GENETIC STRUCTURE AMONG LAKE MICHIGAN'S LAKE WHITEFISH SPAWNING AGGREGATES

by

Justin A. VanDeHey

A Thesis

submitted in partial fulfillment of the requirements of the degree

MASTER OF SCIENCE

IN

NATURAL RESOURCES (FISHERIES)

College of Natural Resources

UNIVERSITY OF WISCONSIN

Stevens Point, Wisconsin

17 May 2007

APPROVED BY THE GRADUATE COMMITTEE OF

Dr. Brian L. Sloss, Committee Chair
Professor of Fisheries
Wisconsin Cooperative Fishery Research Unit

Dr. Michael J. Hansen
Professor of Fisheries
University of Wisconsin-Stevens Point

Dr. Christopher F. Hartleb
Professor of Fisheries Biology and Aquaculture
University of Wisconsin-Stevens Point

Mr. Paul J. Peeters
Fisheries Team Supervisor
Wisconsin Department of Natural Resources

ABSTRACT

Lake whitefish (*Coregonus clupeaformis*) have comprised an important commercial fishery on Lake Michigan since the early 1800s. Concerns exist regarding the commercial harvest of potentially shared stocks by Michigan and Wisconsin state-licensed and Tribal commercial fishing operations. Previous studies indicated potential stock structure, however, questions still exist regarding the number, identity, distribution, and discreteness of lake whitefish stocks in Lake Michigan. Stocks represent the basic biological unit of focus for fish management efforts and, subsequently, stock delineation is an integral part of sound, science-based, fisheries management programs. The objectives of the study were to determine the usefulness of lake whitefish microsatellite and mitochondrial DNA genetic diversity to discriminate among spawning stocks of lake whitefish and to describe the genetic population structure of spawning lake whitefish aggregates in northern Lake Michigan and Green Bay in terms of genetic stock identification and degree of stock isolation. I assumed that distinct spawning aggregates represented potential stocks and that differences at molecular markers underlie population differentiation. Twelve resolved microsatellite DNA loci exhibited adequate levels of diversity for population differentiation in terms of their allelic richness and heterozygosity, met Hardy-Weinberg equilibrium expectations and, therefore, the utility of the markers was considered sufficient. Two mitochondrial DNA gene regions (D-loop and ND5) exhibited low sequence diversity (π = 0.0002 and 0.0025, respectively) and low numbers of haplotypes (7 and 5, respectively). The observed diversity did not meet *a priori* levels of genetic diversity deemed necessary for population differentiation and the molecule was abandoned for further use. Genetic stock identification using the twelve

microsatellite loci indicated 5-7 potential genetic stocks were present. The resolved stocks corresponded to geographically proximate populations clustering into genetic groups. Analysis of molecular variance suggested six genetic stocks present in the lake as a significant portion of variation was attributable between groups of populations, but not within groups of populations. Within stock analyses (pairwise F_{st}) suggested all but one stock delineated by AMOVA was a stable grouping (i.e., no between population differences) with the exception being the Hog Island, Traverse Bay grouping (NOE stock). This putative stock showed significant differences between the two populations indicating gene flow between the NOR stock (Naubinway and Epoufette) and Hog Island but not between NOR and Traverse Bay. Elk Rapids was the most genetically divergent population and actually masked the isolation by distance that exists between all other populations within Lake Michigan. Genetic mixture analysis showed mixed stock analysis using the resolved stocks and molecular markers can be accomplished with high accuracy. The coupling of genetic data from this study with current demographic data and a comprehensive mixed-stock analysis will allow for more efficient and effective management of this economically and socially important resource.

