

MtDNA

In concordance with a previous RFLP analysis of blue marlin mtDNA (100 individuals; Graves and McDowell, 1995), this expanded whole-molecule analysis (500 individuals) revealed high levels of mtDNA divergence within blue marlin. MtDNA haplotypes could be assigned to Atlantic or ubiquitous clades, as with previous whole molecule and cytochrome *b* studies (Graves and Mcdowell, 1995; Finnerty and Block, 1992). The Atlantic clade was more diverse than the ubiquitous clade, sharply increasing the diversity of samples from the Atlantic. Nucleon diversities reported here are larger than those reported by Graves and McDowell (1995) for both Atlantic (0.93) and Pacific (0.69) samples.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Analysis of Molecular Variance. As with microsatellite loci, most mtDNA genetic variation was distributed within populations, and inter-annual and intra-ocean variance components were small. In contrast to microsatellite analyses, θ_S (intra-ocean) was less informative than Φ_S in detecting mtDNA divergence. Pairwise comparisons and phenetic clustering depicted Australian populations as intermediate to Atlantic and Pacific samples. This intermediate position was influenced strongly by the presence of three Atlantic clade individuals within Australian samples. The divergence of these individuals from others in the Pacific enhanced the overall divergence of the AU samples from other Pacific samples. This pattern warrants further sampling in Australian waters to determine accurately the possible mixing of individuals from different oceans. The different distribution of the two clades led to highly significant and pronounced inter-ocean divergence. When corrected for haploid inheritance, the number of effective female migrants was estimated at 1.88 individuals per generation, a value that indicates some gene flow between oceans.

Phylogeography

Recent gene flow between oceans is suggested by the presence of ubiquitous mtDNA clade alleles in both oceans. Contemporary inter-ocean gene flow also was inferred from allele distributions at blue marlin scnDNA loci (Chapter 2). Experimental longline fishing south of the Cape of Good Hope during the austral summer indicated the presence of a number of marlin (Penrith and Cram, 1974; Talbot and Penrith, 1962), providing a possible inter-ocean migratory route. In addition, movement of blue marlin from Atlantic to Indian (Anonymous, 1994), and Pacific to Indian oceans (Julian

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Pepperrell personal communication, 1998) has been inferred from tag and recapture data. However, limited Atlantic to Pacific movement of blue marlin was inferred from the restricted geographic distribution of the large group of Atlantic mtDNA haplotypes (the Atlantic clade). The distribution of the second mode of GATA08 alleles (limited primarily to the Atlantic) supports the notion that there is little current migration from Atlantic to Pacific. For other microsatellite loci, it was more difficult to determine the direction of gene flow, as modes could not be identified unambiguously.

Graves and McDowell (1995) suggested that blue marlin experienced extensive isolation during the Pleistocene, during which divergent clades of mtDNA alleles formed. Climatic reconstructions from 18,000 years ago suggest that austral winter sea-surface temperatures were approximately 14°C at the Cape of Good Hope, at least three degrees colder than present temperatures (CLIMAP, 1976). A decrease in summer temperatures of this magnitude probably limited inter-ocean gene flow, as summer temperatures are currently at the lower end of blue marlin tolerance (Nakamura, 1985; Rivas, 1975; Penrith and Cram, 1974). Using an evolutionary rate of 0.5 to 1.0 % per million years (Wilson et al., 1985), Finnerty and Block (1992) estimated a divergence time of 1.5 to 3.0 million years between the mitochondrial lineages detected using cytochrome *b* data. Using a rate of 2% per million years for the whole mitochondrial genome (Brown et al., 1979), the data of Graves and McDowell (1995) yielded a divergence of 600,000 years.

An 'Atlantic' clade also has been reported in other pelagic fishes, including the bigeye tuna (Alvarado Bremer et al., submitted), swordfish (Alvarado Bremer et al., 1996; Rosel and Block, 1996), and sailfish (Graves and McDowell, 1995). However, the degree of fidelity of the "Atlantic" clade to the Atlantic included values of 100% for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sailfish (Graves and McDowell, 1995) and swordfish (Alvarado Bremer et al., 1996; Rosel and Block, 1996), 99.5% for the blue marlin (this study), and 90% for bigeye tuna (Alvarado Bremer et al., 1997). The frequency of the Atlantic clade within the Atlantic included values of 80% in the sailfish, 73% in bigeye tuna, 40% in blue marlin, and 21% in swordfish (Rosel and Block, 1996). Finally, the amount of sequence divergence between clades estimated from whole molecule RFLP analysis included values of 0.65% for sailfish and 1.2% for blue marlin (Graves and McDowell, 1995), to over 6% for the hypervariable D-loop molecule in bigeye tuna (Alvarado Bremer et al., submitted) and swordfish (Alvarado Bremer et al., 1996). These varying results indicate that the processes leading to clade generation may not have affected all species similarly. However, the presence of highly divergent, predominantly Atlantic, haplotypes across taxa strongly suggests that genetic patterns have been shaped by historical vicariant events.

The same environmental history that led to formation of mitochondrial clades may have produced the bimodality evident at several microsatellite loci. Given tetranucleotide microsatellite mutation rates of 0.001 to 0.0001 mutations per generation (Goldstein et al., 1995), and a generation time for blue marlin of five years, temporal divergence estimates between modes, averaged over five loci, ranged from 580,000 to 5.8 million years. These estimates are consistent with a period of Pleistocene isolation. Accuracy of the divergence estimates is limited by ambiguities concerning the mutational processes that generate microsatellite alleles.

There are at least five reasons that could explain the greater mitochondrial diversity in the Atlantic clade versus the ubiquitous clade. (i) A higher mutation rate in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the Atlantic-clade would enhance its diversity. However, there is little reason to suspect that mutation rates differ between clades.

(ii) A greater coalescence time for the Atlantic clade would have allowed more time for mutations to accumulate. In this scenario, the Atlantic clade would need to be ancestral to the ubiquitous clade, and would require extirpation of the Atlantic clade in the Pacific, since the presence of fossils from the Isthmus of Panama during the late Miocene suggest inter-ocean gene flow at that time (Harry Fierstein, personal communication, 1998). However, one would not necessarily expect drift to have randomly eliminated the Atlantic clade across a number of taxa.

(iii) Larger historical population sizes within the Atlantic Ocean would have allowed a greater accumulation of Atlantic diversity in accordance with mutation-drift equilibrium. However, current stock size estimates from production analyses (Cramer and Prager, 1994; Skillman, 1989) indicate that the Pacific supports a much larger population of blue marlin, and there is little reason to suspect that the Atlantic would have supported greater population sizes in the past.

(iv) A historical population expansion from Atlantic to Pacific would have resulted in a shallow coalescence of Pacific alleles, where Atlantic clade alleles represent remnants. However, this would require substitution of all prior Pacific alleles present before the expansion, by Atlantic alleles.

(v) Greater population fluctuation and bottlenecking, both of which reduce effective population sizes and diminish genetic diversity (Rogers and Harpending, 1992; Slatkin and Hudson, 1991; Shriver et al., 1997), could have occurred to a larger extent in the Pacific. Large-scale, episodic oceanographic fluctuations influencing temperature

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and productivity profiles have been proposed as important factors controlling size of blue marlin populations (Hopper, 1990). Although recent studies propose a connection between tropical oceanographic fluctuations and biological production (Barber and Chavez, 1986; Hayward, 1997), the impact of such fluctuations on large tropical pelagic species requires further investigation. The extremely broad distribution and large population size of blue marlin, along with the presumed stability of the tropical pelagic marine ecosystems (Briggs, 1974) argue against an impact of localized extinctions on overall genetic diversity.

In contrast to mtDNA, microsatellite loci did not provide evidence for greater diversity within Atlantic-derived modes. The Atlantic clade of GATA08 was less diverse than the ubiquitous clade, and at GATA60 and GATA90, modal diversity was generally higher in the Pacific. The presence of relatively diverse clades in the Pacific suggests that Pacific populations have not undergone severe fluctuations in population size. Further investigation into variability patterns from these and other nuclear loci may help elucidate the causes underlying historically greater mitochondrial Atlantic variance. Relative to mtDNA, less pronounced evidence for historical population bottlenecks also was observed at microsatellite loci in other species (Kolman and Bermingham, 1997; Bentzen et al., 1996).

Sex-Biased Dispersal.

This study detected two to four-fold greater mitochondrial than nuclear inter-ocean divergence in blue marlin, suggesting the possibility of greater male dispersal. Sexes may differ in the degree of genetic divergence between oceans if males are more

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

migratory. Inter-ocean mitochondrial gene flow was estimated to be twice as large for males than females. However, this difference could be explained from sampling alone. Furthermore, differences were not detected in allele frequency between sexes within either ocean. Larger sample sizes, in particular of Atlantic females, are needed to reduce sampling error in these analyses. A further limitation of these tests is that migrant males do not pass on their mtDNA, such that only single-generation migrant males contribute to the difference among sexes. The probability of sampling migrants, given low overall levels of gene flow between oceans, is small.

Uni-Directional Gene Flow.

The restricted geographic distribution of alleles at GATA08 and mtDNA (Atlantic clades) suggests that there is little migration from Atlantic to Pacific. If Atlantic-clade alleles present in the Pacific represent migrants, their relative frequencies in both oceans could reflect contemporary gene flow. Such a conclusion is sensitive to the level of support for allelic phylogenies, and the placement of 'migrant' alleles was stable with respect to clade assignment in both neighbor-joining and UPGMA clustering. Occurrence of Atlantic-clade individuals in Australian waters is consistent with the intermediate geographical position of Australia and an African migration corridor. Further sampling of blue marlin in the Indian Ocean is necessary to examine the role of Australia as a mixing zone. The two "Atlantic mode" alleles of GATA08 present in the Pacific were widely distributed, also consistent with the hypothesis of contemporary mixing between oceans.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Test for independence of mtDNA clades and microsatellite allele frequencies

It is feasible that nuclear alleles are not randomly distributed among individuals displaying alternative mitochondrial clades. Significant association between mitochondrial clades and nuclear alleles was not detected in Atlantic samples. This indicates that migrants from the Pacific are spawning with local individuals. The result is consistent with expectations of two unlinked markers within a single randomly mating Atlantic population.

Statistical Considerations and Mutation, Migration, and Drift Equilibrium

Values of R and θ in tests of population structure potentially assess the relative impact of mutation and drift, respectively, on genetic divergence (Slatkin, 1995). For loci with low mutation rates, and for populations with lower divergence times and/or larger numbers of migrants, divergence consists of frequency differences at individual alleles resulting from genetic drift. The θ statistic is sensitive to individual allele frequency shifts, thus identifying the influence of drift. For loci with high mutation rates and/or populations with larger divergence times, microsatellite divergence includes the generation of new modes of alleles through mutation. R is sensitive to changes in mean repeat number, a parameter affected by modal generation.

Modes at microsatellite loci were distributed similarly among samples within oceans. At this level of analysis, divergence consisted of drift of individual alleles, which is likely why θ produced more significant divergence estimates than R. For a number of other empirical studies, θ was more sensitive than R in detection of divergence among closely related populations (FitzSimmons et al., 1997; Forbes et al., 1995; Perez-Lezaun

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

et al., 1997). However, both mutation and drift appear to have important roles in genetic divergence between blue marlin in the two oceans. For three loci (GATA60, GATA08, and GATA90) inter-ocean divergence consisted of the differential distribution of divergent modes of alleles. Moderate to highly divergent R values resulted from shifts in mean repeat number between oceans, typically indicating the influence of mutation in diverging populations. θ detected weaker divergence at these loci, as individual frequency differences were small. For unimodal loci, however, the impact of drift was evident, as θ was more sensitive than R to inter-ocean population structure.

Mutation likely played a role in the divergence of Atlantic and Pacific blue marlin during the Pleistocene, through generation of modes of over an extended period of allopatry. With recent mixing, however, the Pleistocene-derived modes would be distributed between oceans according to a new migration-drift equilibrium.

To further investigate the role of mutation and drift in estimates of inter-ocean divergence, modes were binned into single allele classes and inter-ocean divergence recalculated. If modes formed during historical periods of isolation, the genetic divergence (number of steps) between modes would not be relevant to current estimates of gene flow. Binned divergence estimates measure the degree of inter-ocean divergence based on modal frequencies alone. Binned divergence values for GATA08, GATA60, and GATA90 were 0.164, 0.245, and 0.093, respectively, and were similar to estimates based on R (0.184, 0.269, and 0.130, respectively). The similarity in values indicates that the degree of mode sharing between ocean samples (determined by drift) and not distance of separation between modes (determined by mutation), was most important in determining R values, and hence the description of current gene flow among oceans. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

R statistic effectively linked related groups of alleles according to the strict stepwise mutation model, and subsequently described their distribution among oceans. Numbers of migrants (N_em) from the R values (1.47) were low relative to values of $N_e\mu$ (10.5), indicating that at equilibrium, mutation may accumulate divergence between oceans faster than migration can disperse it. The close match of R and modal binning estimates also confirms significance of the modal comparisons used throughout this study.

For mtDNA, binning of alleles into clades similarly indicated that drift has likely had a more important role than mutation in contributing to current inter-ocean divergence. Divergence based on estimators that linked alleles by restriction site differences (Φ) performed better than θ in detecting inter-ocean population structure. This was evident when comparing overall and pairwise inter-ocean divergence estimates and clustering results for the two methods. As with microsatellite modes, if the divergent mitochondrial clades formed under historical conditions, the distance between clades would not be relevant to measures of current inter-ocean gene flow. Based on frequency alone, and where the Atlantic clade occurs in 40% of Atlantic and 0.5% of Pacific individuals, a θ of 0.387 results, yielding an N_em_f of 0.79. This estimate is larger than that which incorporated restriction site differences between alleles and shows that incorporation of mutational differences actually underestimated mitochondrial divergence. It is possible that incorporation of allelic relationships inferred from construction of a parsimony network may have described more accurately mitochondrial divergence (Excoffier et al., 1992). Mitochondrial N_em_f values (0.79) were lower than the $N_e\mu$ estimate (2.2), indicating that over a period of time, and assuming

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

migration/mutation-drift equilibrium, mutation may play a role in the differentiation of mtDNA between oceans.

This chapter focused on the population genetics of hypervariable nuclear and mitochondrial markers within blue marlin. Overall AMOVAs for both marker classes showed that both inter-annual and intra-ocean variance comprised only a minor proportion of overall genetic variance. Inter-annual variance in allele frequencies approached significance in a number of comparisons, however, indicating that a substantial noise level must be accounted for in comparisons of samples among locations within oceans. Between oceans, both mitochondria and microsatellite loci displayed striking patterns of genetic divergence. Inter-ocean divergence and gene flow estimates for mtDNA, GATA08, GATA60, and GATA90 were described adequately simply by the current distribution of ancient clades of alleles. The role of genetic drift in the current distribution of alleles was implicated in the recent mixing of clades among oceans. This process is explored further in the next chapter, which focuses on the comparison of allozyme, scnDNA, mtDNA, and microsatellite markers. Comparisons are made with respect to sensitivity to population structure at intra-ocean and inter-ocean levels, and regarding the possible role of selection, drift, migration, and mutation on the distributions of molecular diversity for each marker class.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LITERATURE CITED

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38:1507-1514.

Alvarado Bremer, J. R., J. Mejuto, T. W. Greig, and B. Ely. 1996. Global population structure of the swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial DNA control region. J. Exper. Mar. Biol. Ecol. 197:295-310.

Alvarado Bremer, J. R., B. Stequert, N. W. Robertson, and B. Ely. Submitted. Genetic evidence for inter-ocean subdivision of bigeye tuna (*Thunnus obesus* Lowe) populations.

Angers, B., and L. Bernatchez. 1997. Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. Mol. Biol. Evol. 14:230-238.

Anonymous. 1994. Cooperative Game Fish Tagging Program Annual Newsletter: 1992. Southeast Fisheries Science Center, Miami, FL.

Bairoch, A. 1989. PC/GENE. Intelligenetics, Inc., University of Geneva, Switzerland.

Barber, R. T., and F. R. Chavez. 1986. Ocean variability in relation to living resources during the 1982-83 El Nino. Nature. 319:279-285.

Barker, J. S., S. S. Moore, D. J. Hetzel, D. Evans, S. G. Tan, and K. Byrne. 1997. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein-coding loci. Anim. Genet. 28:103-115.

Bentzen, P., C. T. Taggart, D. E. Ruzzante, and D. Cook. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences. 53:2706-2721.

Birky, C. W., T. Maruyama, and P. Fuerst. 1983. An approach to population genetic and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics. 103:513-527.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd, and L. L.

Briggs, J. C. 1974. Marine Zoogeography. McGraw-Hill, Inc., USA.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Brown, W. M., M. George Jr., and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences, USA. 76:1967-1971.

Callen, D. F., A. D. Thompson, Y. Shen, H. A. Phillips, R. I. Richards, J. C. Mulley, and G. R. Sutherland. 1993. Incidence and origin al "null" alleles in the (AC)n microsatellite markers. Am. J. Hum. Genet. 52:922-927.

Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature. 368:455-457.

Chakraborty, R., M. Kimmel, D. N. Stivers, L. J. Davison, and R. Deka. 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences, USA. 94:1041-1046.

Chakraborty, R., and J. V. Neel. 1989. Description and validation for a method for simultaneous estimation of effective population size and mutation rate from human population data. Proc. Nat. Acad. Sci. U. S. A. 86:9407-9411.

CLIMAP. 1976. The surface of the ice-age earth. Science. 191:1131-1137.

Cramer, J., and M. H. Prager. 1994. Exploratory surplus-production analysis of blue and white marlin fisheries in the North Atlantic. Pp. 557-560 *in* ICCAT, ed. Collective volume of scientific papers. Vol. XLI. Report of the second ICCAT billfish workshop. ICCAT, Madrid, Spain.

Edwards. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics. 12:241-253.

Estoup, A., L. Garnery, M. Solignac, and J. M. Cornuet. 1995. Microsatellite variation in honey bee (Apis mellifera L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics. 140:679-95.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics. 131:479-491.

Finnerty, J. R., and B. A. Block. 1992. Direct sequencing of mitochondrial DNA detects highly divergent haplotypes in blue marlin *(Makaira nigricans)*. Mol. Mar. Biol. Biotechnol. 1:206-214.

FitzSimmons, N. N., C. Moritz, C. J. Limpus, L. Pope, and R. Prince. 1997. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. Genetics. 147:1843-1854.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Forbes, H. F., J. T. Hogg, F. C. Buchanan, A. M. Crawford, and F. W. Allendorf. 1995. Microsatellite evolution in congeneric mammals: Domestic and bighorn sheep. Mol. Biol. Evol. 12:1106-1113.

Goldstein, D. B., A. Ruiz Linares, L. L. Cavalli-Sforza, and M. W. Feldman. 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Nat. Acad. Sci. U. S. A. 92:6723-6727.

Graves, J. E., and J. R. McDowell. 1995. Inter-ocean genetic divergence of Istiophorid billfishes. Marine Biology. 122:193-203.

Guo, S. W., and E. A. Thompson. 1993. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometics. 48:361-372.

Hayward, T. L. 1997. Pacific Ocean climate change: atmospheric forcing, ocean circulation and ecosystem response. Trends Ecol. & Evol. 12:150-154.

Hopper, C. N. 1990. Patterns of Pacific blue marlin reproduction in Hawaiian waters. Pp. 29-39 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part II. Contributed Papers. National Coalition for Marine Conservation Inc., Savannah, GA.

Jin, L., and R. Chakraborty. 1995. Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. Heredity. 74:274-285.

Kijas, J. M. H., J. C. S. Fowler, C. A. Garbett, and M. R. Thomas. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques. 16:656-662.

Kolman, C. J., and E. Bermingham. 1997. Mitochondrial and nuclear DNA diversity in the Choco and Chibcha Amerinds of Panama. Genetics. 147:1289-1302.

Lansman, R. A., R. O. Shade, J. F. Shapira, and J. C. Avise. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications. J. Mol. Evol. 17:214-226.

Lehmann, T., W. A. Hawley, and F. H. Collins. 1996. An evaluation of evolutionary constraints on microsatellite loci using null alleles. Genetics. 144:1155-1163.

McConnell, S. K., P. H. O'Reilly, L., J. M. Wright, and P. Bentzen. 1995. Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North American and European populations. Can. J. Fish. Aquat. Sci. 53:1863-1872.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

McElroy, D. P., P. Moran, E. Bermingham, and I. Kornfield. 1991. REAP: The restriction enzyme analysis package. University of Maine, Orono.

Michalakis, Y., and L. Excoffier. 1996. A generic estimation ofpopulation subdivision using distances between alleles with special reference to microsatellite loci. Genetics. 142:1061-1064.

Nakamura, I. 1985. Vol 5: Billfishes of the world: An annotated and illustrated catalogue of marlins, sailfishes, spearfishes, and swordfishes known to date. FAO Fish. Synop.

Nauta, M. J., and F. J. Weissing. 1996. Constraints on allele size at microsatellite loci: implications for genetic differentiation. Genetics. 143:1021-32.

Nei, M., and J. C. Miller. 1990. A simple method for estimating the average number of nucleotide substitutions within and between populations from restriction data. Genetics. 125:873-879.

O'Connell, M., and J. M. Wright. 1997. Microsatellite DNA in fishes. Rev. Fish Biol. Fish. 7:331-363.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. 4:347-354.

Paetkau, D., and C. Strobeck. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. Mol. Ecol. 4:519-520.

Pemberton, J. M., J. Slate, D. R. Bancroft, and J. A. Barrett. 1995. Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol. Ecol. 4:249-252.

Penrith, M. J., and D. L. Cram. 1974. The Cape of Good Hope: A hidden barrier to billfishes. Pp. 175-187 *in* R.S. Shomura and F. Williams, eds. Proceedings of the International Billfish Symposium, Kailua-Kona, HI, 9-12 August 1972. Part 2. Review and contributed papers. U.S. Dep. Commer., NOAA Tech. Rep. NMFS SSRF-675, Seattle, WA.

Perez-Lezaun, A., F. Calafell, E. Mateu, D. Comas, R. Ruiz-Pacheco, and J. Bertranpetit. 1997. Microsatellite variation and the differentiation of modern humans. Hum. Genet. 99:1-7.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. J. Hered. 86:248-249.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reynolds, J., B. S. Weir, and C. C. Cockerham. 1983. Estimation for the coancestry coefficient: basis for a short-term genetic distance. Genetics. 105:767-779.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution. 43:223-225.