ACKNOWLEDGMENTS

A special thanks goes to the Great Lakes Fishery Commission who provided the entirety of funding support for this project. I would like to thank my graduate advisor, Dr. Brian Sloss tremendously for his contributions to this project, building me up as a professional, mentoring me both on research and in school, teaching me how to become a good researcher and helping me to be a better person. I would like to thank Paul Peeters for all of his assistance in sampling, for sharing his invaluable knowledge of Lake Michigan and the commercial fishery with me and for also being a great mentor. I would also like to thank the members of my graduate committee, Dr. Michael Hansen and Dr. Christopher Hartleb, for all their valuable input and time on this project, and especially on this thesis. I would also like to thank Dr. Trent Sutton and Eric Volkman (Purdue University), Ken Royseck, Tim Kroeff and Mike Donofrio from the WDNR, Philip Scheenberger, Randy Claramunt, Patrick Hanchin, Chris Schelb, Dawn Dupras, Karen Koval and Richard O'Neal from the MDNR, Eric Olsen (GTBNR), Steve Lenart (LTBBOI), Archie Martell (LRBOI), and Mark Ebener and staff at CORA for all of their assistance in sample collection, knowledge, and insight into his unpublished tagging data. I would like to thank all the commercial fishermen throughout Lake Michigan for their cooperation and help in sample collection, especially the Hickeys, Kings, Petersons, Webergs, Henricksons, Greg Ruleau and Mo Hermes. I would like to thank Andrea Bernard, at the University of Guelph for insight into her research on Lake Ontario's lake whitefish stock structure. I would like to thank Ryan Franckowiak for all his assistance in the laboratory. I would also like to thank Ed Murphy, Mike Hughes, Jeremy Hammen, Eric Schmechel, and the rest of my fellow graduate students for their assistance in the

field and in the lab. I would like to thank Dan Traynor for hard work and dedication as my technician for the 2005 sampling season. I would also like to thank my fiancé Breanna Riedel for her tireless efforts reviewing papers and presentations, listening to me complain and overall in dealing with me.

TABLE OF CONTENTS

TITLE PAGE i

COMMITTEE SIGNATURES ii

ABSTRACT iii

ACKONWLEDGMENTS v

TABLE OF CONTENTS vii

LIST OF TABLES ix

LIST OF FIGURES xi

LIST OF APPENDICES xiii

INTRODUCTION 1

OBJECTIVES 11

MATERIALS AND METHODS 12

Study Design 12
Research Area 12
Sample Collection 13
DNA Extraction 15
Microsatellite Genotyping 15
Mitochondrial DNA Sequencing 18
Analysis – Objective 1 20
Microsatellite genetic diversity measures 20
Hardy-Weinberg equilibrium and linkage disequilibrium 22
Mitochondrial DNA genetic diversity measures 25
Marker utility 26
Analysis – Objective 2 27
Population differentiation based on genic differentiation 28
Genetic distance 29
Analysis of molecular variance 30
Estimates of F_{st} 31
Additional genetic analyses 32
Effective population size 33
Isolation by distance 34

Genetic mixture analysis 35

RESULTS 39

Usefulness of Genetic Markers 39
Microsatellite genetic diversity measures 39
Hardy-Weinberg equilibrium 42
Mitochondrial DNA genetic diversity measures 43
Genetic Stock Identification 44
Population differentiation based on genic differentiation 44
Genetic distance 45
Analysis of molecular variance 46
Estimates of F_{st} 47
Effective population size 47
Isolation by distance 48
Genetic mixture analysis 48

DISCUSSION 50

Usefulness of Genetic Markers 50
Genetic Stock Identification of Lake Michigan Lake Whitefish 52
Population Specific Issues 55
Effective Population Sizes of Lake Michigan Lake Whitefish 57
Genetic Mixture Analysis 58
Management Implications 59

FUTURE RESEARCH 62

Additional Research on Current project 62
Mixed Stock Analysis 63
Historic Stock Structure Identification 63
Standardization of Genetic Data Across the Great Lakes Region 64

SUMMARY 66

LITERATURE CITED 68

LIST OF TABLES

TABLE 1. Lake Michigan commercial fish harvest by kgs and value in the year 2000.. 78

TABLE 2. Great Lakes commercial fish harvest by kgs and value (U.S. dollars) in the year 2000 (from Kinnunen 2003) 79

TABLE 3. a) Microsatellite and b) Mitochondrial DNA primers, primer sequences, observed allele size range in base pairs (microsatellites), number of alleles observed (A; microsatellites), number of haplotypes (H; mtDNA) and references 80

TABLE 4. PCR reaction cocktail recipes, fluorescent labels and thermocycler temperature profiles for all developed multiplexes and singlet PCR reactions 81

TABLE 5. Population statistics for all 12 sampled populations 82

TABLE 6. Allelic richness estimates based on rarefaction (Kalinowski 2005) and mean values for each locus and each population. rarefaction was based on second smallest sample size (n = 70; Hog Island) 83