Rivas, L. R. 1975. Synopsis of biological data on blue marlin, *Makaira nigricans* (Lacepede), 1802. Pp. 1-27 *in* R.S. Shomura and F. Williams, eds. Proceedings of the International Billfish Symposium, Kailua-Kona, HI, 9-12 August 1972. Part 3. Species Synopses. U.S. Dep. Commer., NOAA Tech. Rep. NMFS SSRF-675, Seattle, WA.

Rogers, A. R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 93:552-559.

Rosel, P. E., and B. A. Block. 1996. Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. Marine Biology. 125:11-22.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics. 142:1357-62.

Ruzzante, D. E., C. T. Taggart, and D. Cook. 1996. Spatial and temporal variation in the genetic composition of a larval cod (*Gadus morhua*) aggregation: cohort contribution and genetic stability. Can. J. Fish. Aquat. Sci. 53:2695-2705.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstruction of phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Sambrook, J., E. F. Fritsch, and T. Maniatus. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, USA.

Schneider, S., J. Kueffer, D. Roessli, and L. Excoffier. 1997. Arlequin. A software for population genetic data analysis., Geneva.

Shriver, M. D., L. Jin, R. Chakraborty, and E. Boerwinkle. 1993. VNTR allele frequency distributions under the stepwise mutation model: A computer simulation approach. Genetics. 134:983-993.

Shriver, M. D., L. Jin, R. E. Ferrell, and R. Deka. 1997. Microsatellite data support an early population expansion in Africa. Genome Res. 7:586-591.

Skillman, R. A. 1989. Status of Pacific billfish stocks. Pp. 179-195 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part I. Fishery and stock synopses, data needs and management. National Coalition for Marine Conservation Inc., Savannah, GA.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics. 139:457-462.

Slatkin, M., and R. R. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. 129. :555-562.

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical Taxonomy. W. H. Freeman, San Fransisco.

Talbot, F. H., and M. J. Penrith. 1962. Tunnies and marlins of South Africa. Nature. 193:558-559.

Tautz, D., M. Trick, and G. A. Dover. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature. 322:652-656.

Viard, F., P. Bremond, R. Labbo, F. Justy, B. Delay, and P. Jarne. 1996. Microsatellites and the genetics of highly selfing populations in the freshwater snail *Bulinus truncatus*. Genetics. 142:1237-1247.

Waldbieser, G. C. 1995. PCR-based identification of AT-rich tri- and tetranucleotide repeat loci in an enriched plasmid library. Biotechniques. 19:742-744.

Waples, R. S. 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. J. Hered. In Press

Weber, J., and C. Wong. 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2:1123-1128.

Weir, B. 1996. Genetic data analysis II. Methods for discrete population genetic data. Sinauer Associates, Inc., Sunderland, Massachusetts.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358-1370.

Wilson, A. C., R. L. Cann, S. M. Carr, M. George, and U. B. Gyllenstein. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26:375-400.

Wright, J. M., and P. Bentzen. 1994. Microsatellites: genetic markers for the future. Rev. Fish Biol. Fish. 4:384-388.

Wright, S. 1978. Evolution and the genetics of populations. Vol. 4., Variability in and among natural populations. Univ. Chicago Press, Chicago.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1. Blue marlin samples from mtDNA, microsatellite (MICROS), and scnDNA marker classes. Levels for AMOVA were ocean, location, and year. 'N' refers to the number of individuals sampled.

OCEAN	LOCATION		YEAR	Abbrev.	MtDNA[1] N	MICROS N	ScnDNA N
ATLANTIC	Puerto Rico	Puerto Rico	1991	PR91	31		
	Jamaica	Jamaica	1991	JM91	28	54	48
		Jamaica	1992	JM92	54	54	55
		Jamaica	1993	JM93	18	42	42
		Jamaica	1994	JM94	43	46	45
		Jamaica	1995	JM95	21	24	24
		Total			164	220	214
	Eastern USA	New Jersey	1994	NJ94	9	11	11
		N. Carolina	1992	NC92	12	12	12
		Total			21	23	23
		Atlantic Total			216	243	237
PACIFIC	Hawaii	Hawaii	1990	HI90	26		
		Hawaii	1992	HI92	32	34	30
		Hawaii	1993	HI93	43	45	48
		Hawaii	1994	HI94	28	25	25
		Total			129	104	103
	E. Pacific	Mexico	1994	MX94	11	48	49
		Mexico	1995	MX95	23	24	24
		Ecuador	1995	EC95	19	19	19
		Total			53	91	92
	Australia	Australia	1991	AU91	10	10	10
		Australia	1994	AU94	14	16	15
		Total			24	26	25
		Pacific Total			206	221	220
		Overall Total			422	464	457

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1 (Continued). Blue marlin samples from mtDNA, microsatellite (MICROS), and scnDNA marker classes.

		MtDNA[1] N
Atlantic	Females	19
	Males	100
Pacific	Females	55
	Males	21

[1] MtDNA samples from Graves and McDowell, unpublished data.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2. Microsatellite clone sequences.

CLASSIFICATION	LOCUS	# REPEATS	REPEAT CONTENT
Perfect	GATA10	19	(GATA)19
	GATA22	31	(GATA)31
	GATA23	11	(GATA)11
	GATA52	15	(GATA)15
	GATA53	17	(GATA)17
	GATA62	19	(GATA)19
	GATA96	32	(GATA)32
	GATA08*	17	(GATA)17
			(GATA)7 GATG (GATA)9
Imperfect	GATA90	29	(GATA)6 GACA (GATA)22
	GATA29	24	(GATA)6 GACA (GATA)17
	GATA61	24	(GATA)2 GACA (GATA)22
	GATA57	28	(GATA)3 GCTA (GATA)2 GCTA (GATA)21
	GATA01	26	(GATA)31 GATT (GATA)2 GATT (GATA)3
	GATA89	32	(GATA)18 GATC (GATA)5 GATC GATA GATC (GATA)5
	GATA60	31	(GATA)4 GCTA GATA GCTA (GATA)5 GCTA (GATA)18
	GATA04	43	(GATA)20 GACA GATA GACA (GATA)6 GACA GATA GACA (GATA)3 GACA GATA GATT (GATA)5
Compund	GATA105	26	(GGTA)10 (GATA)16
	GATA92	28	(GATA)7 (GATG)8 (GATA)5 (GACA)8
Compund Imperfect	GATA47	33	(GATA)12 AATA (GATA)15 (GACA)4 GATA
	GATA14	53	(GATA)17 GACA (GATA)10 (GACA)2 (GATA)9 (GACA)3 (GATA)8 (GACA)2 GATA
	GATA83	55	GATA (GACA)14 (GATA)3 (GATA)19 (GACA)5 (GATA)3 GACA GATA (GACA)3 (GATA)2 GACA (GATA)2

* The two sequences reported for GATA08 represent two alleles at the same locus.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3. *f* values, representing the deviation of genotypic proportions from Hardy-Weinberg equilibrium for each sample. Positive values indicate heterozygote deficiency. Bold refers to significant *f* values after Bonferroni correction. Sample abbreviations follow table 1.

f	GAT01	GAT08	GAT10	GAT60	GAT90		BM47	BM81	WM08	BM32-2	
AU91	0.27	-	-	-	0.11		-	-0.33	-0.67	-	
AU94	-	0.12	-	-	-		-	0.45	-0.19	0.38	
HI92	-	-	**0.16**	-	-		-	0.18	-	-	
HI93	-	-	0.13	-	-		-	0.15	-0.15	-0.12	
HI94	-	-	0.17	-	-		0.34	-	-	-0.43	
MX94	-	-	-	-	-		-	**-0.39**	-0.17	-	
MX95	-	-	**0.39**	-	-		-	-	-	-0.11	
EC95	-	-	0.12	-	-		-	-	-	0.18	
JM91	-	-	-	-	-		0.15	0.24	-0.26	0.15	
JM92	-	-	**0.06**	-	-		-	-0.11	-0.18	-	
JM93	-	-	0.18	-	-		0.14	-	-	-	
JM94	-	-	0.13	-	-		-	-	0.13	-	
JM95	-	-	-0.12	-	-		-0.15	-0.21	-0.24	-	
NC92	-	-	-	0.15	-		-	-0.33	-	-0.29	
NJ94	0.22	-	0.13	-	-		-	0.13	-0.20	0.27	
						Overall					Overall
Mean	0.02	0.01	**0.11**	-0.01	0.00	0.027	0.00	-0.01	-0.12	0.01	-0.032
Variance	0.01	0.00	0.01	0.00	0.00	0.006	0.01	0.06	0.04	0.04	0.037

Table 4. Single locus measures of diversity for Atlantic (ATL) and Pacific (PAC) samples of blue marlin, and for groups of alleles comprising modes (microsatellites) or clades (mtDNA). 'n' refers to the total number of allele counts. '*H*' refers to expected heterozygosity (nuclear loci), or the probability that two randomly drawn haplotypes will be different (mtDNA). 'Modal Freq.' refers to frequency of alternative modes or clades within sample indicated. Repeat number variance (*Var.*) and standard deviation (*Stdev.*), and mean of the mismatch distribution are indicated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Locus	Group	Class	n	Modal Freq.	H	#alleles	Repeat # Mean	Repeat # Var.	Repeat # Stdev.	Mismatch Mean
GATA01		ATL	500		0.93	24	34.4	19.7	4.4	4.8
		PAC	432		0.92	21	33.5	13.8	3.7	4.2
		Over All	932		0.93	27	34.0	17.1	4.1	4.5
GATA10		ATL	494		0.90	22	16.9	31.2	5.6	4.7
		PAC	436		0.90	18	16.0	13.6	3.7	3.4
		Over All	930		0.91	28	16.4	23.0	4.8	4.8
GATA08	ATL	1-36	380	0.819	0.94	24	20.3	27.5	5.2	6.0
		40-58	84	0.181	0.94	18	48.4	17.0	4.1	4.8
		Over All	464		0.96	42	25.5	144.4	12.0	12.5
	PAC	1-36	413	0.995	0.93	26	19.0	22.8	4.8	5.2
		40-58	2	0.005	-	2	-	-	-	-
		Over All	415		0.93	28	19.2	27.0	5.2	5.4
	Over Oceans	1-36	793	0.902	0.94	29	19.6	25.5	5.0	5.7
		40-58	86	0.098	0.94	18	48.4	17.1	4.1	4.8
		Over All	879		0.95	47	22.5	98.4	9.9	9.7
GATA60	ATL	5-30	374	0.744	0.86	11	26.1	7.6	2.8	2.5
		31-53	129	0.256	0.88	11	34.2	6.8	2.6	2.9
		Over All	503		0.92	22	28.2	20.1	4.5	4.6
	PAC	5-30	151	0.368	0.86	10	26.9	4.1	2.0	2.3
		31-53	259	0.632	0.91	19	35.5	13.2	3.6	3.9
		Over All	410		0.94	29	32.4	27.1	5.2	5.8
	Over Oceans	5-30	525	0.575	0.87	11	26.5	5.9	2.4	2.4
		31-53	388	0.425	0.90	19	34.9	9.9	3.1	3.6
		Over All	913		0.94	30	30.2	27.8	5.3	5.7
GATA90	ATL	8-28	413	0.892	0.92	18	18.2	13.3	3.6	4.2
		29-42	50	0.108	0.85	9	33.6	3.5	1.9	2.1
		Over All	463		0.94	28	19.9	35.0	5.9	6.3
	PAC	8-28	299	0.717	0.92	21	19.6	19.0	4.4	4.9
		29-42	118	0.283	0.89	12	33.1	6.2	2.5	2.8
		Over All	417		0.95	33	23.4	52.4	7.2	8.2
	Over Oceans	8-28	712	0.809	0.93	21	18.9	16.0	4.0	4.6
		29-42	168	0.191	0.88	12	33.3	4.8	2.2	2.6
		Over All	880		0.95	33	21.5	46.4	6.8	7.5
Overall* Microsatellite		ATL	486	0.524	0.93	27.6	25.0	50.1	7.1	6.6
		PAC	442	0.476	0.93	25.8	24.9	26.8	5.2	5.4
		Over All	928		0.94	33	24.9	42.5	6.5	6.4
MtDNA	ATL	ACLD	90	0.423	0.98	60	-	-	-	5.3
		UCLD	123	0.577	0.81	37	-	-	-	1.7
		Over All	213		0.93	97	-	-	-	6.1
	PAC	ACLD	3	0.015	-	3	-	-	-	-
		UCLD	200	0.985	0.76	39	-	-	-	1.7
		Over All	203		0.77	42	-	-	-	1.8
	Over Oceans	ACLD	93	0.224	0.97	63	-	-	-	5.2
		UCLD	323	0.776	0.79	64	-	-	-	1.7
		Over All	416		0.87	127	-	-	-	4.5
			n		H	#alleles	Mean (Repeat #)	Var. (Repeat #)	Stdev. (Repeat #)	Mean Mismatch

* Averages over loci.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5. Diversity by year and location. Sample abbreviations follow Table 1. *H* refers to expected heterozygosity. Also shown are the number of alleles (# ALL), repeat variance (*Var*), and repeat mean (*Mean*). Bolded values are discussed in text. *h* refers to nucleon diversity for mtDNA. Average (Avg) and total (Tot) values are shown over all loci.

Locus	GATA01				GATA08				GATA10				GATA60				GATA90				OVERALL				Repeat #				MtDNA		
			Repeat #				Repeat #				Repeat #				Repeat #				Repeat #		avg	tot	avg	tot	avg	tot	avg				Mismatch
Taxon	*H*	*# All*	*Var*	*Mean*	*H*	*# All*	*Var*	*Mean*	*H*	*# All*	*Var*	*Mean*	*H*	*# All*	*Var*	*Mean*	*H*	*# All*	*Var*	*Mean*	*H*	*H*	*# All*	*# All*	*Var*	*Var*	*Mean*	*Taxon*	*h*	*#All*	*Mean*
AU91	0.79	7	14	33.4	0.83	8	8.3	19.5	0.75	8	11	**13.2**	0.90	12	38	**34.1**	0.91	12	75	25.2	0.84	0.95	9.4	28	29.2	96.7	25.4	AU91	0.93	8	4.0
AU94	0.90	13	13	32.9	0.89	14	28	**17.3**	0.86	12	29	16.3	0.91	16	24	31.6	0.86	15	35	**21**	0.88	0.96	14	35	25.8	76.6	23.9	AU94	0.92	10	3.0
AU	0.90	14	13	33.1	0.90	15	24	17.9	0.84	13	25	15.2	0.92	20	29	32.5	0.90	21	51	22.4	0.89	0.96	16.6	38	28.3	83.2	24.4	AU		14	3.3
																												HI90	0.59	5	0.8
HI92	0.91	17	15	33.8	0.88	14	35	20	0.89	14	19	16.2	0.92	18	24	32.3	0.94	23	57	24.8	0.91	0.96	17.2	35	30.2	77.1	25.5	HI92	0.81	16	2.0
HI93	0.90	16	12	33.2	0.93	20	26	18.6	0.88	14	8.3	15.8	0.92	18	19	32.8	0.94	27	50	22.4	0.92	0.96	19	37	23	75.1	24.6	HI93	0.68	14	1.5
HI94	0.90	14	11	34.1	0.92	18	26	19.6	0.90	14	23	16.6	0.94	23	41	31.9	0.93	20	44	23.8	0.92	0.96	17.8	37	29.1	75.1	25.2	HI94	0.86	13	2.0
HI	0.91	19	13	33.6	0.92	22	29	19.3	0.90	19	15	16.1	0.94	24	26	32.4	0.95	30	51	23.5	0.92	0.96	26.8	75.7	22.8	39	25	HI			1.5
MX94	0.92	16	15	33.9	0.93	23	30	18.9	0.90	15	9.4	16.1	0.94	23	31	31.9	0.95	26	53	23.4	0.93	0.97	20.6	39	27.8	77.1	24.9	MX94	0.89	7	1.7
MX95	0.91	16	16	32.5	0.89	16	19	19.9	0.87	12	9.8	16.1	0.90	16	22	32.4	0.94	22	54	22.9	0.90	0.96	16.4	34	24.5	68.2	24.8	MX95	0.73	10	2.2
EC95	0.90	13	14	33.7	0.91	19	23	19.8	0.87	10	6.5	15.9	0.92	17	29	33	0.93	21	56	24.8	0.91	0.96	16	36	25.7	75.6	25.5	EC95	0.719	8	1.1
EPAC	0.93	19	15	33.5	0.92	25	26	19.4	0.89	15	8.8	16	0.94	26	28	32.3	0.95	29	54	23.6	0.93	0.96	22.8	41	26.5	74.3	25	EPAC			1.7
																												PR	0.96		6.4
JM91	0.91	20	24	34.3	0.94	27	131	24.7	0.90	17	31	16.9	0.91	18	25	27.4	0.93	20	36	20.4	0.92	0.97	20.4	45	49.4	85.9	24.7	JM91	0.87	14	6.1
JM92	0.93	22	28	34.1	0.94	27	141	24.3	0.93	22	51	18	0.91	18	18	28.4	0.92	22	32	19.3	0.93	0.97	22.2	43	54	87.3	24.9	JM92	0.92	32	6.2
JM93	0.92	18	15	34.4	0.95	30	136	26	0.89	15	26	16.6	0.91	19	22	28.6	0.94	22	35	19.6	0.92	0.97	20.8	43	46.8	86.9	25	JM93	0.94	13	5.6
JM94	0.92	19	15	34.4	0.95	29	170	26.9	0.88	15	17	16	0.91	19	19	28.1	0.92	20	36	19.9	0.92	0.97	20.4	43	51.7	93.5	25.1	JM94	0.93	26	6.2
JM95	0.91	15	19	34.5	0.94	24	103	24.1	0.84	13	16	15.8	0.92	17	19	28.6	0.93	19	37	20.2	0.91	0.96	17.6	38	38.8	80.9	24.7	JM95	0.95	13	5.7
JM	0.93	26	21	34.3	0.95	41	140	25.3	0.90	25	31	16.8	0.92	24	21	28.2	0.93	27	35	19.8	0.93	0.97	28.6	51	49.2	87.4	24.9	JM			6.0
NC92	0.89	11	13	34.6	0.92	16	152	**29.1**	0.86	12	56	17.5	0.84	10	12	28.8	0.90	14	30	19.6	0.88	0.96	12.6	34	52.5	92.1	25.9	NC92	1	12	6.5
NJ94	0.88	11	11	34.8	0.90	15	231	25.5	0.89	11	17	17.7	0.88	12	22	28.2	0.90	13	46	**21**	0.89	0.96	12.4	35	65.4	98.3	25.4	NJ94	0.94	7	7.3
US	0.91	13	12	34.7	0.94	25	189	27.4	0.89	14	37	17.6	0.90	14	16	28.5	0.92	19	37	20.2	0.91	0.97	17	42	58.1	94.7	25.7	US			6.8

Table 6A. AMOVA showing each variance component (σ^2) as a fraction of overall variance. σ_P, σ_S, σ_{SS}, σ_I, and σ_G refer to the components of variation due to the division of oceans, locations within oceans, years at a location, individuals within years, and alleles (genes) within individuals, respectively.

ScnDNA						Microsatellite						
	BM47	BM81	BM32-2	WM08	Overall	GATA01	GATA08	GATA10	GATA60	GATA90	Overall	MtDNA
% Total	-0.305	5.164	10.244	12.261	8.554	2.027	18.355	1.242	26.928	13.055	14.535	20.910
σ^2_P	0.000	0.023	0.053	0.063	0.138	0.351	19.879	0.288	8.669	6.310	31.205	0.529
% Total	0.000	-1.138	-0.010	-0.811	-0.557	-0.107	0.000	0.082	-0.005	0.102	-0.002	-0.148
σ^2_S	0.000	-0.005	0.000	-0.004	-0.009	-0.019	0.000	0.019	-0.002	0.049	-0.005	-0.004
% Total	0.555	0.842	-0.387	1.157	0.520	-0.258	0.000	0.198	-0.013	0.226	-0.006	-0.767
σ^2_{SS}	0.001	0.004	-0.002	0.006	0.008	-0.045	0.000	0.046	-0.004	0.109	-0.013	-0.019
% Total	3.471	-1.366	-0.483	-7.808	-2.350	6.445	-9.787	40.844	0.454	-2.666	-3.815	
σ^2_I	0.005	-0.006	-0.003	-0.040	-0.038	1.116	-10.599	9.464	0.146	-1.289	-8.190	
% Total	96.279	96.497	90.646	95.201	93.833	91.893	91.433	57.633	72.636	89.284	89.289	80.006
σ^2_G	0.139	0.424	0.469	0.485	1.517	15.911	99.026	13.354	23.383	43.157	191.691	2.024

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6B. Measures of heterogeneity among blue marlin samples. θ refers to Weir and Cockerham's (1984) F_{ST} estimator, R refers to the *Rho* statistic of Rousset (1996), and Φ refers to the mitochondrial estimator of Excoffier et al. (1992) with restriction site information considered. Subscripts P, S, and SS refer to the effects of oceans, locations within oceans, and years at a location, respectively. p values refer to the probability $F = 0$. Significance of θ was evaluated using exact probability tests. Significance of R was evaluated using exact R permutation tests. '-' indicates p values > 0.1, F-statistics which are negative, or F-statistics which are < 0.001. Sample abbreviations follow Table 1. The average difference in mean repeat number between ocean samples (*Avg. Diff. In Rpt # Means*) and Delta mu (*Dmu*) statistic of Goldstein et al. (1995) are shown.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LOCUS		GATA01		GATA08		GATA10		GATA60		GATA90		Overall		MtDNA		MtDNA
Sample		*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	
JM	*θ SS-JM*	0.004	0.09	0.001	-	0.002	-	-	-	0.003	-	-	-	0.002	-	*θ SS-JM*
	R SS-JM	-	-	0.002	-	0.015	0.063	0.002	-	-	-	0.005	-	-	-	*Φ SS-JM*
US	*θ SS-US*	0.007	-	0.01	-	-	-	0.031	-	0.013	-	0.012	-	0.009	-	*θ SS-US*
	R SS-US	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*Φ SS-US*
HI	*θ SS-HI*	0.001	-	-	-	0.002	-	0.008	0.009	-	-	-	0.05	0.025	0.027	*θ SS-HI*
	R SS-HI	0.003	-	0.004	-	-	-	-	-	0.019	-	0.009	-	0.006	-	*Φ SS-HI*
AU	*θ SS-AU*	0.035	-	0.016	-	-	-	-	-	0.009	-	-	-	-	-	*θ SS-AU*
	R SS-AU	-	-	-	-	0.041	0.039	0.13	-	0.059	0.043	0.027	-	-	-	*Φ SS-AU*
MX/EC	*θ SS-EPAC*	-	-	-	-	-	-	0.005	-	-	-	-	-	-	-	*θ SS-EPAC*
	R SS-EPAC	0.022	-	0.022	-	-	-	-	-	-	-	-	-	0.01	-	*Φ SS-EPAC*
Overall	*θ SS-OVERALL*	0.002		-		-		-		-		-		0.005		*θ SS-OVERALL*
	R SS-OVERALL	-		-		0.002		-		-		-		-		*Φ SS-OVERALL*
ATL	*θ S-ATL*	-	-	0.003	0.054	-	-	0.004	-	-	-	-	-	0.003	-	*θ S-ATL*
	R S-ATL	-	-	0.005	-	-	-	-	-	-	-	-	-	-	-	*Φ S-ATL*
PAC	*θ S-PAC*	0.001	-	-	-	0.004	0.047	0.005	0.044	0.002	0.09	-	-	-	-	*θ S-PAC*
	R S-PAC	-	-	0.003	-	0.001	-	-	-	-	-	-	-	0.013	0.036	*Φ S-PAC*
Overall	*θ S-OVERALL*	-		-		-		-		-		-	-	0.001		*θ S-OVERALL*
	R S-OVERALL	-		-		-		-		-		-	-	-		*Φ S-OVERALL*
Oceans	*θ P*	0.0009	0.0003	0.019	< 0.0000	0.014	< 0.0000	0.017	< 0.0000	0.009	< 0.0000	0.013	< 0.0000	0.03	0.0003	*θ P*
	R P	0.02	0.0008	0.183	< 0.0000	0.014	0.034	0.27	< 0.0000	0.13	< 0.0000	0.153	< 0.0000	0.21	< 0.0000	*Φ P*
Avg. Diff. In Rpt # Means		0.9		6.3		0.9		4.2		3.5		3.8				
	Dmu	0.8		40.1		0.8		17.3		12.6		14.3				

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6C. Diversity and divergence summary over marker clases. H_e refers to expected heterozygosity. θ refers to Weir and Cockerham's (1984) unbiased F estimator, R refers to the *Rho* statistic of Rousset (1996), and Φ refers to the mitochondrial estimator of Excoffier et al. (1992) with restriction site information considered. p values were obtained by exact tests as discussed in text. *VAR* refers to variance and *STDEV* to standard deviation of the mean F values for each nuclear marker class.