TABLE 7. Private allelic richness estimates based on rarefaction (Kalinowski 2005) for the second smallest sample size (n=70; Hog Island) for each population at each locus and total private allelic richness values for each locus and for all loci..... 84

TABLE 8. Mitochondrial DNA haplotypes and their percent occurrence, haplotypic diversity and nucleotide diversity for the ND 5 and D-loop gene regions 85

TABLE 9. a) ND 5 haplotype distribution; b) D-loop haplotype distribution..... 86

TABLE 10. Population pairwise comparison of allele frequency distributions (genic differentiation) across all loci 87

TABLE 11. Pairwise genetic distance matrix based on Cavalli-Sforza and Edwards (1967) chord distance D_c..... 88

TABLE 12. Analysis of molecular variance groupings, sum of squares, percent of variation, and p-values 89

TABLE 13. Pairwise F_{st} values (above diagonal) and their corresponding p-values (below diagonal) 92

TABLE 14. Effective population size estimates based on linkage disequilibrium..... 93

TABLE 15. Variations of tests for isolation by distance with Z, R, r^2, and p-values estimated from reduced major axis regression..... 94

TABLE 16. Individual assignment testing to both population and stock of origin estimated using genetic mixture analysis (Kalinowksi 2003)..... 95

TABLE 17. Genetic composition of all sampled populations based on genetic mixture analysis..... 96

TABLE 18. Proportional assignment of simulated commercial harvests using GMA (Kalinowski 2003) 97

LIST OF FIGURES

FIGURE 1. Historical Lake Michigan lake whitefish commercial harvest (Baldwin et al. 2002; from Great Lakes Fishery Commission database). 100

FIGURE 2. Lake whitefish commercial management zones. Zones in Wisconsin waters include WI-1, WI-2, and WI-3. All other zones are in Michigan waters (WFMs) and were originally established by the 1836 Consent Decree 101

FIGURE 3. Non-tribal and tribal commercial harvest of lake whitefish from Michigan waters, reported by statistical grids, Lake Michigan, 1985. 102

FIGURE 4. Lake Michigan commercial fishing grids 103

FIGURE 5. Primary and secondary study sites for Lake Michigan, lake whitefish study. Primary sites (denoted by gray boxes) include the lakeside of the Door County peninsula (WI; NMB) and Big Bay de Noc (BBN) Michigan. Secondary study sites (denoted by gray circles) include Lower Menominee River (MR), Cedar River (CR), Naubinway, (NAU), Epoufette (EPO), Hog Island (HGIS), Traverse Bay (TB), Elk Rapids (ER), Ludington (LUD), Muskegon (MUS) and Saugatuck (SAU) 104

FIGURE 6. Allele frequency distribution for locus Cocl-lav 18. X-axis numbers correspond to populations (1 = Epoufette, 2 = Naubinway, 3 = Traverse Bay, 4 = Elk Rapids, 5 = Hog Island, 6 = Saugatuck, 7 = Muskegon, 8 = Ludington, 9 = Menominee River, 10 = Cedar River, 11 = NMB, 12 = BBN. Y-axis numbers correspond to allele size in base pairs. The size of each circle represents the frequency of that allele in each population 105

FIGURE 7. Allele frequency distribution for locus Cocl-23. X-axis numbers correspond to populations (1 = Epoufette, 2 = Naubinway, 3 = Traverse Bay, 4 = Elk Rapids, 5 = Hog Island, 6 = Saugatuck, 7 = Muskegon, 8 = Ludington, 9 = Menominee River, 10 = Cedar River, 11 = NMB, 12 = BBN. Y-axis numbers correspond to allele size in base pairs. The size of each circle represents the frequency of that allele in each population 106

FIGURE 8. Neighbor-joining tree based on Cavalli-Sforza and Edwards (1967) chord distance (D_c). Branch support represents the percent recovery of that node based on 5,000 bootstrap pseudoreplicates. Support ≤50% is not shown. Most likely genetic groupings of populations are indicated by gray ovals. 107

FIGURE 9. Neighbor-joining tree based on Cavalli-Sforza and Edwards (1967) chord distance (D_c) without the Menominee River sample. Branch support represents the percent recovery of that node based on 5,000 bootstrap pseudoreplicates. Support ≤50% is not shown. Most likely genetic groupings of populations are indicated by gray ovals 108

FIGURE 10. Isolation by distance for all populations. Figure 10a is genetic (D_c) distance against geographic distance (kms). Figure 10b is the log genetic distance versus log geographic distance. Regression line was calculated using reduced major axis regression 109

FIGURE 11. Isolation by distance without the Menominee River population to determine potential influences from an admixture population. Figure 11a is genetic (D_c) distance versus geographic (kms) distance. Figure 11b is log genetic versus log geographic distance. Regression line was calculated using reduced major axis regression 110

FIGURE 12. Isolation by distance without the Elk Rapids population. Figure 12a is genetic (D_c) distance versus geographic (kms) distance. Figure 12b is log genetic versus log geographic distance. Regression line was calculated using reduced major axis regression 111

FIGURE 13. Plot of residuals from simulated genetic mixture analysis. Outside dashed lines represent the largest residual values. Inside dashed lines represent a 1% error mark. twenty-two of 36 comparisons yielded an error of ≤ 1% 112

FIGURE 14. Six genetic management zones (GMZs) based on genetic stock identification and spawning site locations, over-layed on the statistical and commercial management zones currently in place for Lake Michigan 113

LIST OF APPENDICES

APPENDIX 1. Allele frequencies for all loci for each population..... 114

INTRODUCTION

Fisheries scientists and managers have long been concerned about population structure and the implications of subcomponents of fishery harvest (i.e., multiple species, different stocks, or both) for sustainable, long-term management of the resource(s). "Despite the fact that species are fundamental taxonomic and biological units, they are seldom entirely panmictic. Rather, sexually reproducing species typically are composed of subpopulations (stocks) that are at least partially reproductively isolated and differentiated from one another," (Shaklee and Currens 2003). This partial reproductive isolation, when coupled with temporal and spatial restrictions in the distribution of subpopulations, provides the basis for local adaptation through natural selection and is the central tenet of the stock concept.

The utility of the stock concept in resource management is based on scientific evidence showing the overall productivity and evolutionary potential of a species is dependant on maintaining the abundance and diversity of its component stocks (Shaklee and Currens 2003). Contemporary fisheries management has relied on the stock concept and the operational definition of a stock as "local populations that maintain recognizable genetic differentiation by separation of their spawning time or place" (Bailey and Smith 1981). As such, stocks are usually composed of a population(s) of fish spawning in the same location and time and exhibiting similar growth rates, mortality rates, morphological/meristic features, and age structure (Van den Avyle 1993). Thus, a fish stock represents a genetically cohesive aggregate of fish exhibiting population dynamics and genetic characteristics that are independent of other stocks. Stocks represent the basic biological unit for fish management efforts and, subsequently, stock discrimination

is an integral part of sound, science-based fisheries management programs (Spangler et al. 1981).

Lake whitefish (*Coregonus clupeaformis*) are members of the Salmonidae and are native to the Great Lakes basin, all Canadian provinces, and Alaska. Lake whitefish have silver sides and a green to brown back with relatively clear fins. They are benthic feeders that primarily feed on macroinvertebrates such as *Diporeia* spp. and Chironomidae. Average length of an adult lake whitefish is approximately 50 cm with a mean weight of approximately 1.25 kg (Wisconsin Sea Grant 2007). The average age of commercially harvested (i.e., mature) lake whitefish in Lake Michigan during the 2005 harvest year was between 6 and 7 years in Wisconsin waters (P. J. Peeters, WDNR personal communication) and between 7 and 8 years in Michigan waters (Wright and Ebener, in press; P. J. Schneeberger, MDNR, personal communication).

Lake whitefish exhibit a common reproductive ecology with other Great Lakes salmonids in terms of timing of spawn (Becker 1983) and philopatry to their natal spawning grounds (Ebener 1980; Ebener and Copes 1985; Scheerer et al. 1985; Walker et al. 1993). Lake whitefish generally spawn from late October through early December on near-shore reefs consisting of gravel or small stones at depths of approximately 2-18 m (Becker 1983). While spawning, two or three males swim alongside a single female and both sexes broadcast their gametes over the spawning reef. Tagging studies showed a high degree of site fidelity in terms of mark-recapture of spawning lake whitefish across multiple years from various spawning aggregates in Lake Michigan (Ebener 1980; Rowe 1984; Ebener and Copes 1985; Scheerer et al. 1985; Walker et al. 1993). Assuming this site-fidelity is consistent with the reef(s) where the individuals were themselves spawned

(i.e., natal philopatry) and only low-levels of straying among reefs occur, population/ stock-specific differences should exist across large geographic regions such as the Lake Michigan basin.