ALLOZYME	*#Alleles*	*He*	θ	p
ACOH	2	0.11	0.107	<0.001
ADH	2	0.40	0.146	<0.001
G3PDH	3	0.43	0.050	0.011
IDDH	2	0.24	0.000	0.999
OVERALL	2	0.30	0.077	<0.001
VAR	0.25	0.02	0.004	
STDEV	0.50	0.15	0.064	
SCNDNA			θ	p
BM47	2	0.14	0.000	0.890
BM81	2	0.43	0.052	<0.001
BM32-2	2	0.47	0.102	<0.001
WM08	2	0.45	0.123	<0.001
OVERALL	2	0.37	0.086	<0.001
VAR		0.02	0.003	
STDEV		0.16	0.055	
MICROSATELLITE			R	p
GATA01	27	0.93	0.020	<0.001
GATA08	47	0.95	0.180	<0.001
GATA10	24	0.91	0.012	<0.001
GATA60	30	0.94	0.270	<0.001
GATA90	33	0.95	0.130	<0.001
OVERALL	32	0.94	0.145	<0.001
VAR	79.70	0.00	0.012	
STDEV	8.93	0.02	0.109	
MTDNA			Φ	p
	127	0.86	0.210	<0.002

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 7. Results of inter-ocean pairwise comparisons. For description of divergence estimators, see Table 6B. *STDEV* refers to standard deviation. *Avg. n* refers to the average haploid sample size. '*Power Index*' consists of the number of significant inter-ocean pairwise comparisons divided by the average sample size.

Microsatellite			ScnDNA		MtDNA	
LOCUS	*R*	θ	LOCUS	θ	Φ	θ
GATA01	0	4	BM47	0	36	4
GATA08	31	30	BM81	9	-	-
GATA10	0	14	BM32-2	24	-	-
GATA60	46	33	WM08	25	-	-
GATA90	30	18	-	-	-	-
Average	21.4	19.8	*Average*	14.5	36	4
STDEV.	20.5	11.9	*STDEV.*	12.1	-	-
Avg. n	62	62	*Avg. n*	61	24	24
Power Indx.	0.35	0.32	*Power Indx.*	0.24	1.5	0.17

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 8. Results of phenogram analyses. "NJ' and 'UPG' refer to the neighbor-joining and UPGMA clustering methods, respectively. '+' Indicates perfect clustering of populations into ocean of origin. '-' Indicates imperfect clustering of populations into ocean of origin. Distance estimators include linearized Rst (*Lin Rst*), linearized θ (*Lin* θ), Delta mu (*Dmu*), Nei's (1978) genetic distance (Nei, 1978), and linearized Φ_{ST} (*Lin* Φ_{ST}).

MICROSATELLITE DNA							ScnDNA					MtDNA			
LOCUS	*Lin Rst*		*Lin* θ		*Dmu*		LOCUS	*Nei, 78*		*Lin* θ		*Lin* Φ_{ST}		*Lin* θ	
	NJ	UPG	NJ	UPG	NJ	UPG		NJ	UPG	NJ	UPG	NJ	UPG	NJ	UPG
GATA01	-	-	-	-	-	-	BM47	-	-	-	-	+	+	+	-
GATA08	+	+	+	+	+	+	BM81	-	-	-	-				
GATA10	-	-	+	+	-	-	BM32-2	+	+	+	+				
GATA60	+	+	+	+	+	+	WM08	+	+	+	+				
GATA90	-	-	+	+	+	+									

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 9. Results of tests for indepedence among microsatellite alleles and mitochondrial clades. *R* refers to the *Rho* statistic of Rousset (1996). Significant heterogeneity among groups was evaluated by using exact *R* tests, and exact probability (exact prob.) tests.

GATA01	*R*	0.005
exact *R*	*p*	0.280
exact prob.	*p*	0.520
GATA08	*R*	0.005
exact *R*	*p*	0.702
exact prob.	*p*	0.202
GATA10	*R*	-0.005
exact *R*	*p*	0.738
exact prob.	*p*	0.828
GATA60	*R*	-0.002
exact *R*	*p*	0.501
exact prob.	*p*	0.230
GATA90	*R*	0.011
exact *R*	*p*	0.219
exact prob.	*p*	0.446

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 10. Mitochondrial haplotype frequency data by sex. Group' refers to groups generated from trimmed haplotype tree, of figure 2F. 'n' refers to haploid sample size, and 'FQ' to haplotype frequency. Ubiquitous clade includes groups 1 to 33 and Atlantic clade includes groups 34 to 67.

	MALE				FEMALE			
	ATLANTIC		PACIFIC		ATLANTIC		PACIFIC	
Group	n	FQ	n	FQ	n	FQ	n	FQ
1	2	0.02						
2	7	0.07	1	0.05	1	0.05	8	0.14
3			1	0.05				
4	1	0.01					1	0.02
5	1	0.01					1	0.02
6	6	0.06	3	0.16	1	0.05	6	0.11
7								
8	1	0.01					2	0.04
9	2	0.02	1	0.05			4	0.07
10							1	0.02
11	25	0.26	11	0.58	3	0.16	25	0.44
12								
13	1	0.01						
14					1	0.05		
15	1	0.01	1	0.05	1	0.05	1	0.02
16	2	0.02					3	0.05
17	2	0.02						
18	1	0.01					1	0.02
19	3	0.03					1	0.02
20	1	0.01						
21					1	0.05		
22								
23	1	0.01						
24	1	0.01						
25					3	0.16		
26								
27							1	0.02
28								
29								
30								
31							1	0.02
32							1	0.02
33	1	0.01						

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 10 Continued.

	MALE				FEMALE			
	ATLANTIC		PACIFIC		ATLANTIC		PACIFIC	
Group	n	FQ	n	FQ	n	FQ	n	FQ
34	1	0.01						
35	1	0.01						
36								
37								
38	1	0.01						
39	1	0.01						
40	2	0.02						
41			1	0.05				
42								
43								
44								
45								
46	2	0.02						
47	2	0.02						
48								
49	4	0.04						
50	1	0.01			1	0.05		
51	1	0.01			1	0.05		
52	1	0.01						
53	1	0.01			2	0.11		
54	1	0.01						
55	7	0.07			2	0.11		
56								
57	3	0.03						
58	1	0.01						
59	4	0.04			1	0.05		
60	2	0.02						
61					1	0.05		
62	1	0.01						
63								
64								
65								
66	1	0.01						
67	1	0.01						
Total	98		19.0		19		57	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

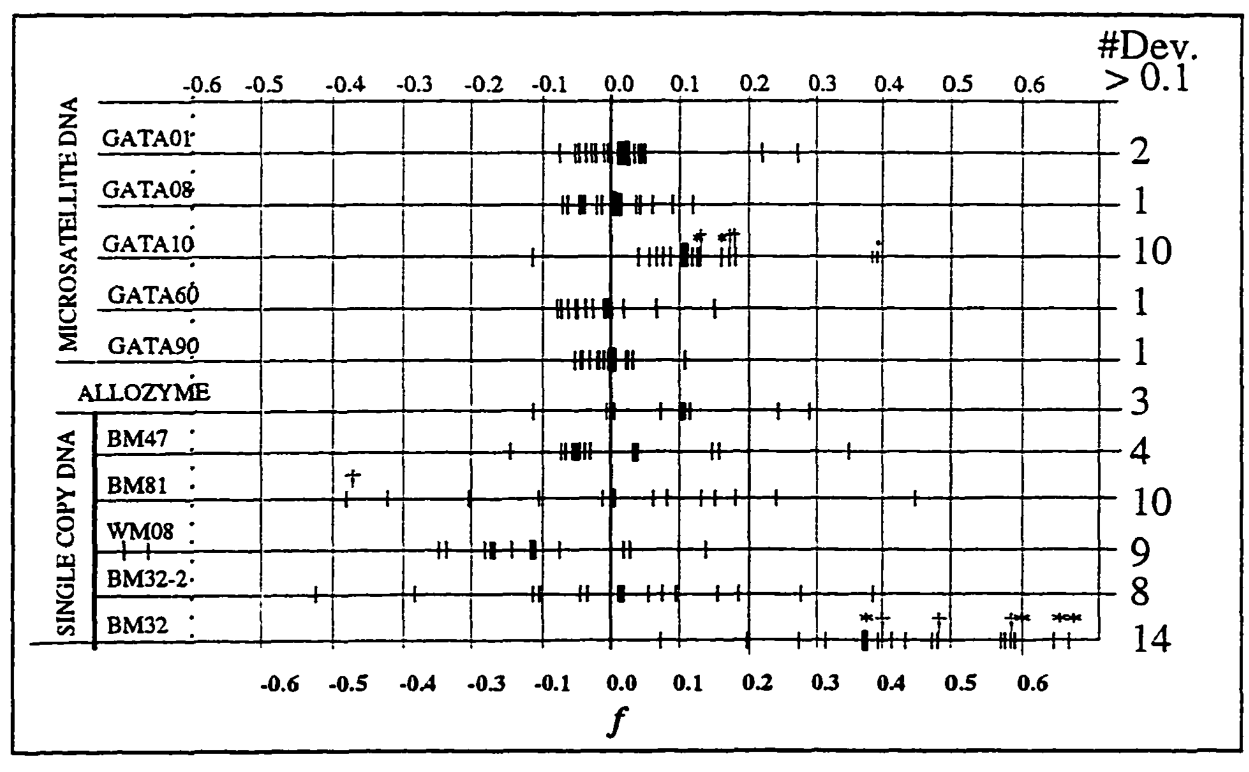
Table 11. Estimates of female (Fem.) and male inter-ocean divergence, and tests for homogeneity in allele frequency among sexes within an ocean. 'A' and 'P' refer to Atlantic and Pacific samples, respectively. The sample sizes for each comparison are shown (n). *p* values represent exact probability values for the divergence estimator θ (Weir and Cockerham, 1984) and exact Φ values for the divergence estimator Φ, which incorporates restriction site information among alleles (Excoffier et al., 1992). '-' indicates divergence estimates less than 0.001, or significance probabilities greater than 0.05.

		A. vs P. Fem.	A. vs P. Male	Male vs Fem. A.	Male vs Fem. P.
	n's	19 vs 55	100 vs 21	100 vs 19	21 vs 55
mtDNA	θ	0.036	0.044	-	0.003
	p	*< 0.00001*	-	-	-
	Φ	0.326	0.167	-	-
	p	*< 0.00001*	*0.0006*	-	-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 1. Distribution of *f* values, the local inbreeding coefficient (Weir and Cockerham, 1984). Vertical bars indicate position of each sample for each locus for five microsatellite and four scnDNA loci. Atlantic and Pacific samples for four allozyme loci were combined on a single axis. Values were non-significant unless indicated otherwise. "†" Indicates values were non-significant after correction for multiple tests (initial α = 0.0033). "*" Indicates significant after corrections for multiple tests. Significance probabilities represent exact probabilities of conformance to Hardy-Weinberg expectations. Thicker line indicates mean *f* for each set of values. '# Dev. > 0.1' indicates sum of all deviations greater than 0.1 from the equilibrium value of 0.

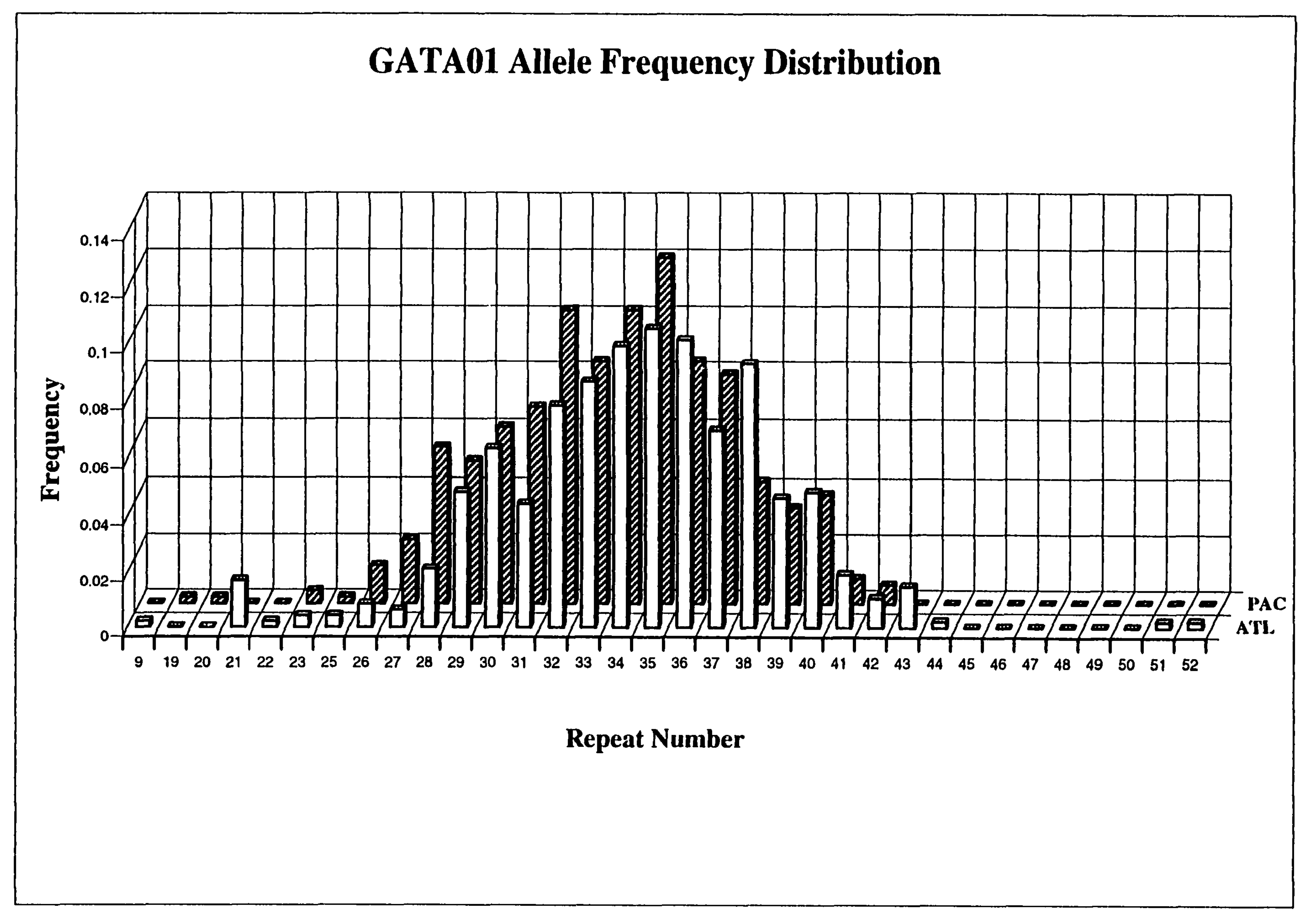
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

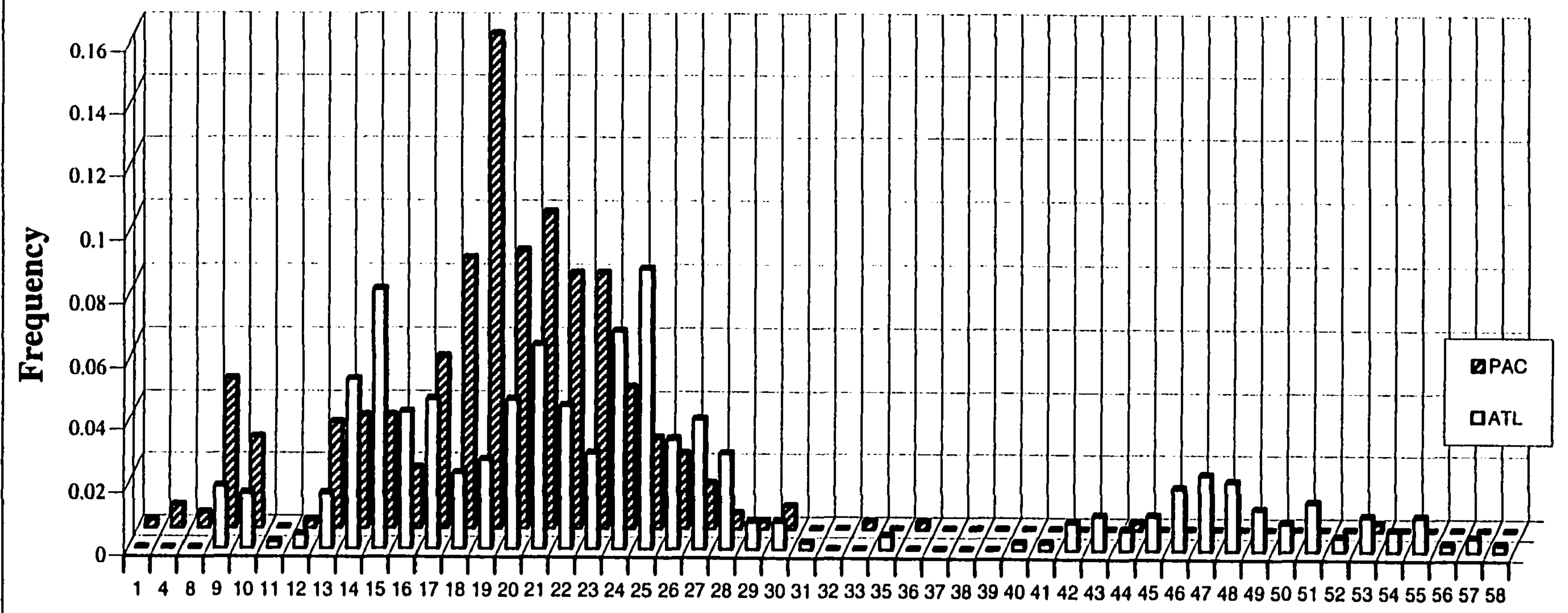
Figure 2A-2F. Allele frequency distributions for each locus in samples from the Atlantic and Pacific. (A) GATA01; (B) GATA08; (C) GATA10; (D) GATA60; and (E) GATA90. (F) Similarity of mtDNA composite haplotypes based on UPGMA clustering. The exterior two branching levels were collapsed, and Atlantic and ubiquitous clades indicated. Size of circles represents the frequency of each haplotype group within Atlantic and Pacific samples.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



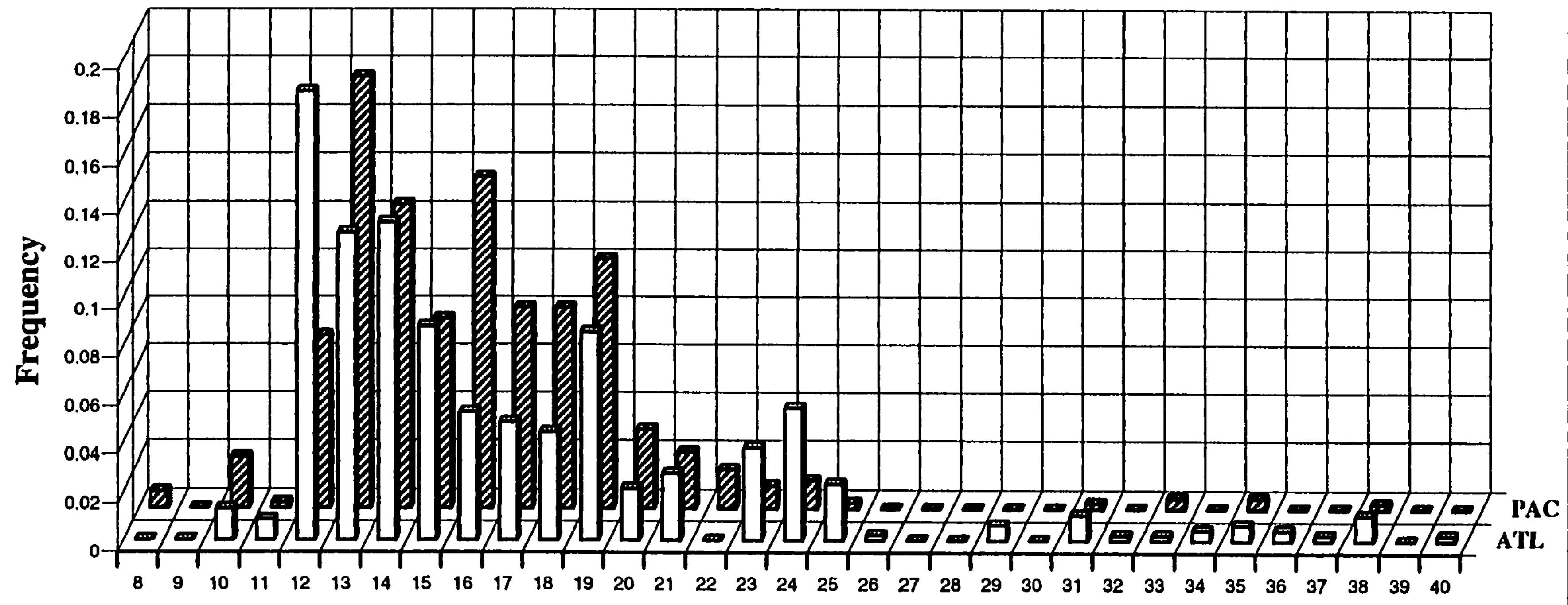
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

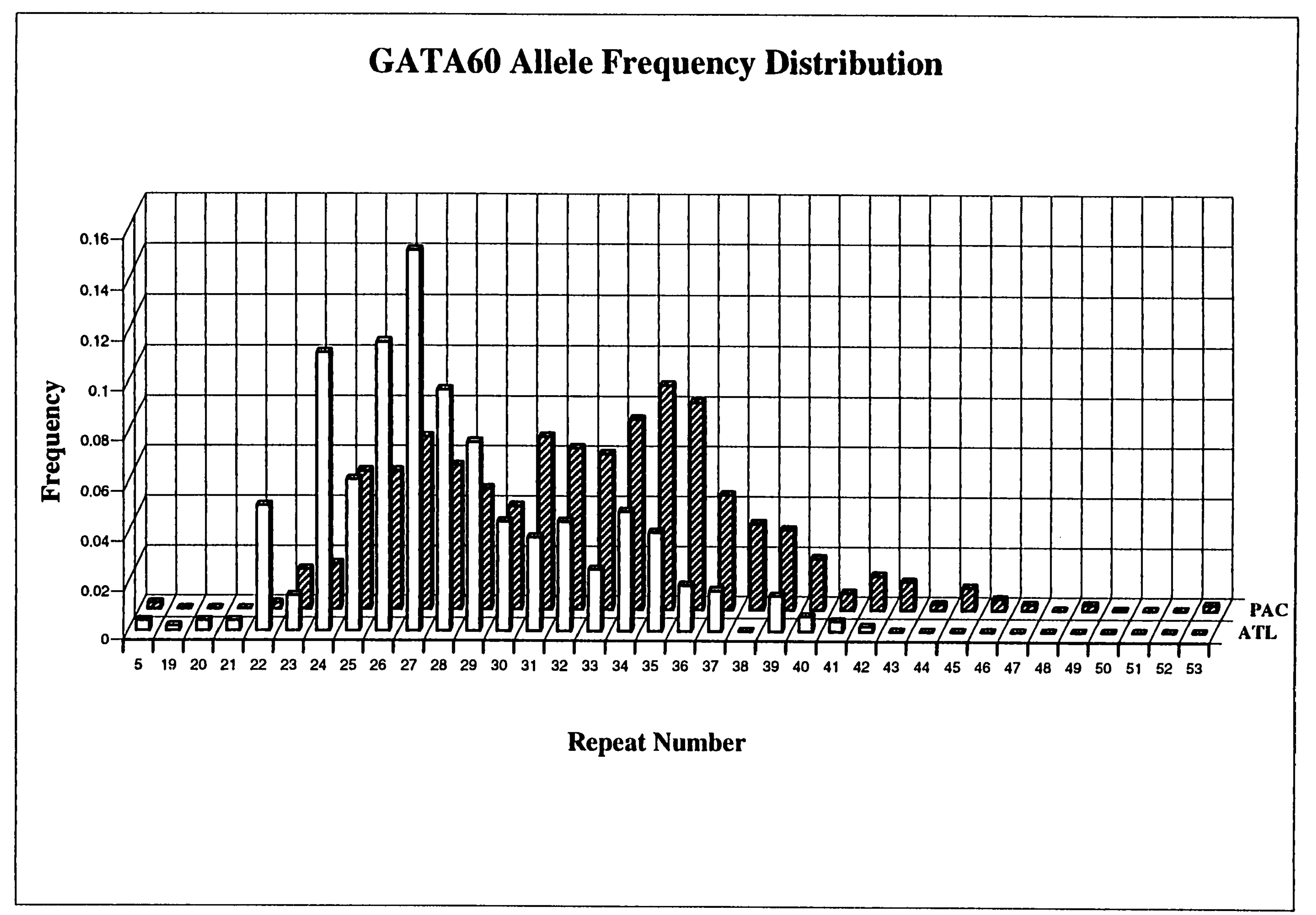


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

GATA10 Allele Frequency Distribution

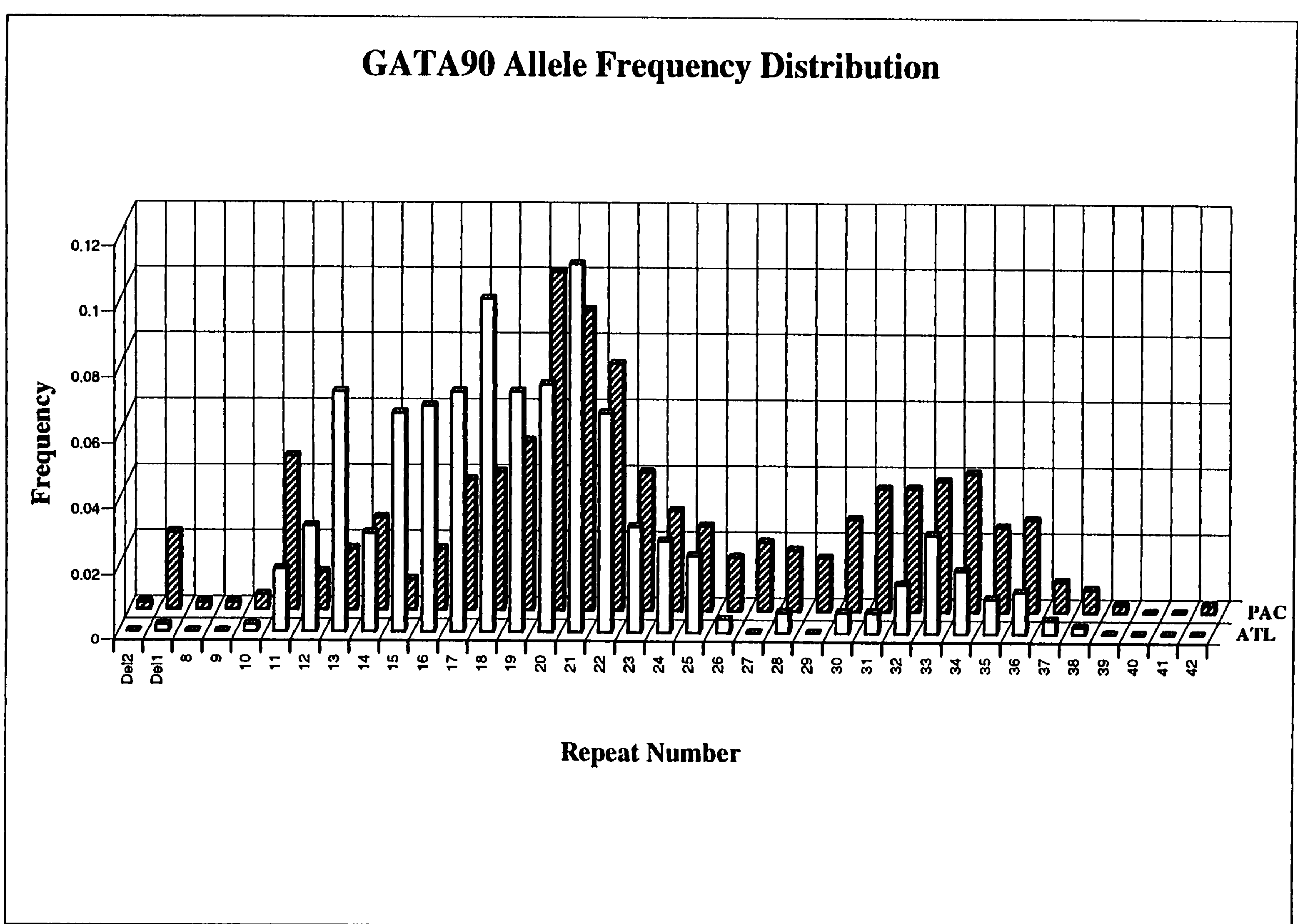


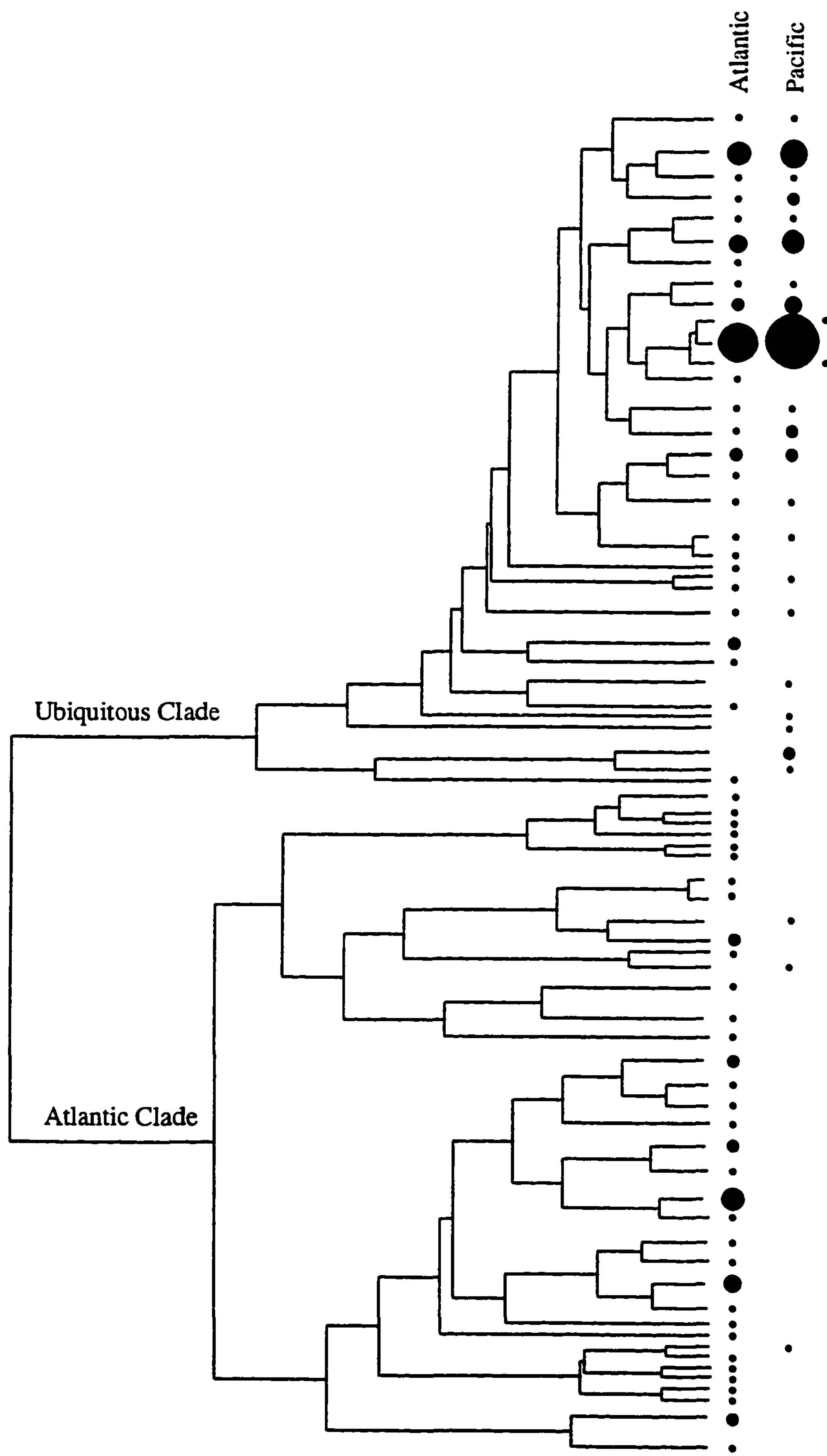
Repeat Number



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

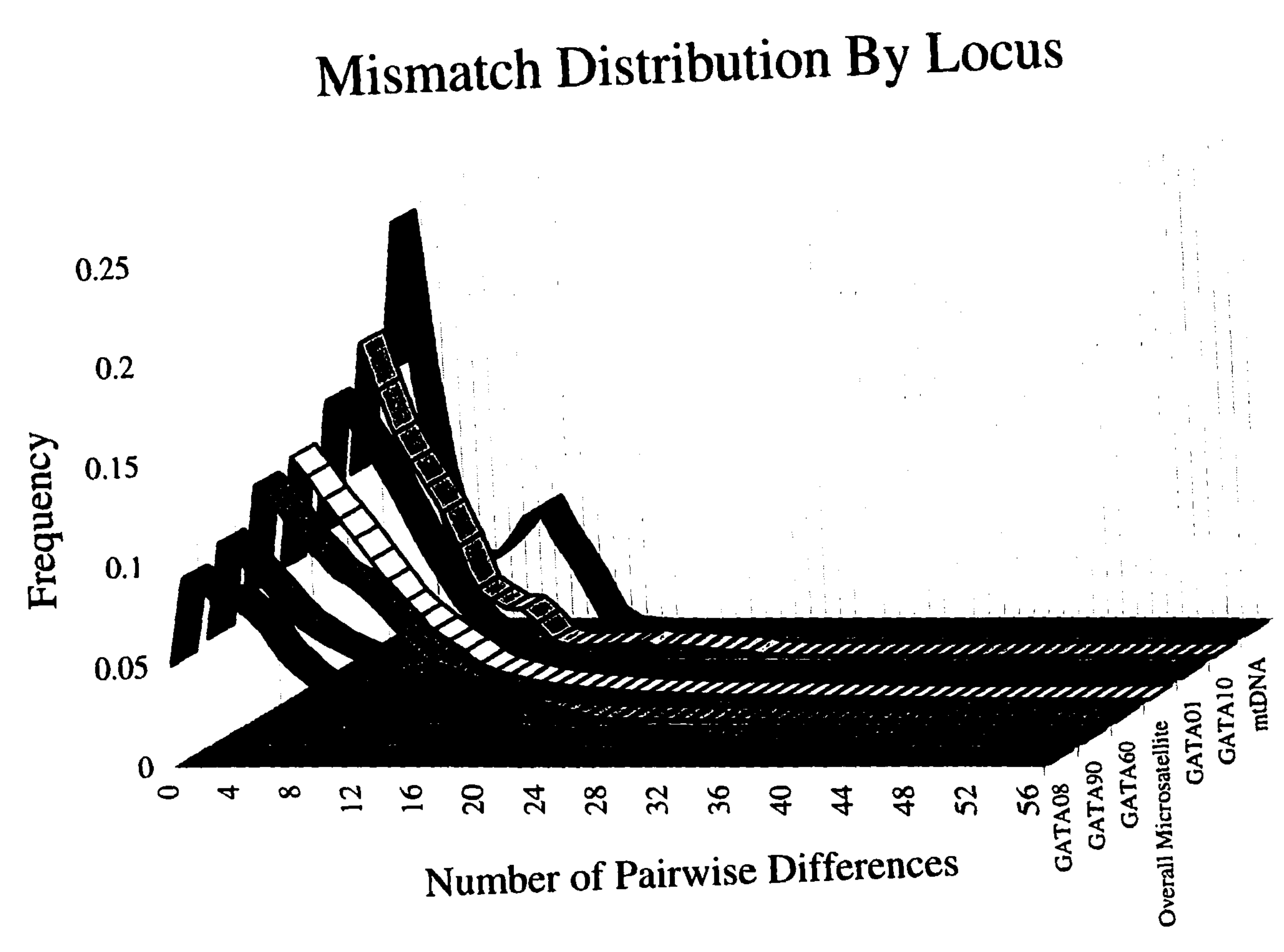




Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

FIGURE 3. Frequency distribution of pairwise differences for each microsatellite locus and for mtDNA.

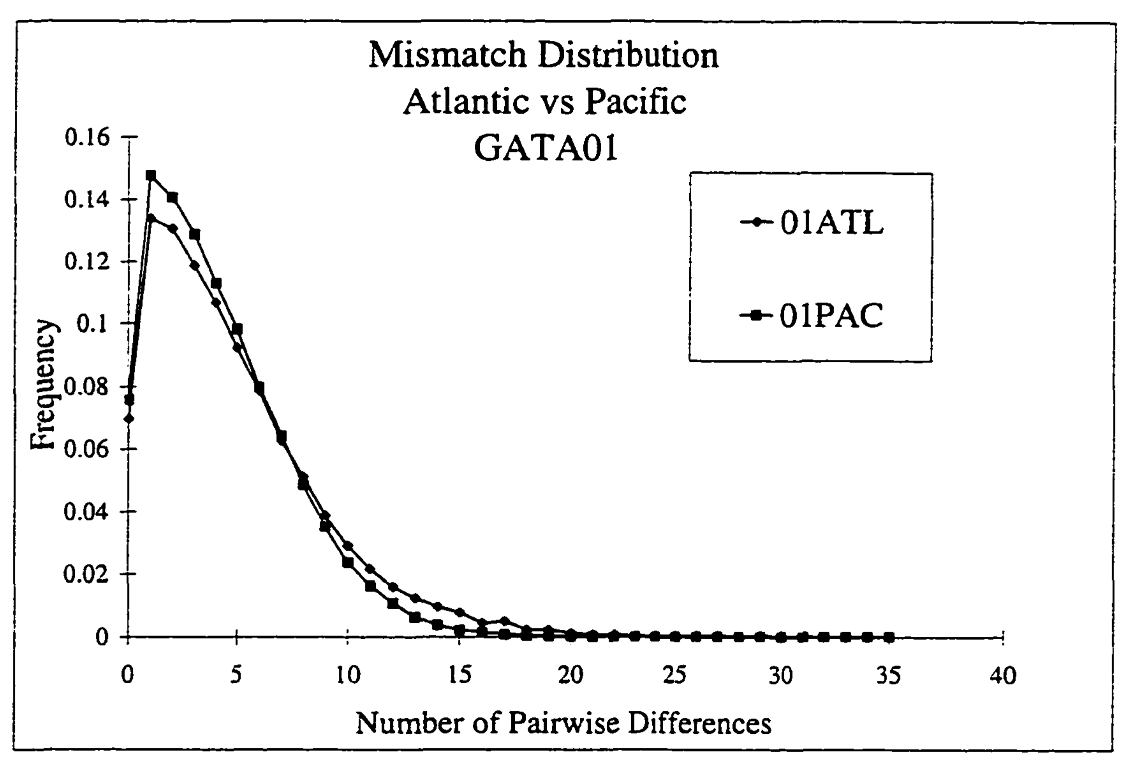
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