The lake whitefish has been at the heart of the Great Lakes commercial fishery since the 1800s. In Lake Michigan, lake whitefish support both state-licensed and tribal commercial fisheries in Michigan and Wisconsin. The 2000 harvest season was the highest on record and produced 2,174,096 kg of fish on Lake Michigan alone with a dockside value of over $5 million US (Kinnunen 2003; Table 1). Presently, lake whitefish comprise the primary commercial fishery on the Great Lakes in terms of total kgs harvested and total dockside value (Table 2). In the 2002-2003 quota year, the commercial harvest of lake whitefish from Wisconsin waters was 600,100 kg (WDNR 2004) and harvest from Michigan waters was 241,331 kg (Schneeberger 2004). In addition to the sale of lake whitefish flesh, a market exists for lake whitefish roe. Roe sales for 2001 - 2004 averaged $8.82/kg for Wisconsin commercial fishermen (A. D. Blizel, WDNR, personal communication).

These harvest levels represent a remarkable recovery of a once-diminished resource. In the mid-1800s through the early-1900s, a variety of stressors including over-fishing, loss of habitat and introduction or invasion of non-indigenous species, led to the collapse of the lake whitefish fishery in the Great Lakes basin (Reckhan 1995; Spangler and Peters 1995; Ebener 1997; Figure 1). This recovery has been credited in part to sea lamprey control, more intense and focused management, and the Clean Water Act of 1977.

Presently, Lake Michigan lake whitefish are managed through the use of

management zones and quotas rather than being managed on a stock basis. Management of commercial fishing on Lake Michigan is an inter-jurisdictional effort with the Department of Natural Resources from Wisconsin (WDNR), Michigan (MDNR), Illinois (IDNR), and Indiana (INDNR) all managing part of the lake. Tribal resource management and concerns are handled by the Chippewa Ottawa Resource Authority (CORA), which also supervises the activities of individual tribes such as the Grand Traverse Band of Indians (GTBNR), the Little Traverse Bay Bands of Odawa Indians (LTBBODI), and the Little River Band of Ottawa Indians (LRBOI). Most of the commercial lake whitefish fishery in Lake Michigan falls under the control of the WDNR, MDNR, and CORA. Lake Michigan currently has 13 commercial lake whitefish management zones (Figure 2). Wisconsin has three lake whitefish management zones and Michigan has 10 lake whitefish management zones. Most of the Wisconsin commercial fishing falls within one zone (WI-2; Figure 2; P.J. Peeters, personal communication), whereas Michigan's harvest is focused on six zones (WFM-00, WFM-01, WFM-02, WFM-03, WFM-04 and WFM-05; Figure 3). Wisconsin's commercial management zones were created based on commercial fisheries present for different species (e.g., the yellow perch, *Perca flavescens*, fishery in Green Bay corresponds to WI-1) and the borders of these zones follow Lake Michigan's grid system (Figure 4). Michigan's commercial lake whitefish management zones were created based on mark-recapture studies, knowledge of the locations of spawning populations, and fishing patterns of the commercial fishery, with the knowledge that several reproductively isolated populations likely occur within some management units (Ebener et al. 2005).