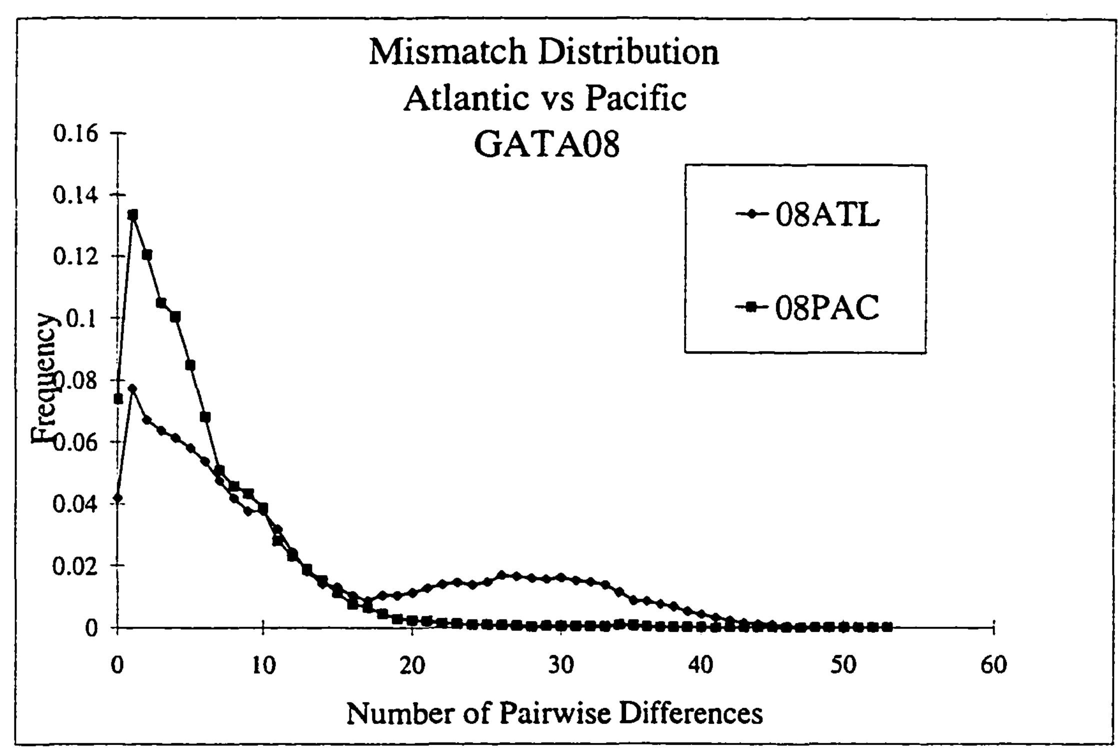


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

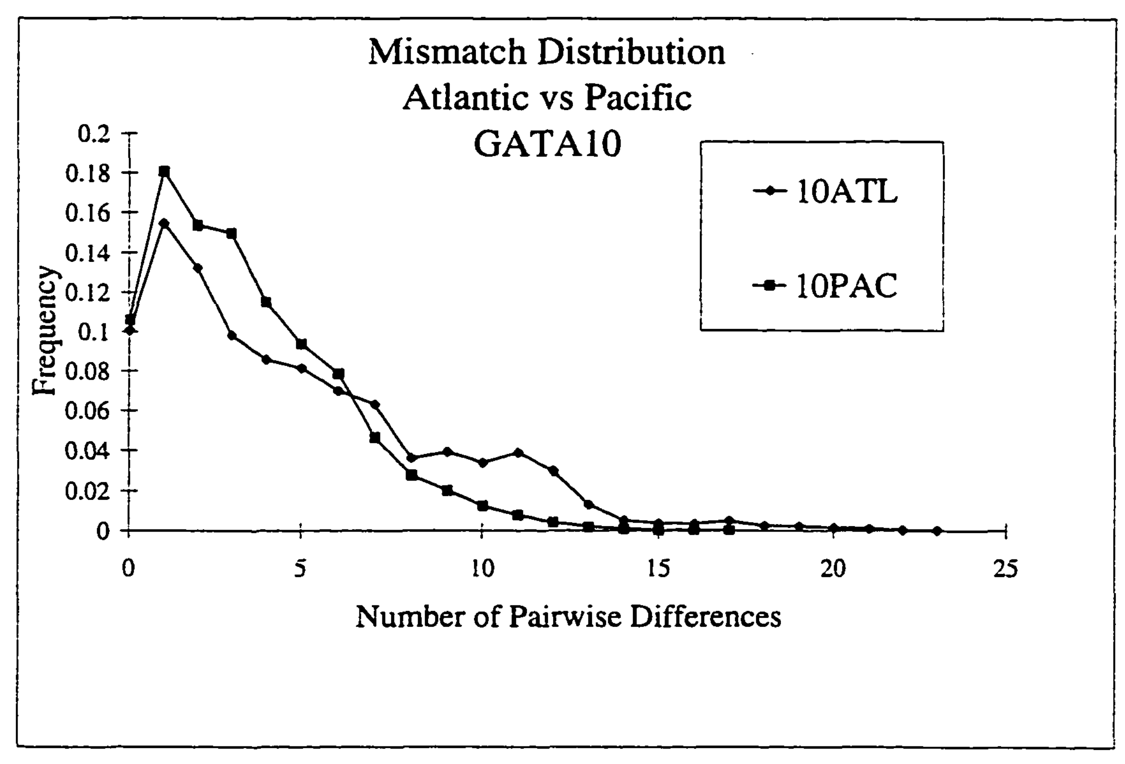
FIGURE 4A-4E. Mismatch distribution for each microsatellite locus and mtDNA over Atlantic and Pacific samples. (A) GATA01; (B) GATA08; (C) GATA10; (D) GATA60; (E) GATA90; (F) mtDNA.

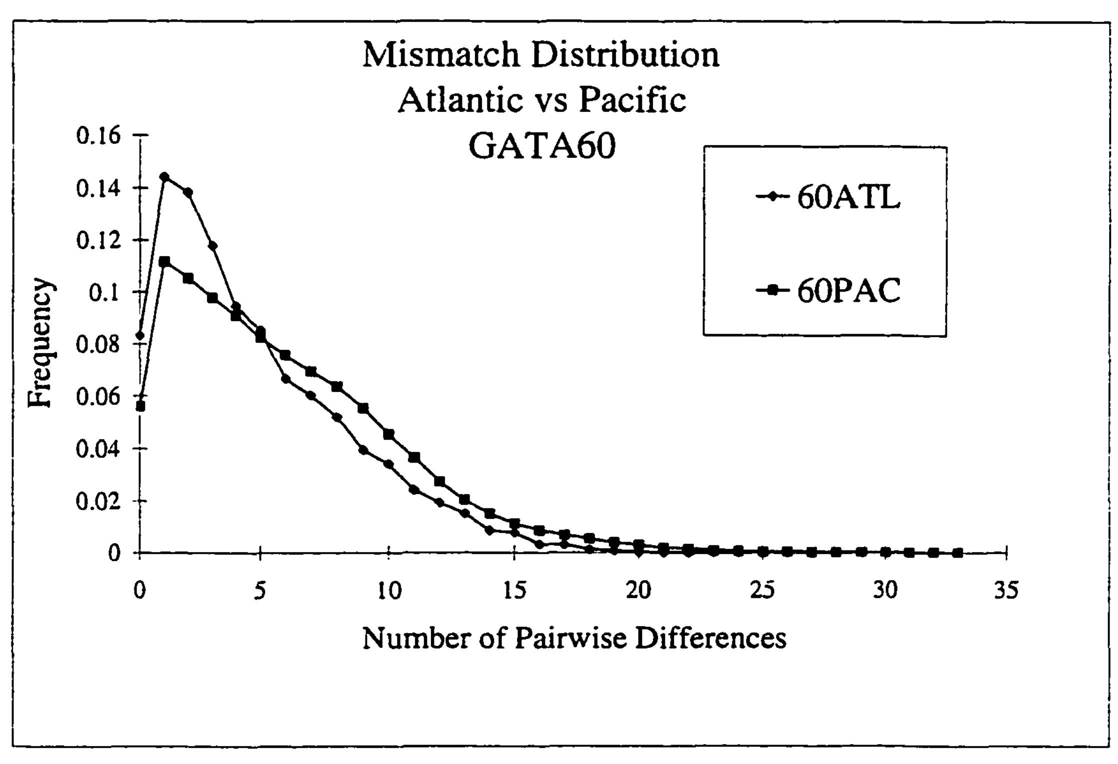
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



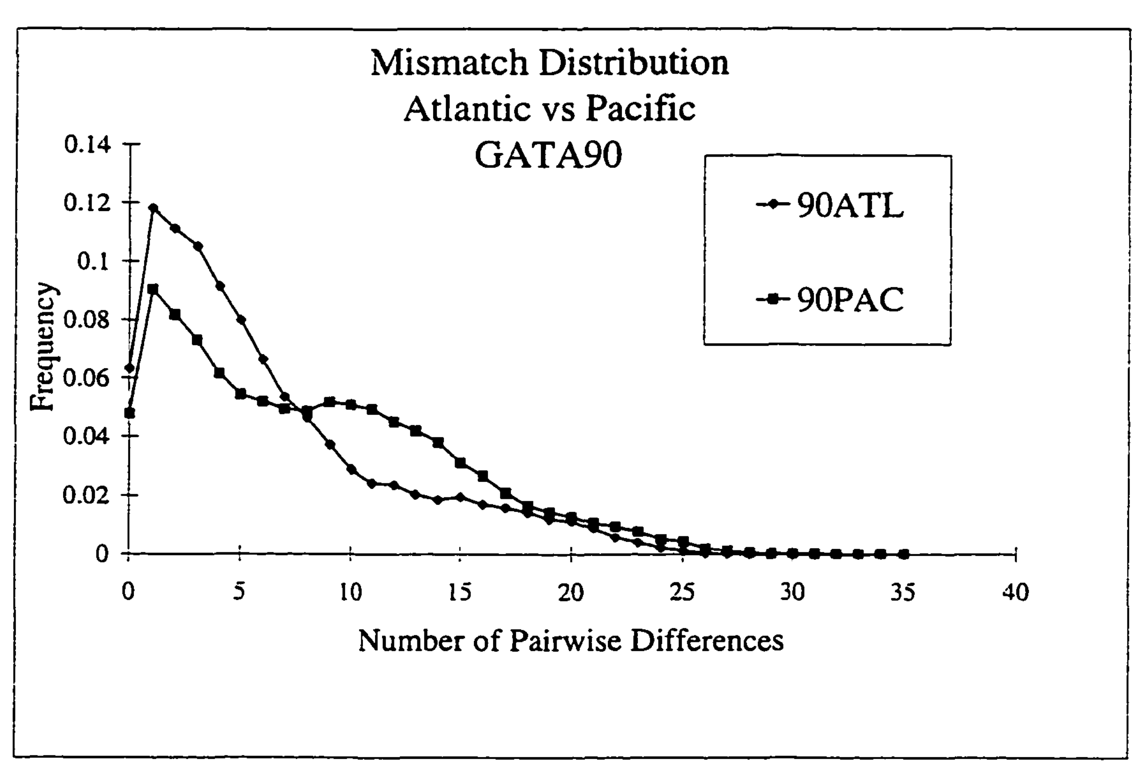


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.





Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

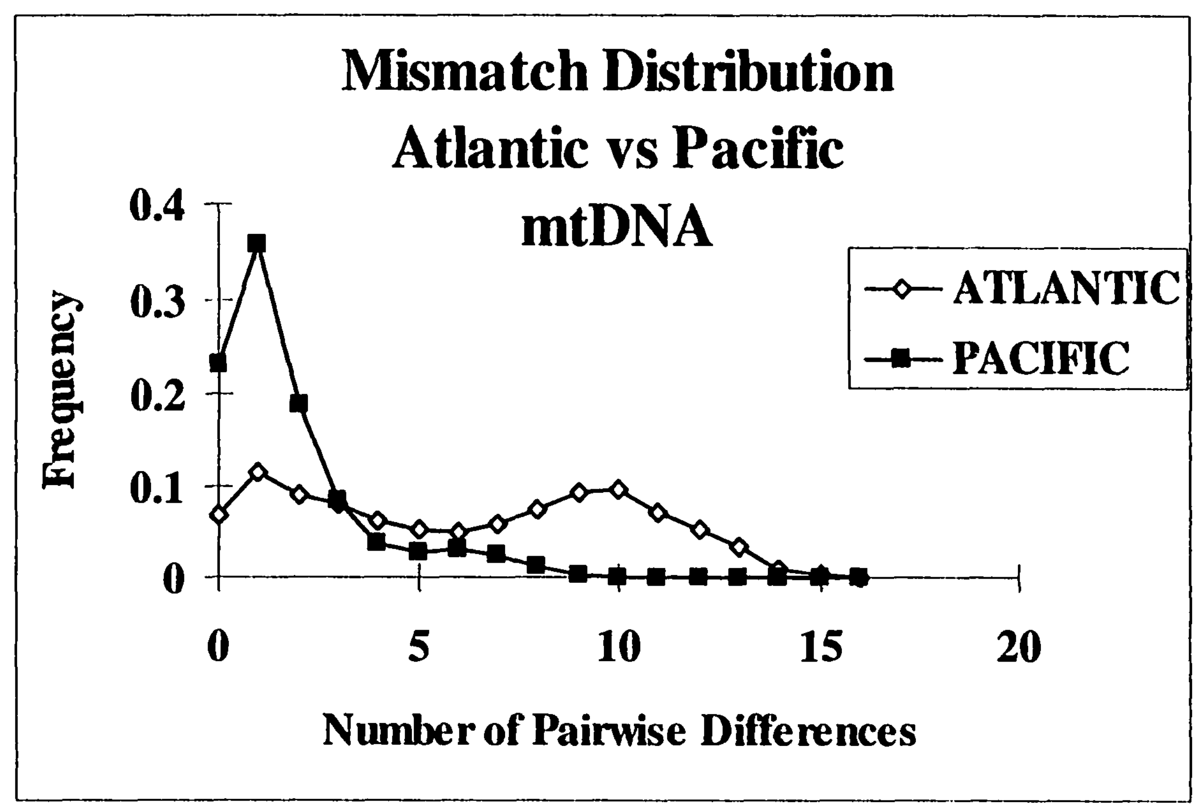
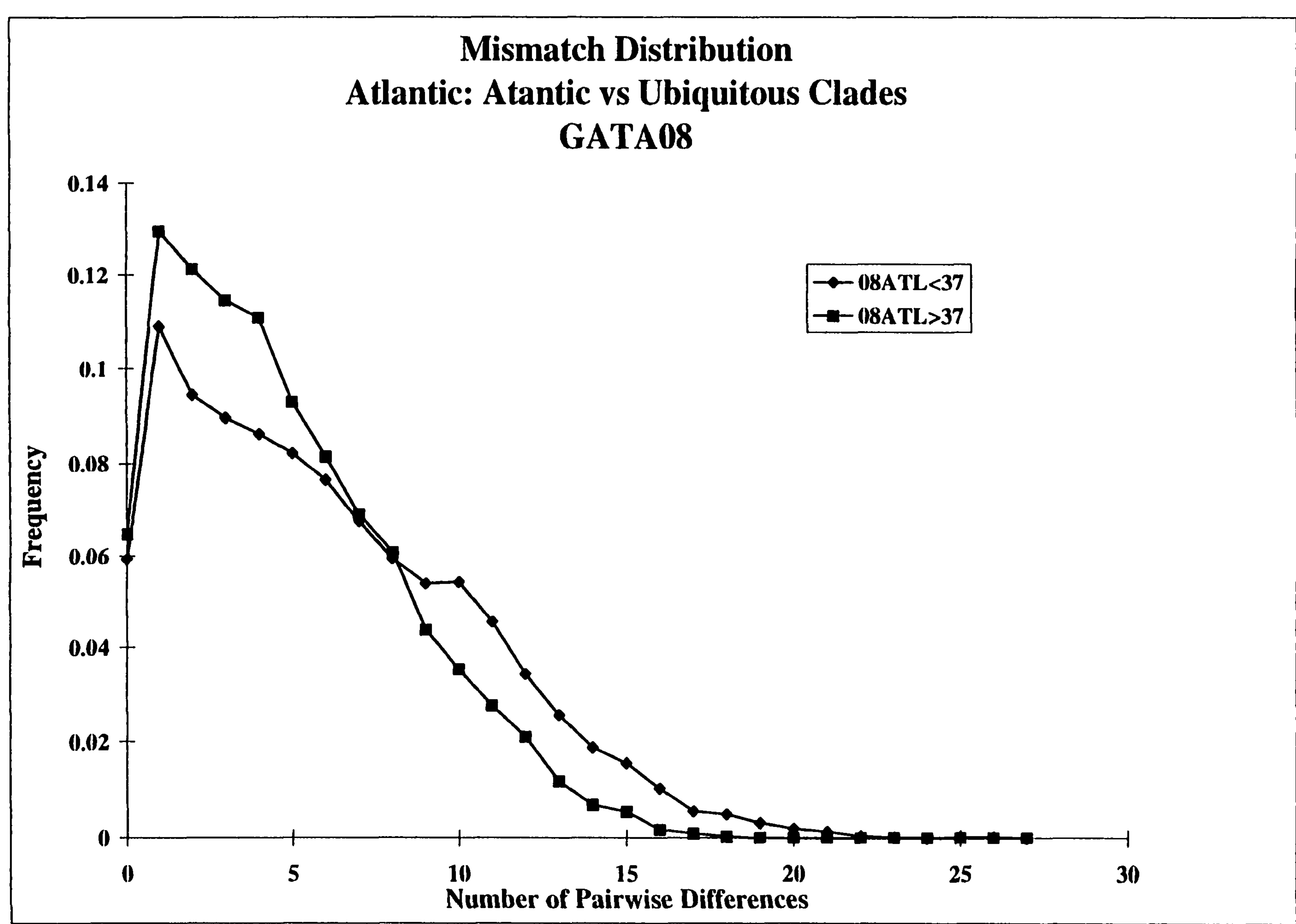


FIGURE 5. Mismatch distribution for 'ubiquitous' versus 'Atlantic' clades of alleles. (A) GATA08; (B) mtDNA. The GATA08 'Atlantic clade is defined as > 37 repeats.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

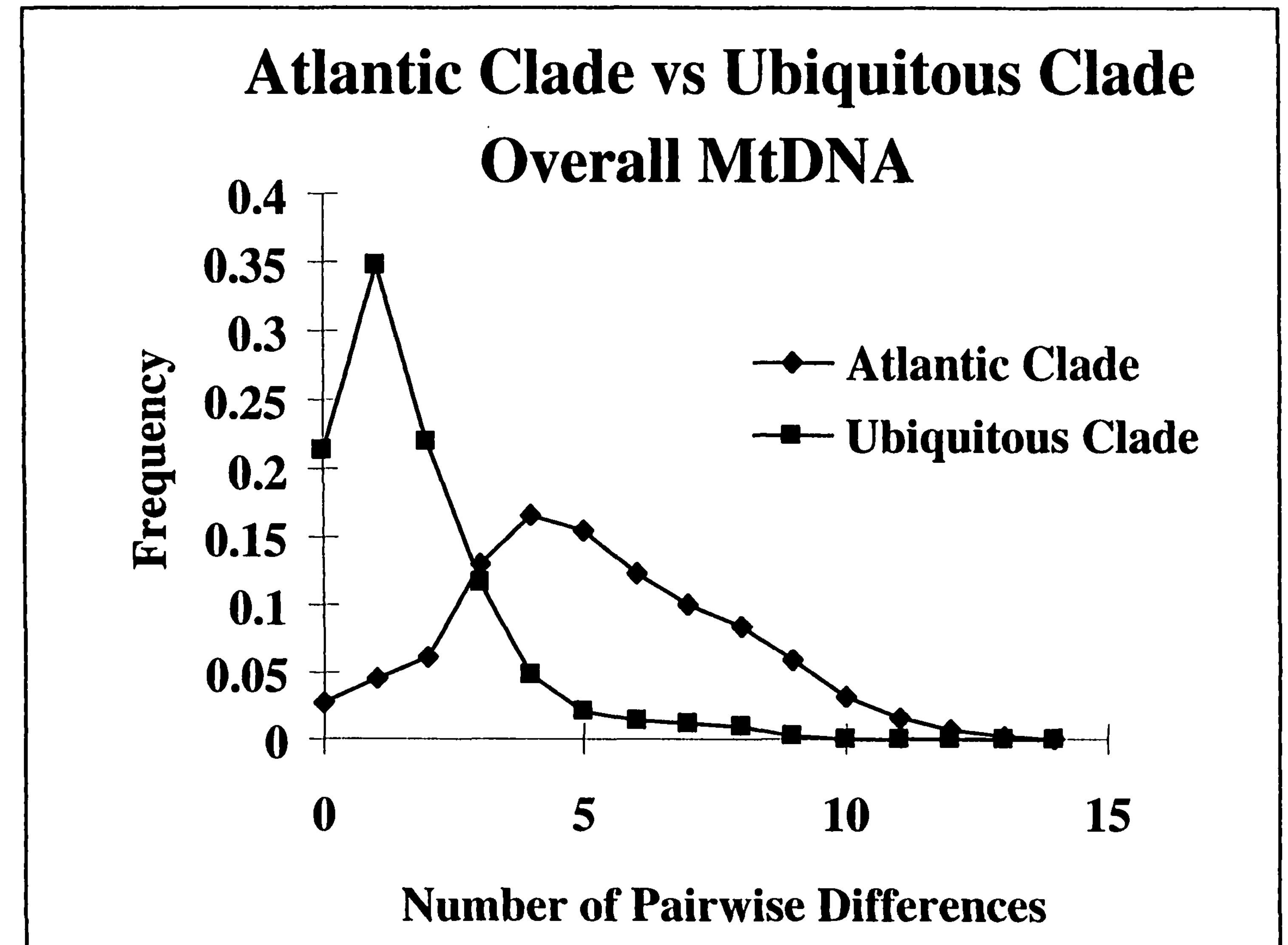


FIGURE 6. Mismatch distribution over all microsatellite loci, for Atlantic and Pacific samples.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

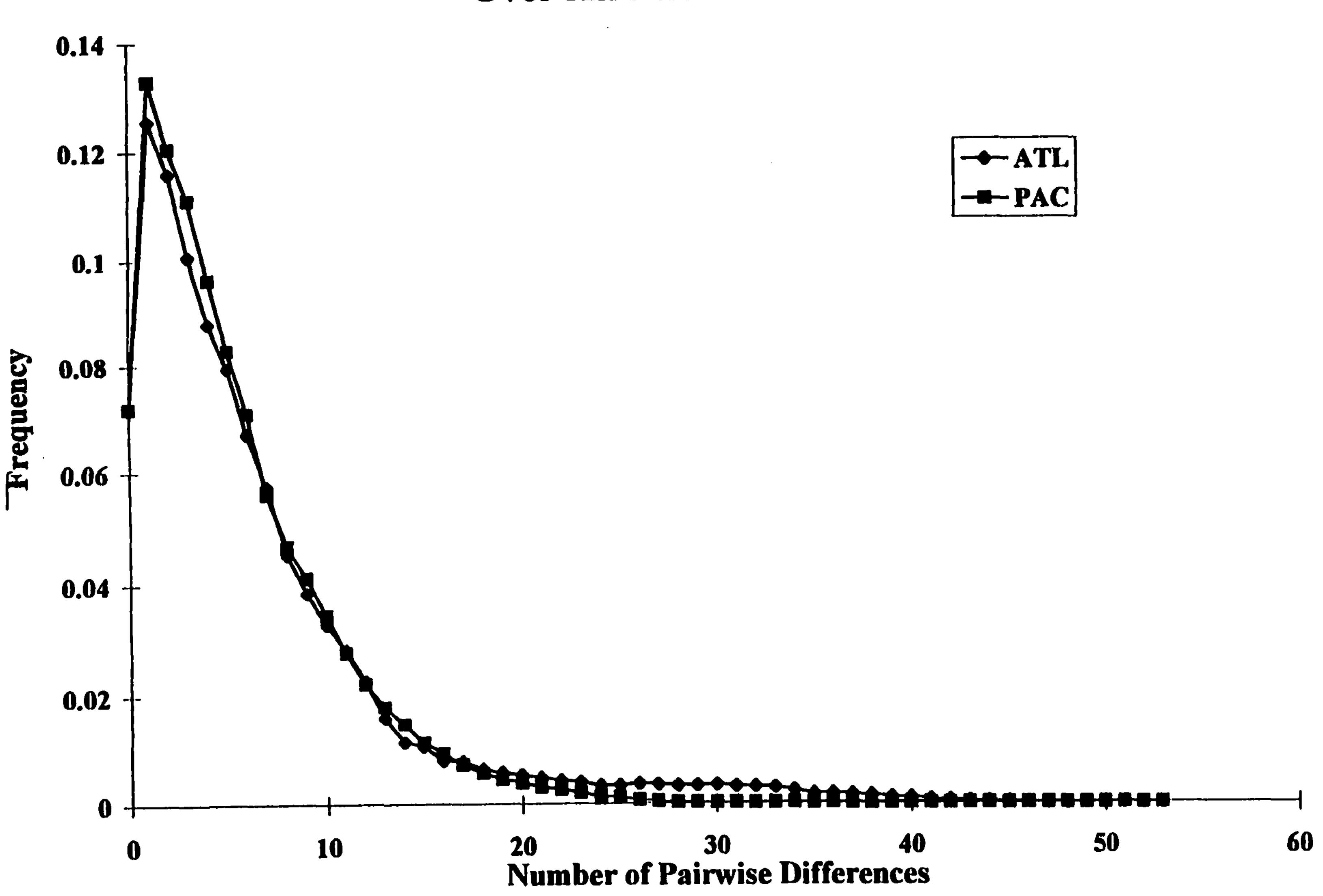
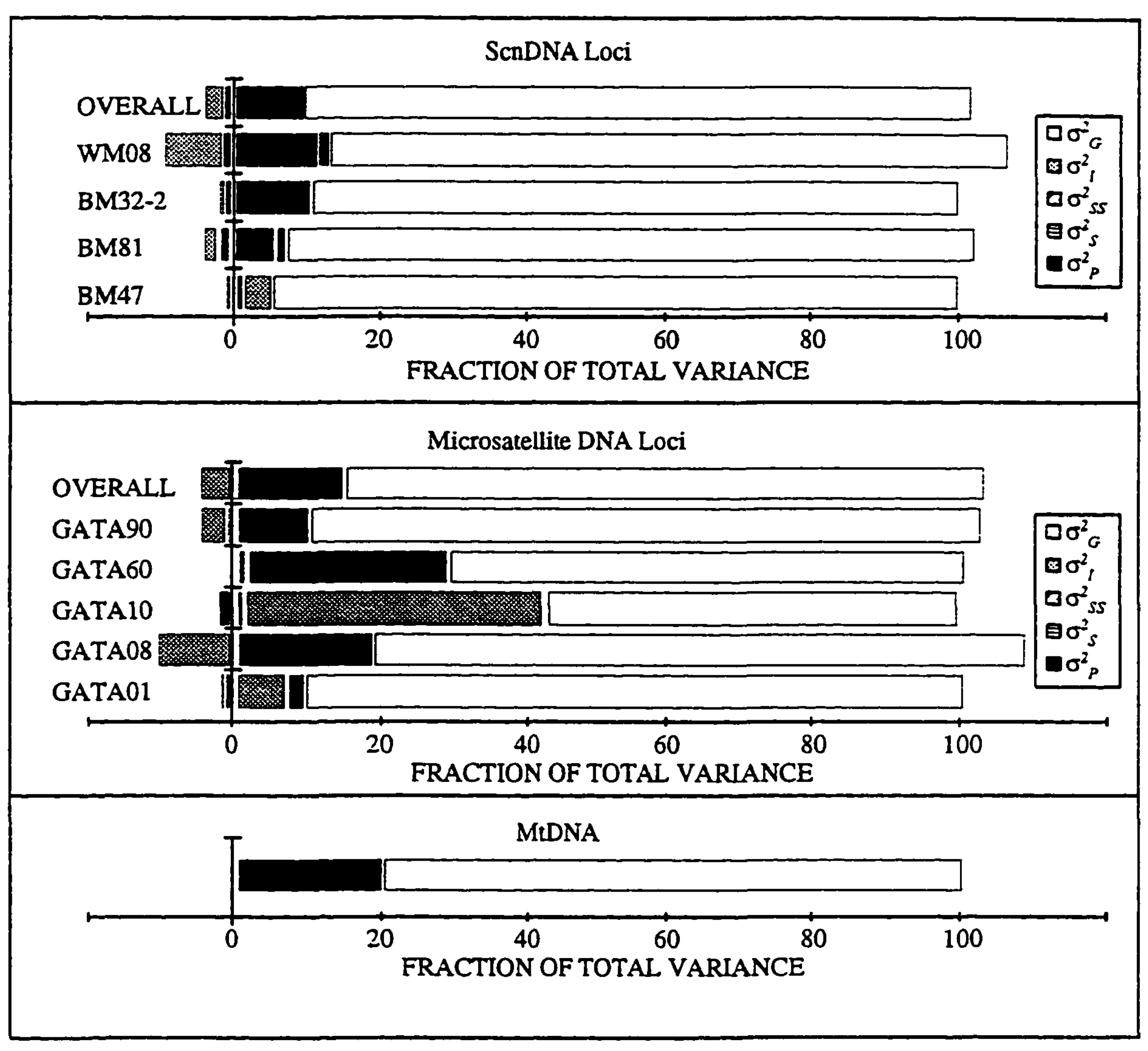


FIGURE 7. Hierarchical analysis of variance. Source of variance component are: σ^2_G, genes within individuals; σ^2_I, individuals within subsubpopulations (years); σ^2_{SS}, subsubpopulations within subpopulations (geographic locations); σ^2_S, subpopulations within populations (oceans); and σ^2_P, between populations. Shaded areas represent proportion of total variance attributable to each variance component.

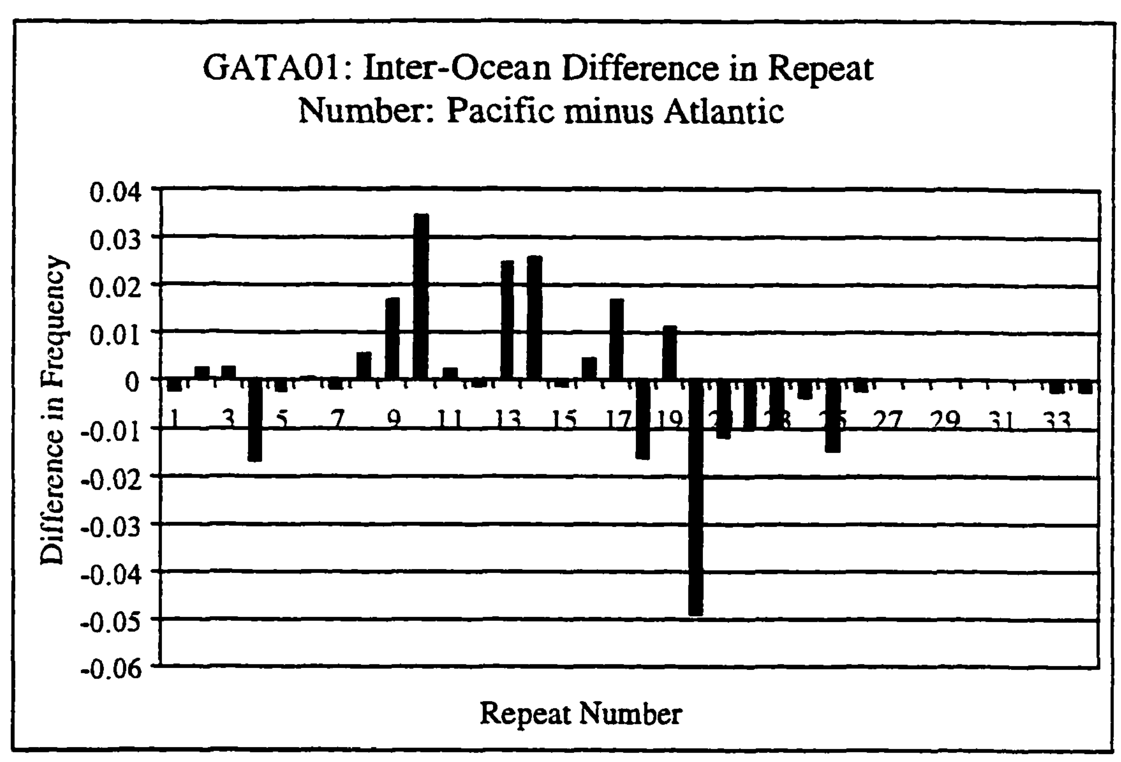
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

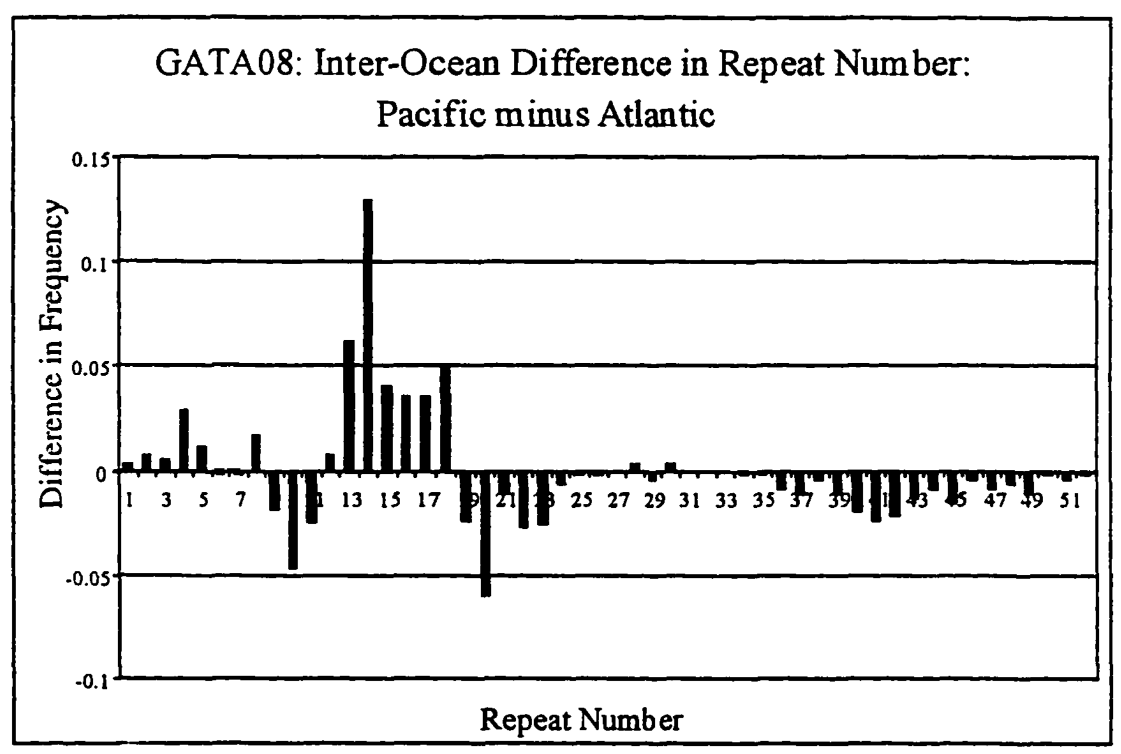


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

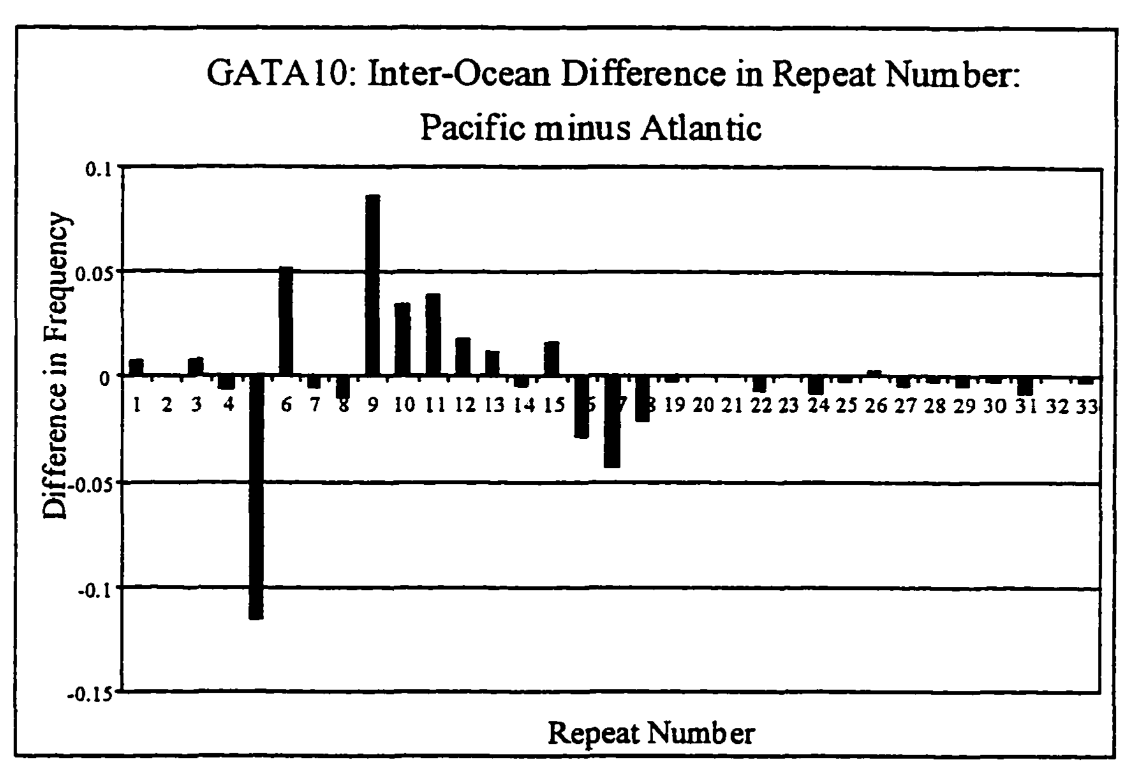
FIGURE 8. Pairwise difference in allele frequency between Atlantic and Pacific samples. Values were calculated as Pacific frequencies minus Atlantic frequencies. (A) GATA01; (B) GATA08; (C) GATA10; (D) GATA60; and (E) GATA90. (F) Pairwise comparisons of allele frequencies at GATA60 between inter-annual samples from Hawaii.

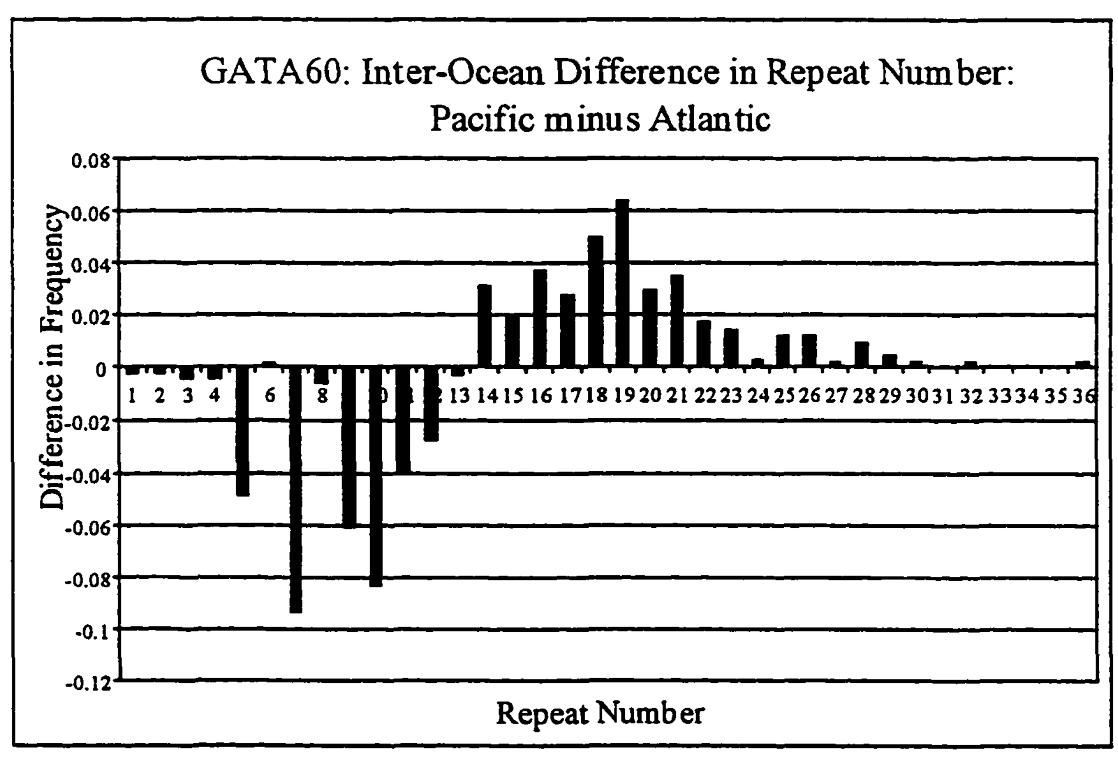
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



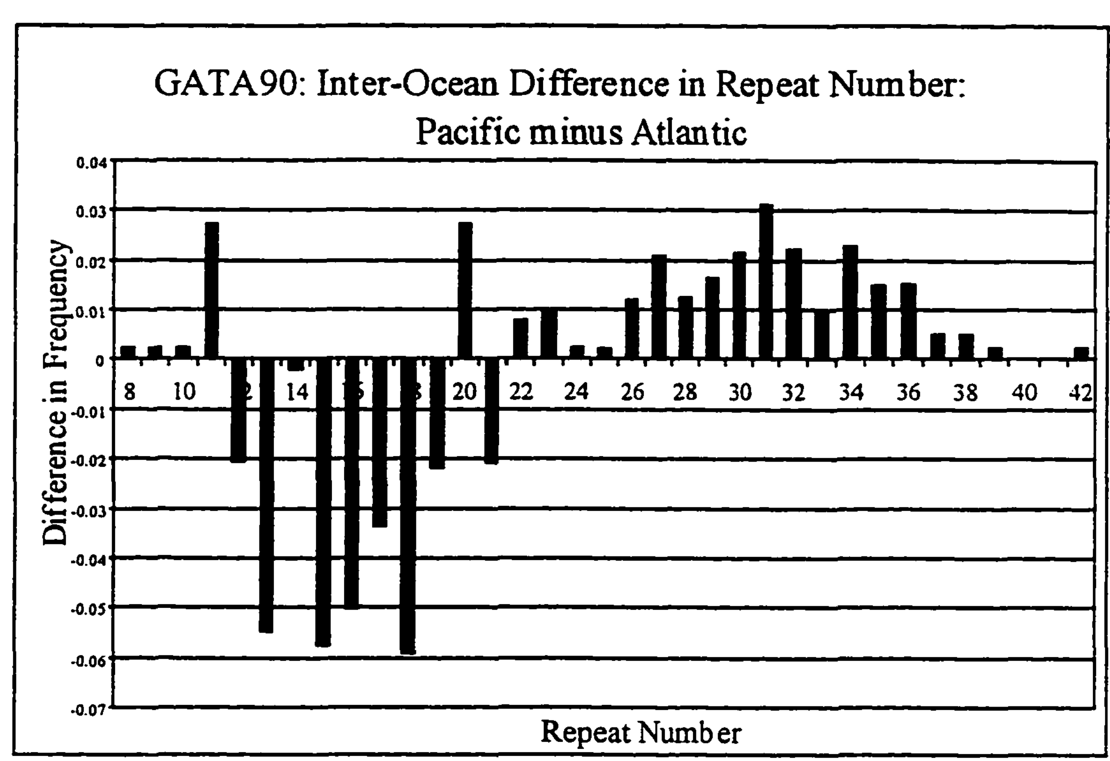


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

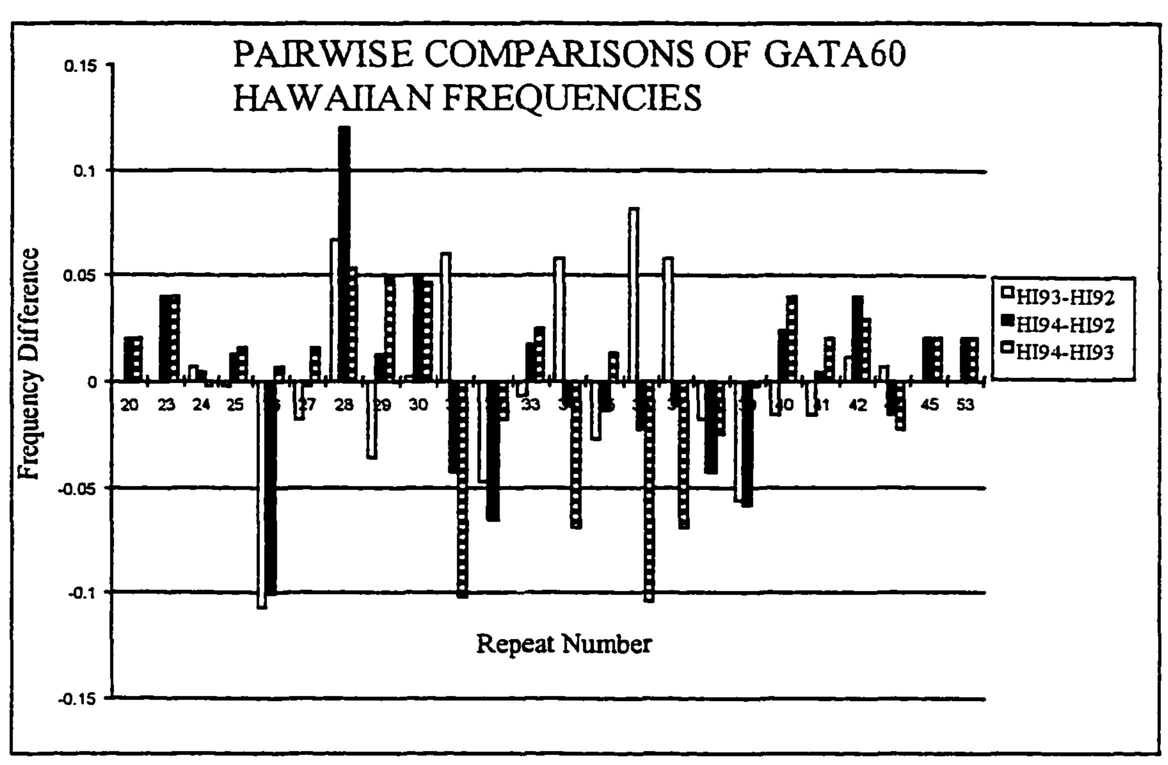




Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

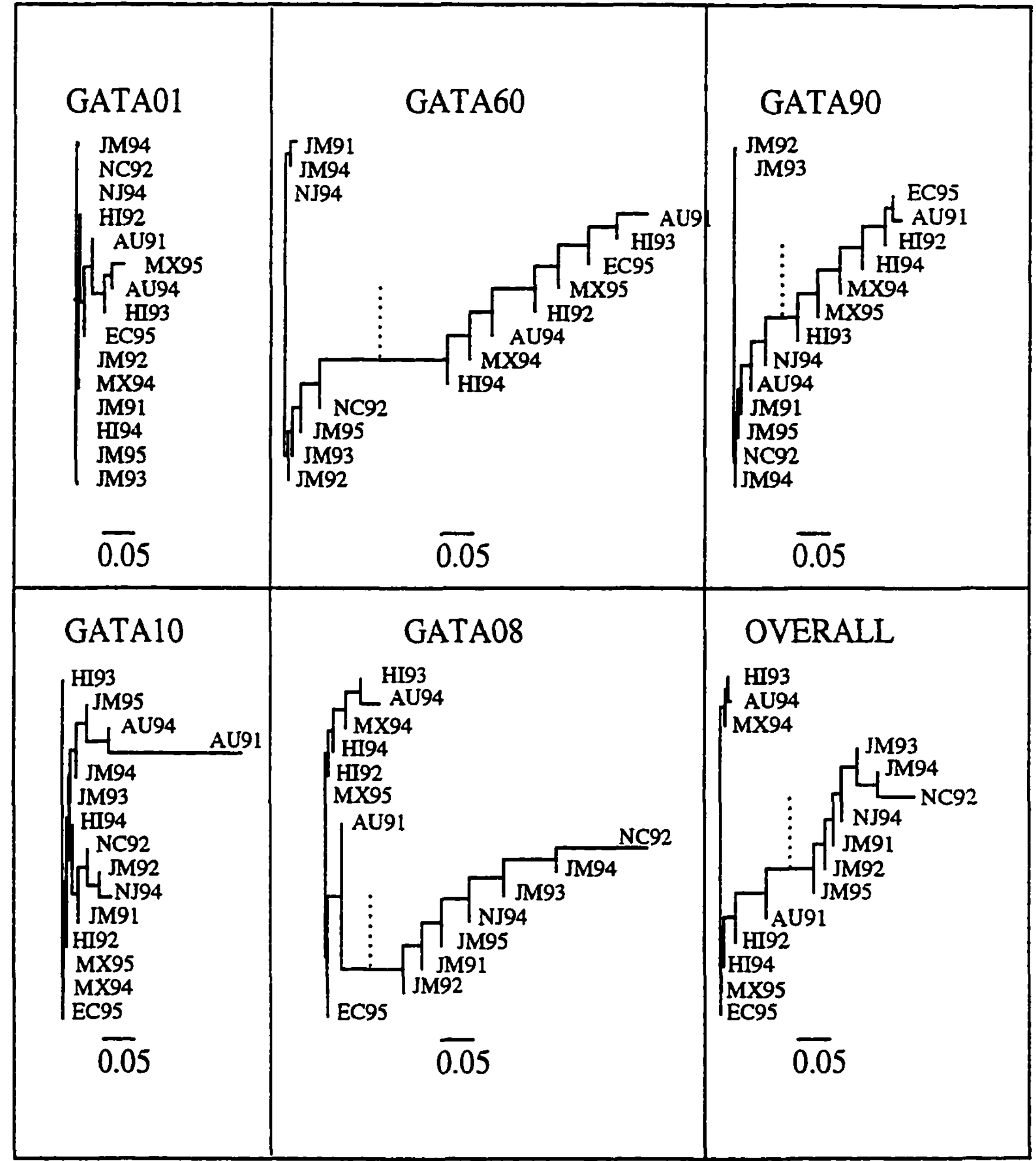


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

FIGURE 9. Neighbor-joining topologies constructed for each microsatellite locus by using (A) Slatkin's linearized R, (B) Slatkin's linearized θ, and (C) *Dmu* distance measures. (9D) ScnDNA linearized θ, microsatellite linearized R, and mtDNA linearized Φ distances measures over all loci. Dashed line indicates assumed midpoint root.

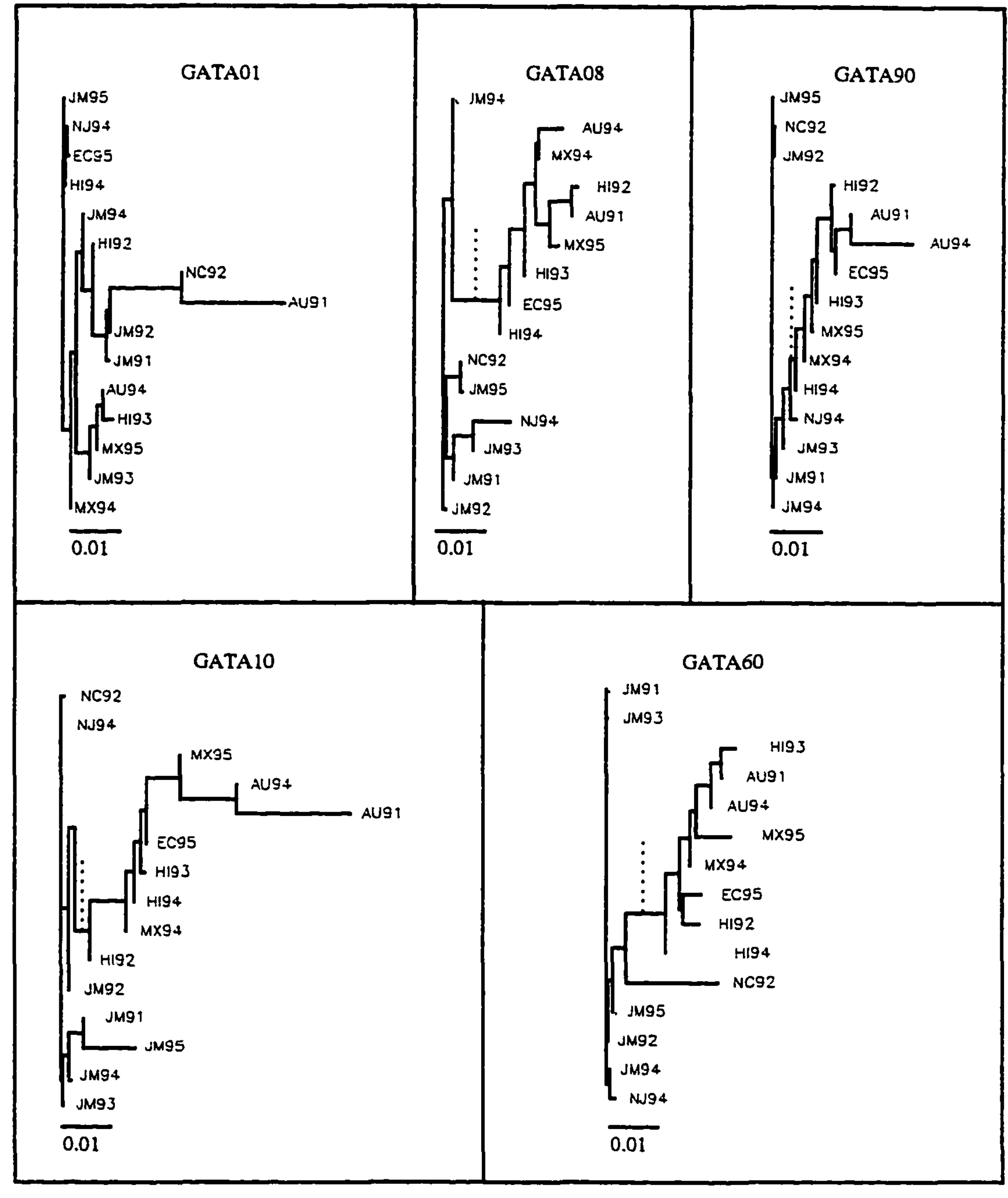
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Linearized *R* distance measure for microsatellites



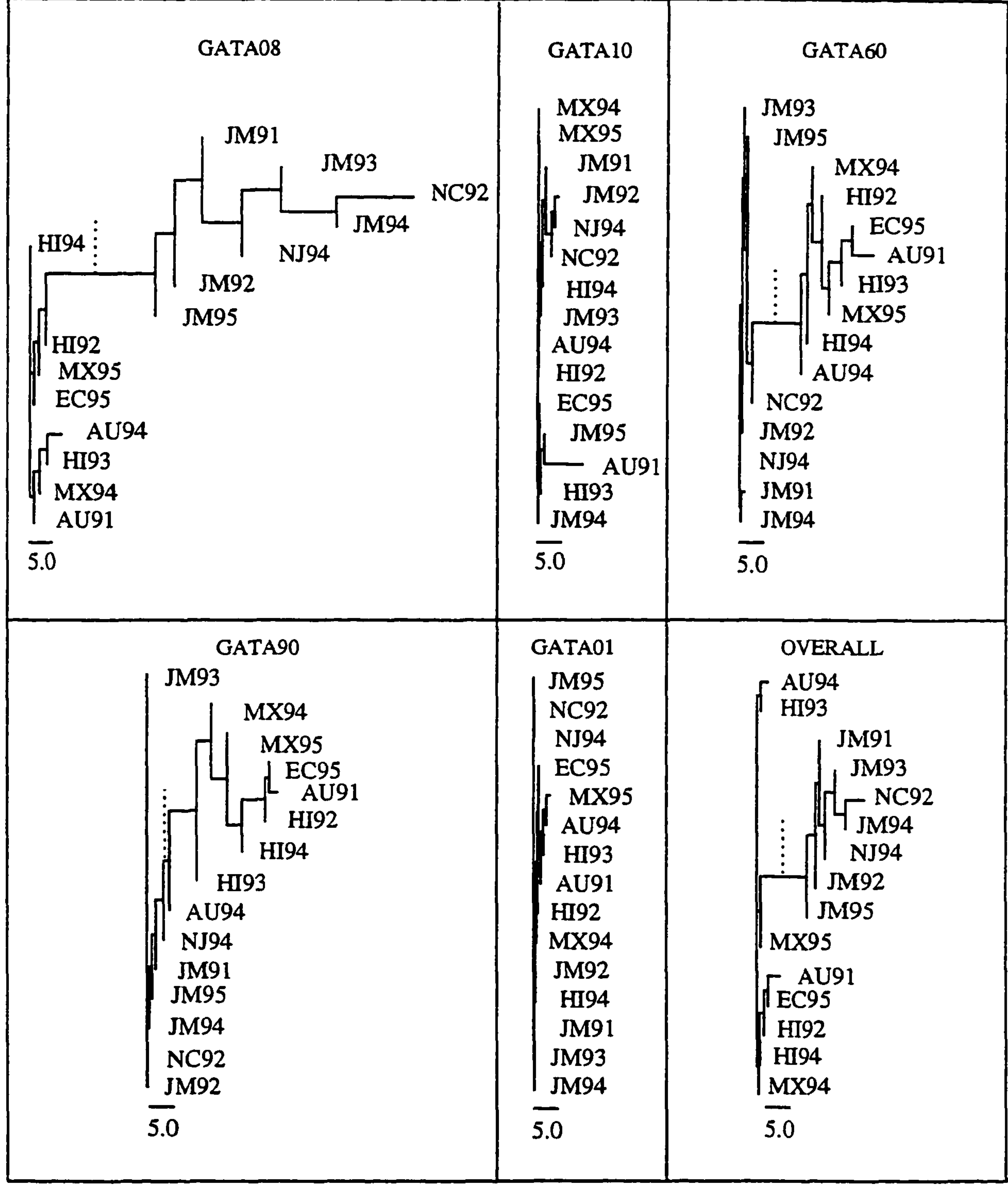
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Linearized θ distance measure for microsatellites

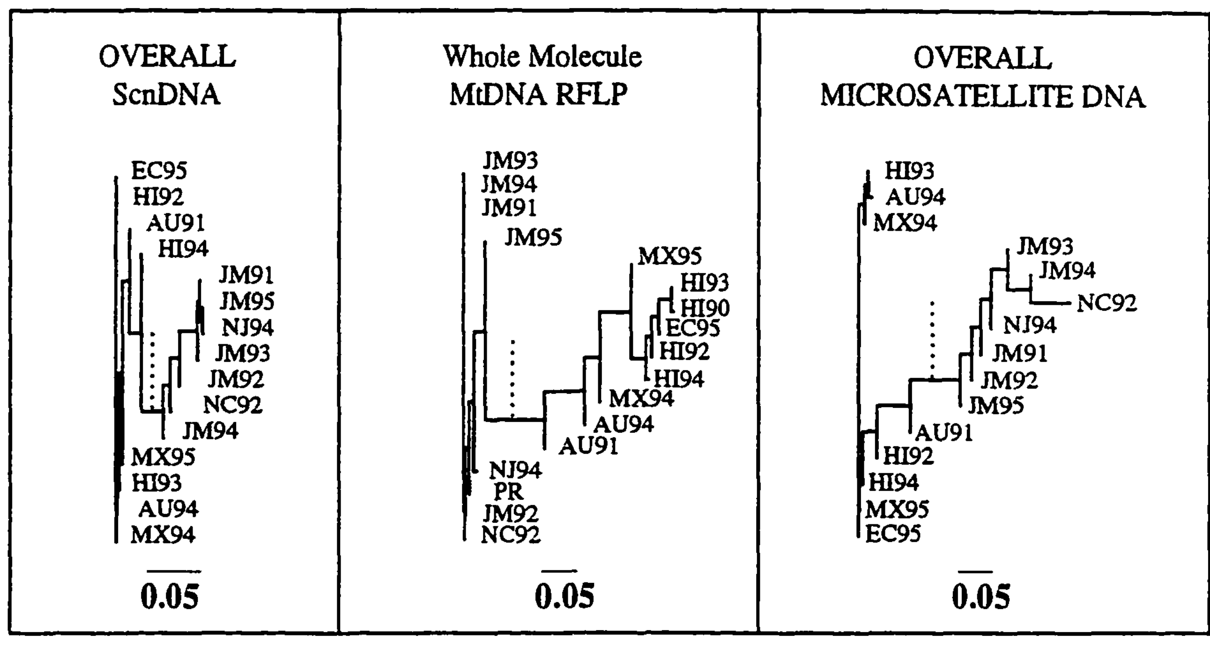


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Delta Mu distance measure for microsatellites



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

FIGURE 10. Results of inter-ocean pairwise tests for population differentiation for microsatellites including (A) exact R, (B) exact probability tests, and for mtDNA (C) exact Φ. Tests were performed among all populations (i), and among geographic locations pooled over temporal replicates (ii). In A and B, for each population comparison, boxes indicate the significance level for GATA01, GATA08, GATA10, GATA60, and GATA90 loci respectively. The initial α level was 0.00045 for each marker and calculation method, corresponding to a total of 105 pairwise comparisons per locus, including inter-annual and intra-ocean tests. Population abbreviations follow Table 1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A. Exact *R* tests: Microsatellites

i.

	JM91	JM92	JM93	JM94	JM95	NC92	NJ94	2n
HI92								68
HI93								90
HI94								50
MX94								96
MX95								48
EC95								38
AU91								20
AU94								32
2n	108	108	84	92	48	24	22	

ii.

	JM	US
HI		
E-Pac		
AU		

P< 0.01, but non-significant after corrections for multiple tests.

Significant after correction for multiple tests.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

B. Exact probability tests: Microsatellites

i.

	JM91	JM92	JM93	JM94	JM95	NC92	NJ94	2n
HI92								68
HI93								90
HI94								50
MX94								96
MX95								48
EC95								38
AU91								20
AU94								32
2n	108	108	84	92	48	24	22	

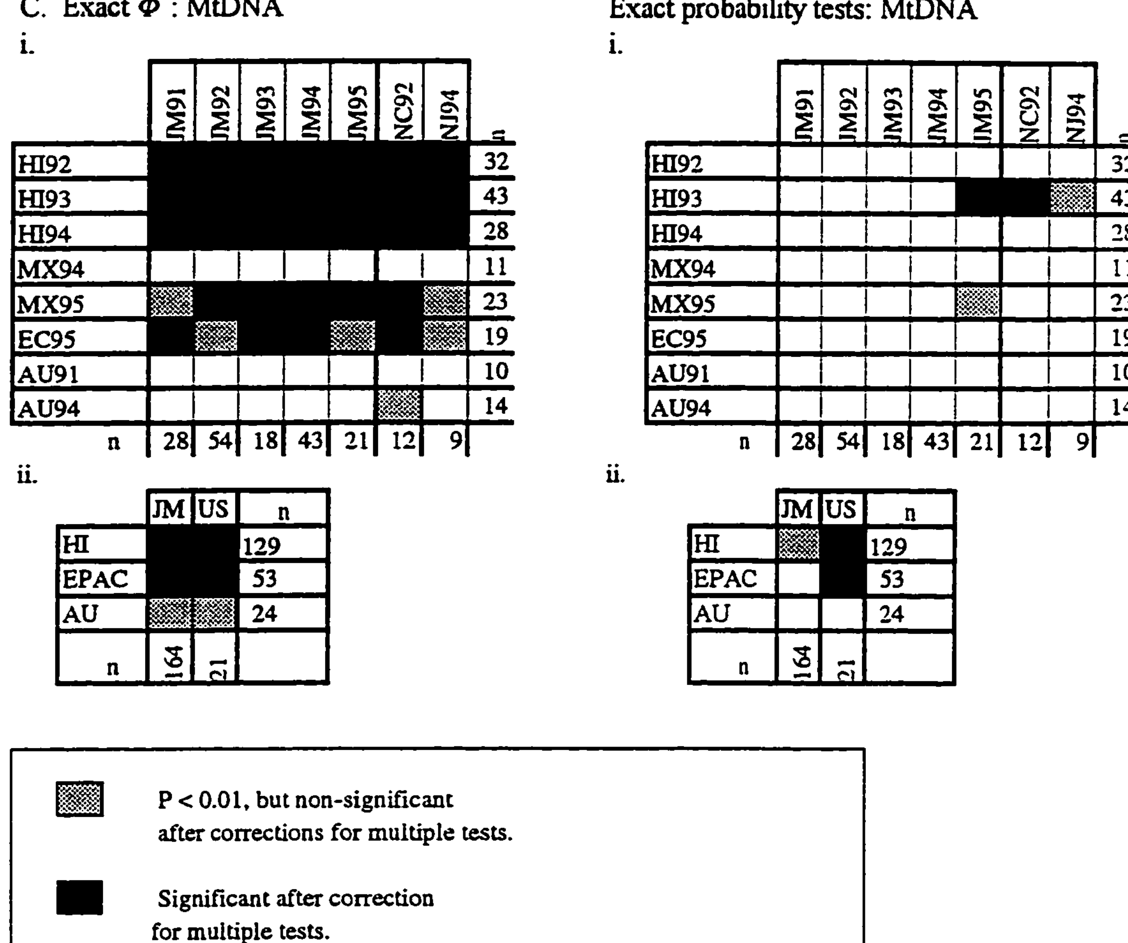
ii.

	JM	USA
HI		
EPAC		
AU		

P< 0.01, but non-significant after corrections for multiple tests.

Significant after correction for multiple tests.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 4. Comparative Evolution of Allozyme, ScnDNA, MtDNA, and Microsatellite DNA Molecular Markers

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Comparative studies of molecular markers can contribute much to our understanding of how the evolutionary history of an organism is recorded by different molecular markers, and how different evolutionary forces shape patterns of molecular diversity within a species. In this chapter, relative mutation rates of allozyme, scnDNA, microsatellite, and mtDNA markers are estimated from diversity measures. Deviations of the genotypic distribution of nuclear markers from Hardy-Weinberg equilibrium are discussed with implications for Mendelian inheritance, occurrence of null alleles at the individual level, and sensitivity to the Wahlund effect at the population level. The ability of each marker class to detect intra-ocean and inter-ocean population structure is assessed through comparison of overall AMOVAs, pairwise comparisons of allele frequencies, and phenetic clustering. Finally, the influence of evolutionary forces on each marker class is discussed.

Mutation Rates

Estimates of mutation rate for markers used in the study varied by four orders of magnitude (0.032 – 42). Estimates based on the proportion of monomorphic loci (Hartl and Clark, 1989), for allozymes ($4N_eu = 0.032$) and scnDNA ($4N_eu = 0.37$) were relatively low. Higher values were estimated for mtDNA ($2N_eu = 4.5$), based on the mismatch mean and use of the infinite sites mutation model (Tajima, 1983), and for microsatellites ($4N_eu = 42$) estimated from twice the repeat variance and the stepwise mutation model (42; Moran, 1975). For the same loci, mutation rates calculated using

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

overall heterozygosity for each marker class yielded similar results (data not shown). Regardless, the marker classes in this study displayed a broad range of estimated mutation rates.

Hardy-Weinberg Equilibrium.

Over all marker classes, deviations from Hardy-Weinberg equilibrium were significant only for sample MX94 at the BM81 locus (heterozygote excess), and HI92, MX95 and JM92 samples at the GATA10 locus (heterozygote deficiencies). This illustrates that allozyme, scnDNA, and microsatellite nuclear markers each displayed genotypic distributions consistent with Mendelian inheritance.

Mixing of genetically distinct populations can result in heterozygote deficiencies within a sample (Wahlund effect). Australian samples were intermediate in allele frequencies at mtDNA, GATA90, BM81, and BM32-2 loci, between Atlantic and other Pacific samples. However, Australian samples were no more deviant from Hardy-Weinberg equilibrium than other samples of the same size. It is likely that much larger sample sizes would be needed detect significant deficiencies in Australian samples, should they exist.

Deviations from Hardy-Weinberg equilibrium also are informative about the accuracy of allele scoring at nuclear loci. Large deviations (> 0.1) were 2.5 times more frequent at scnDNA loci than at microsatellite loci (Mann Whitney U, $p = 0.009$). Furthermore, the difference in average variance of f values between microsatellite (0.037) and scnDNA (0.006) loci was fairly large (Bartlett's heteroscedasticity $X^2 = 3.11$, $p < 0.1$). This suggests that scoring of scnDNA loci may be problematic. One source of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

allele scoring error is the preferential amplification of alternative alleles (O'Connell and Wright, 1997). When alleles are stained with ethidium bromide and visualized under ultra-violet light, sensitivity is limited to products amounting to approximately 10 ng. In contrast, the flourescent detection system used to detect microsatellite alleles in this study is very sensitive, with a detection threshold approximately six orders of magnitude lower than that of ethidium bromide staining. This sensitivity may explain the lower deviation of microsatellite loci from the expectations of Hardy-Weinberg equilibrium.

Marker Comparison: Inter-Annual and Intra-Ocean Divergence

The magnitude of inter-annual differentiation must be accounted for when assigning significance levels of intra-ocean differentiation. For all markers, inter-annual divergence was of the same magnitude as differences between locations within an ocean. This pattern was evident in the overall AMOVA for each class of molecular markers, the level of significance for pairwise comparisons among samples, and the mixing of temporal replicates among locations in phenogram analyses. For all marker classes, the level of temporal variance approached significance for a number of comparisons, setting a substantial level of background noise.

Although very few significant differences were detected in within-ocean comparisons, significant differentiation was detected between Australia and the other Pacific samples at mtDNA. Given moderate levels of inter-ocean gene flow, and the intermediate geographical position of AU samples, one might expect AU gene frequencies to be intermediate to those of the Atlantic and Pacific, as seen for mtDNA. However, AU samples were different from other samples from both oceans for the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

microsatellite loci GATA10 and GATA90. The limitation on this analysis is that both samples were small, allowing for pronounced sampling variance as illustrated by the inconsistent placement of the two AU populations in phenograms.

The definition of intra-ocean stock structure is critical for billfish management (Cramer and Prager, 1994). This study has shown that Australian samples may differ from other Pacific samples, possibly due to an influx of Atlantic migrants. Due to the limitation in sample sizes relative to the variation detected and a small signal between populations, this study is not particularly well suited to definitively address the significance of intra-ocean differentiation (Waples, 1998). Intra-ocean comparisons approaching significance were observed only in cases that involved small populations, suggesting a large role for sampling error. Analysis of mtDNA has revealed significant within-ocean heterogeneity among Atlantic samples of swordfish (Alvarado Bremer et al., 1996, but see Rosel and Block, 1996), and striped marlin (Graves and McDowell, 1994). To adequately address intra-ocean population structure in blue marlin, much larger sample sizes are needed from eastern Australia, Indian Ocean, and south Atlantic Ocean.

Marker Comparison: Inter-Ocean Divergence

The sensitivity of each class of molecular marker to inter-ocean population structure was evaluated using the overall AMOVA, pairwise differences, and construction of trees. Figure 7 (Chapter 3) illustrates the proportion of variance for each level of subdivision revealed by scnDNA, microsatellite DNA, and mtDNA analyses. The division between-oceans (σ^2_P) formed a significant proportion of overall variance for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

each marker class, and varied greatly among loci within marker classes. The largest proportion of variance associated with inter-ocean divergence was detected in the analysis of mtDNA, GATA60, and GATA08 microsatellite loci. The large modal differences between ocean samples at these loci were evident in all subsequent analyses of population structure. Over all loci, the greatest level of divergence was revealed by mtDNA, followed by microsatellite, scnDNA, and allozyme markers (Chapter 3, Table 6C). Variation of inter-ocean divergence estimates within a marker class was greatest for microsatellites, with a standard deviation of 0.11, followed by allozymes (0.064) and scnDNA (0.054) markers. This variation among markers approached significance (Bartlett's test for heteroscedasticity $p < 0.1$)

Using the same populations and individuals, the largest number of significant pairwise comparisons was detected at mtDNA and microsatellite loci, followed by the scnDNA loci. The mtDNA locus was the most divergent marker class on average. A total of 36 of the 56 inter-ocean comparisons was significant despite the presence of 127 alleles, reduced sample sizes at JM92 and MX94, and sample size reduction by a factor of two due to haploid inheritance. An (albeit crude) estimation of the relative power of the mitochondrial data, adjusted for average sample size, shows that the mtDNA tests were at least five times more sensitive than any nuclear test when the divergence between alleles was considered (Chapter 3, Table 7A). However, the performance of mtDNA was greatly reduced when based on frequency data alone. Among nuclear markers, the average number of significant differences per locus was larger for microsatellite (25/56) than scnDNA markers (14.5/56) when the best (most significant) method of analysis (θ, R) was taken for each microsatellite locus, but the mean values were not significantly

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

different (Mann Whitney U, $p = 0.43$). The larger variance among microsatellite divergence estimates observed with the overall AMOVA was evident at this level as well, with standard deviations of microsatellite R estimates (20.5) almost twice as large as those of the scnDNA (12.1). In contrast to the mtDNA data, the microsatellite results revealed similar numbers of significant differences using θ and R measures.

Comparison of phenetic trees was limited to 'perfect' or 'imperfect' clustering of ocean samples. Table 8 (Chapter 3) shows that when the best method was considered, the only loci that were not able to separate populations into oceans correctly, were BM47, BM81, and GATA01. Considering the best method for each locus and marker class, 80% of microsatellite, 50% of the scnDNA, and the single mtDNA locus were able to reconstruct relationships without error.

In summary, microsatellites were more sensitive to inter-ocean population structuring than other nuclear markers. The average microsatellite divergence (R = 0.145) was nearly twice as large as that of scnDNA and allozyme markers ($\theta = 0.08$) although the values were not significantly different. Microsatellites detected a larger percentage of significant inter-ocean differences than single-copy markers regardless of method of calculation, although the mean values were not significantly different. Finally, microsatellites were able to cluster populations into oceans with less error than scnDNA markers. While more sensitive, differences between means were not significant due to the large variance of microsatellite loci. The trend of greater microsatellite divergence indicates that microsatellite homoplasy, which tends to underestimate differences between samples, was not likely a significant factor when considering divergence estimates averaged over loci. More moderately polymorphic molecular markers have the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

statistical advantage that sampling variance in allele frequency estimates is reduced as the number of observations per allele class increases. However, in this study the larger average divergence (effect size) of microsatellites outweighed the statistical advantage of the scnDNA markers, resulting in a higher proportion of significant results. The higher power of microsatellite relative to scnDNA analyses was also reported by Fitzsimmons (1997b) for marine turtles, although sample sizes for scnDNA markers were one third those of microsatellite markers.

Evolutionary Processes

Genetic divergence between blue marlin from the Atlantic and Pacific oceans was much more pronounced at the mitochondrial than nuclear genome. This difference was attributed to the interplay of drift, migration, and mutation among marker classes (Chapter 2). The most appropriate mitochondrial estimator for purposes of marker class comparison was the drift-based estimator that binned clades of alleles into single allele classes. The divergence estimate from this analysis was 0.387, nearly three times that based on microsatellites (0.145), and nearly five times larger than the estimate based on allozyme and scnDNA markers (0.08). While greater male-mediated dispersal is consistent with these results, small sample sizes preclude effective comparisons of male and female inter-ocean gene flow.

Selection. Selection can be inferred through comparison of observed and expected distributions of F values within a marker class. With drift, the distribution of F statistics is predicted to fit a chi-square distribution of n-1 degrees of freedom, where n is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the number of loci (Lewontin and Krakauer, 1973; Pogson et al., 1995; Workman and Niswander, 1970). Using this approach, Pogson et al. (1995) found significantly more variance than expected by drift alone among *F* estimates (scnDNA loci), leading them to hypothesize directional selection on sites linked to the scnDNA loci. For blue marlin, a Kolmogorov-Smirnov (K-S) test (Sokal and Rohlf, 1995) was used to fit observed and expected *F* distributions for each marker class, in intervals calculated as standard deviations from the mean. The K-S test is less sensitive to unbiased estimates of θ that occasionally yield negative values when the true value is close to zero, exceeding the confidence interval of the chi-square distribution. Significant deviations from the chi-square distribution were not detected for scnDNA, allozyme, or microsatellite markers ($p > 0.05$).

Selection also has been inferred from comparisons of the mean and variance of *F* distributions among marker classes (Barker et al., 1997; FitzSimmons et al., 1997; Karl and Avise, 1992; Pogson et al., 1995). Significantly lower mean divergence among allozyme loci than among scnDNA loci was used to implicate balancing selection on allozymes in studies of the Atlantic cod (Pogson et al., 1995) and American oyster (Karl and Avise, 1992; but see McDonald et al., 1996). For blue marlin, means of each marker class (allozyme, scnDNA, and microsatellite) were within a single standard deviation of each other. Microsatellite markers in the blue marlin did display larger variance in inter-ocean divergence estimates (*R*) than did scnDNA loci and allozyme loci, but the differences were not significant statistically (Bartlett's test for heteroscedasticity, $p > 0.05$). For blue marlin, the fit of observed to theoretical (X^2) distribution of *F*-values for each marker class, the non-significant differences in mean values among marker classes,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and the homoscedasticity among marker classes, indicate that the variance expected from drift alone is sufficient to explain the mean and variance of the individual and overall values obtained. These results argue against the hypothesis that anonymous scnDNA polymorphisms are commonly subject to the influence of selection (FitzSimmons et al., 1997; Pogson et al., 1995).

Mutation. If differences in mutation rates significantly affected the inter-ocean genetic divergence in blue marlin, one would expect a correlation between diversity and divergence values. For blue marlin, correlations between heterozygosity and divergence did not differ significantly among loci within each marker class, nor over all nuclear loci (Spearman's rank correlation test, $p > 0.1$). Correlations between divergence estimates and average heterozygosity or estimates of average $4N_e u$ for each marker class, including mtDNA, also were non-significant. Similar findings have been reported by Pogson et al. (1995) and Jorde et al. (1995). Bowcock et al. (1994) and Paetkau et al. (1997) reported significant negative correlations between heterozygosity and divergence for humans and bears, respectively, indicating that high mutation rates may obscure population structure (Jin and Chakraborty, 1995). For blue marlin, absence of significant correlation between diversity and divergence among variable markers suggests that differences in mutation rates do not account for the range of divergence values detected.

Drift. The finding of four-fold greater divergence in mtDNA relative to nuclear DNA can be explained by the four-fold enhanced effect of drift on mtDNA (Birky et al., 1989). However, this explanation assumes that the population (as a whole) has reached a migration/drift equilibrium. For a given rate of migration, the time to reach equilibrium depends on the effective population size (Hartl and Clark, 1989). For blue marlin, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

rate of approach to a migration-drift equilibrium is affected by the unequal effective size of the nuclear and mitochondrial genomes, presumed one-way migration, and potentially different effective population sizes between oceans. For estimators such as θ and R, migration-drift equilibrium is approached at the rate $1/m$, where m represents migration rate in units of migrants per generation (Hartl and Clark, 1989). For populations with effective sizes on the order of 1000 individuals, equilibrium for the nuclear genome is reached within a few hundred generations. Population census sizes for blue marlin are poorly known. Sizes of 40,000 and 400,000 individuals were inferred from Atlantic (Cramer and Prager, 1994) and Pacific (Skillman, 1989) stock production analyses, respectively. If the observed (nuclear) $N_e m$ of 2.25 represents the equilibrium number of migrants and if migration occurs predominantly one-way, it would take approximately 9,000 to 90,000 generations, respectively, for the blue marlin in the two oceans to reach equilibrium (eq. 6.17 of Hartl and Clark, 1989). For mitochondrial genes, equilibrium would be obtained in 2,200 to 22,000 generations. If blue marlin were isolated during the late Pleistocene, i.e. 10,000 to 100,000 years ago, the nuclear and mitochondrial genomes may now be in different states in the approach to equilibrium.

Genetic drift explains many of the observed patterns of molecular diversity in blue marlin. Variance in estimates of F among loci and among marker classes are consistent with variation expected from stochastic processes, and the close match of binned and non-binned R and Φ estimates suggested that drift of clades and modes is more important than mutational events in determining current inter-ocean genetic structure. The role of mutation was de-emphasized, and the difference between nuclear

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and mitochondrial inter-ocean divergence values is concordant with the expected values under migration-drift equilibrium.

Conclusions

1) Inter-annual heterogeneity in allele frequencies approached significance for loci within each marker class, providing a substantial level of noise from which intra-ocean comparisons of allele frequencies must be distinguished

2) Intra-ocean divergence was non-significant in nearly all cases, although mitochondrial evidence that Australian populations may serve as genetic intermediates to Pacific and Atlantic populations was detected. Much larger sample sizes are necessary to adequately address patterns of intra-ocean population structure given the low levels of divergence at this level, high levels of within population variability, and substantial genetic variation among years.

3) Inter-ocean population structuring was highly significant for each marker class. A nuclear estimate of approximately 2.25 absolute migrants was calculated over all marker classes, indicating moderate levels of inter-ocean gene flow.

4) Mitochondrial inter-ocean divergence was larger than nuclear estimates when incorporating allelic relationships (Φ), and when considering the distribution of related groups of alleles.

5) Due to the larger inter-ocean effect-size of several microsatellite loci, microsatellite markers were marginally more efficient than scnDNA markers in detecting population structure. On average, however, mtDNA RFLP analysis remained the most powerful molecular tool.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6) A nuclear counterpart to the mitochondrial 'Atlantic clade' of haplotypes was detected at one microsatellite locus. This finding provided evidence that the Pleistocene isolation similarly affected both mitochondrial and nuclear genomes.

7) Using a number of criteria, selection was not detected among the loci for the markers analyzed. Significant differences were not detected in the mean and variance of the F distribution among different nuclear marker classes. Furthermore, deviations from the theoretical F distribution were not significant. These results indicate that the variance in the F distribution among loci and markers is consistent with that expected from markers predominantly influenced by genetic drift.

8) Differences in mutation rates among loci failed to explain the different levels of inter-ocean divergence detected among loci or marker classes. It is hypothesized that the patterns of modal diversity present within microsatellite and mtDNA allele distributions resulted from mutational events that accumulated during the Pleistocene isolation. The close match of the inter-ocean divergence estimates that incorporate allelic relatedness compared to those that simply described the distribution of modes of alleles demonstrated that mutation has not contributed significantly to inter-ocean divergence since the Pleistocene. It is not clear that migration-drift equilibrium has been reached, given a limited amount of time since major changes in the species range and large population sizes, although a 4:1 ratio in mtDNA to nDNA divergence is in accord with equilibrium expectations.

9) From consideration of migration, mutation, selection, and drift, the current distribution of mitochondrial and nuclear diversity between oceans can be explained by the equilibrium expectations between drift and migration. Mutation and selection may

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

contribute to the diversity of patterns obtained, but their effects cannot be discerned, possibly due to the limited number of loci sampled per marker class. Male mediated gene flow may have contributed to the difference detected among nuclear and mitochondrial markers, but small sizes limited the power of this analysis.

10) Further studies of intra-specific population structure in the blue marlin should focus on obtaining large samples from Australia and Indian Ocean. Further examination of sex-biased dispersal should concentrate on obtaining larger samples of Atlantic females and Pacific males. Further studies of historical phylogeography should focus on rooted sequence analyses of microsatellite modes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LITERATURE CITED

Barker, J. S., S. S. Moore, D. J. Hetzel, D. Evans, S. G. Tan, and K. Byrne. 1997. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein-coding loci. Anim. Genet. 28:103-115.

Birky, C. W., P. Fuerst, and T. Maruyama. 1989. Organelle gene diversity under migration, mutation and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics. 121:613-627.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd, and L. L. Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature. 368:455-457.

Cramer, J., and M. H. Prager. 1994. Exploratory surplus-production analysis of blue and white marlin fisheries in the North Atlantic. Pp. 557-560 *in* ICCAT, ed. Collective volume of scientific papers. Vol. XLI. Report of the second ICCAT billfish workshop. ICCAT, Madrid, Spain.

FitzSimmons, N. N., C. Moritz, C. J. Limpus, L. Pope, and R. Prince. 1997. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. Genetics. 147:1843-1854.

Hartl, D. L., and A. G. Clark. 1989. Principles of Population Genetics. Sinauer Associates, Inc., Sunderland, MA.

Jin, L., and R. Chakraborty. 1995. Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. Heredity. 74:274-285.

Jorde, L. B., M. J. Bamshad, W. S. Watkins, R. Zenger, A. E. Fraley, P. A. Krakowiak, K. D. Carpenter, H. Soodyall, T. Jenkins, and A. R. Rogers. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am. J. Hum. Genet. 57:523-538.

Karl, S. A., and J. C. Avise. 1992. Balancing selection at allozyme loci in oysters: Implications from nuclear RFLP's. Science. 256:100-102.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics. 74:175-195.

McDonald, J. H., B. C. Verrelli, and L. B. Geyer. 1996. Lack of geographic variation in anonymous nuclear polymorphisms in the American oyster, *Crassostrea virginica*. Mol. Biol. Evol. 13:1114-1118.

Moran, P. A. P. 1975. Wandering distributions and their electrophoretic profiles. Theor. Popul. Biol. 8:318-330.

O'Connell, M., and J. M. Wright. 1997. Microsatellite DNA in fishes. Rev. Fish Biol. Fish. 7:331-363.

Paetkau, D., L. P. Waits, P. L. Clarkson, L. Craighead, and C. Strobeck. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics. 147:1943-1957.

Pogson, G. H., K. A. Mesa, and R. G. Boutilier. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics. 139:375-385.

Skillman, R. A. 1989. Status of Pacific billfish stocks. Pp. 179-195 *in* R.H. Stroud, ed. Planning the Future of Billfishes. Research and Management in the 90s and Beyond. Part I. Fishery and stock synopses, data needs and management. National Coalition for Marine Conservation Inc., Savannah, GA.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W. H. Freeman & Co., New York.

Tajima, F. 1983. Evolutionary relationships of DNA sequences in finite populations. Genetics. 105:437-460.

Workman, P. L., and J. D. Niswander. 1970. Population studies on southwest indian tribes. II. Local genetic differentiation in the Papago. Am. J. Hum. Genet. 22:24-49.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

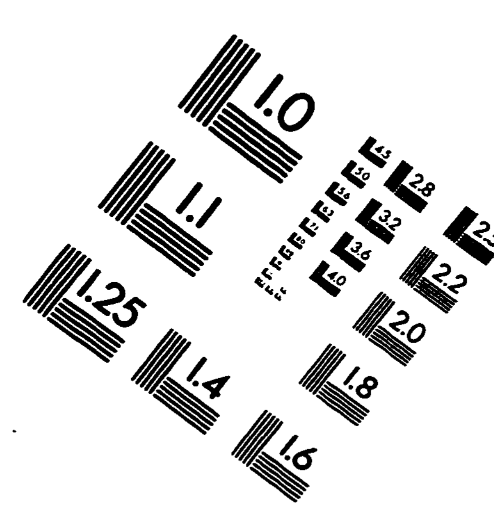
VITA

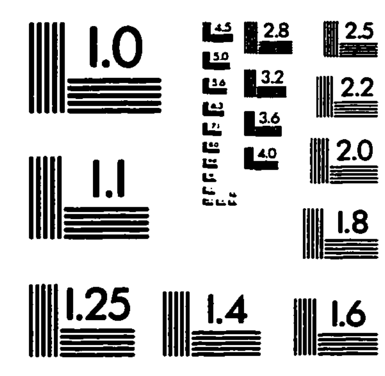
Vincent Patrick Buonaccorsi

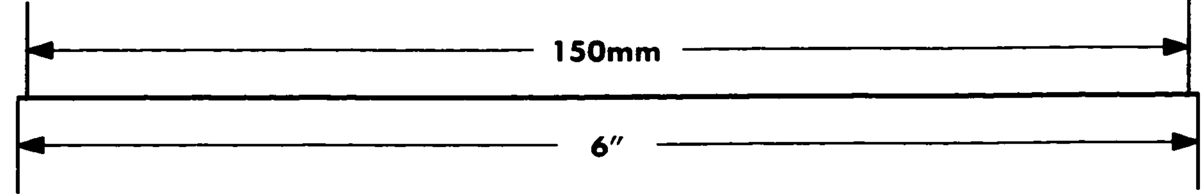
Born in Buffalo, New York, July 21, 1971. Graduated from Loyola Blakefield High School, Baltimore MD, in 1989. Earned a B.S. in Biology, *Magna Cum Laude*, from the University of Notre Dame in 1993. Entered doctoral program in College of William and Mary, School of Marine Science in 1993. Awarded National Research Council Postdoctoral Associateship, at the Southwest Fisheries Science Center, National Marine Fisheries Service, La Jolla CA in 1998.

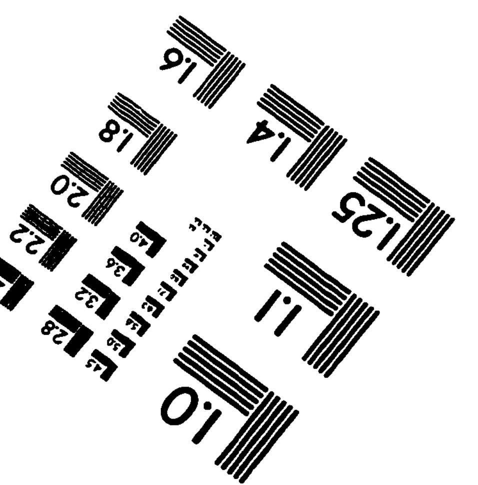
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

IMAGE EVALUATION TEST TARGET (QA-3)

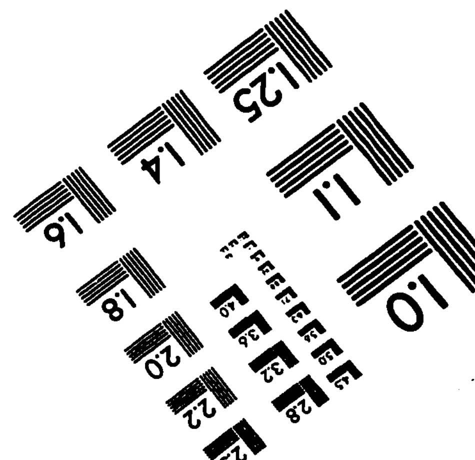








© 1993, Applied Image, Inc., All Rights Reserved



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.