Despite the commercial reliance on lake whitefish, questions abound regarding

the source (i.e., stock) of fish being harvested. In particular, there are concerns regarding shared stocks potentially being harvested by Michigan and Wisconsin state-licensed and native commercial fishing operations. One concern is over the different management techniques currently utilized by the different agencies. For example, Wisconsin's commercial harvest is reported in kg of dressed fish (viscera removed) and Michigan's harvest is reported in round weight (viscera not removed) (Peeters 2001). A second concern involves differences in allocations of quotas between Michigan and Wisconsin. Wisconsin allocates portions of their total quota to commercial management zones (Figure 2) rather than to lake whitefish management zones used by Michigan (Figure 2). Of Wisconsin's total annual harvest (2.47 million lbs; P.J. Peeters, WDNR, personal communication), 9% is allocated to WI-1, 82% to WI-2, and the remaining 9% to WI-3. Individually assigned quotas are distributed to fishermen within each commercial management zone based on historical catch rates and the total annual quota within each zone. In Michigan, harvest is initially divided based on the 1836 Consent Decree between the Tribes and the state of Michigan wherein the MDNR manages the harvest based on statistical catch at age models within the statistical management zones (P.J. Schneeberger MDNR, personal communication). Within each zone, harvest is then divided among license holders based on their five year mean harvest. CORA enforces regulations on the tribal fishermen and manages their portion of the fishery based on a statistical catch at age model allowing no more than 65% mortality on the most vulnerable age class (M.P. Ebener, CORA, personal communication). Individual quotas within the management zones are then based on the statistical catch at age models.

A final difference in management approach involves commercial season length

and timing. Due to the philopatric nature of lake whitefish (Ebener 1980; Ebener and Copes 1985; Walker et al. 1993) and their tendency to congregate in vast numbers on spawning reefs (Becker 1983), commercial fishing seasons close on Lake Michigan to protect spawning fish and prevent over-harvest. Wisconsin's season is closed from October 26th through November 30th (P.J. Peeters, WDNR, personal communication), whereas Michigan's season is closed from November 1st through November 30th (P.J. Schneeberger, MDNR, personal communication), and the Tribal fishermen's season is closed from November 6th at 2400 hours until November 29th at 2400 hours (M.P. Ebener, CORA, personal communication).

The highly migratory nature of Lake Michigan lake whitefish (M.P. Ebener, CORA, personal communication) lends itself to a high probability of a mixed-stock fishery and the commercial exploitation of fish across management zone boundaries. This coupled with the differences in management strategies, and the inter-jurisdictional nature of this fishery could potentially be problematic for management of this highly exploited commercial resource.

Presently, there is some evidence for multiple lake whitefish stocks in and around the Green Bay area and northern Lake Michigan (Figure 4; Imhoff 1977; Leary 1979; Ebener 1980; Imhoff et al. 1980; Ebener and Copes 1985; Casselman et al. 2001; P.J. Schneeberger, MDNR, personal communication; P.J. Peeters, WDNR, personal communication; M. P. Ebener, CORA, personal communication). The evidence is based primarily on the presence of known spawning aggregates of lake whitefish in these locations coupled with the philopatric nature of lake whitefish (Ebener 1980; Ebener and Copes 1985; Walker et al. 1993). The two major stocks thought to be present in this area

are the North and Moonlight Bay stock (NMB) and the Big Bay de Noc stock (BBN). The NMB fish spawn within and around North and Moonlight Bays and Bailey's Harbor on the lake-side of the Door County (WI) peninsula. The BBN fish spawn within Big Bay de Noc, located in Michigan waters of northern Green Bay. Other potential secondary stocks in the Green Bay area (i.e., smaller known or suspected spawning aggregates) have been identified in the Bark River and Cedar River area in Green Bay (P.J. Schneeberger, MDNR, personal communication) which is also in Michigan waters and at the Lower Menominee River (M.C. Donofrio, WDNR, personal communication) that forms the border between Michigan and Wisconsin. Evidence of potential secondary stocks has also been found at Naubinway, Epoufette, Hog Island, Traverse Bay, Elk Rapids, Ludington, Muskegon and Saugatuck, all in Michigan waters.

A primary management issue surrounding these potential stocks is whether or not they represent discrete stocks and, if so, is the lake whitefish fishery a mixed-stock fishery? Tagging studies have shown that both the NMB and the BBN fish exhibit high spawning site fidelity and philopatry during the fall spawning season but are highly migratory throughout much of the rest of the year (M. P. Ebener, CORA, personal communication). If spawning aggregates represent distinct ecological and genetic entities, then these patterns of movements suggest a strong likelihood of a mixed-stock fishery existing in Green Bay and northern Lake Michigan. Because fish do not follow any jurisdictional boundaries, fisheries managers need to know whether or not these stocks are a shared resource.

To better manage this important resource, delineation of lake whitefish stock structure in Lake Michigan is necessary. Optimal long-term management and

conservation depends on knowing the number, distribution, and characteristics of all component stocks and maintaining their integrity, diversity, and abundance. The more subdivided and diverse a species is, the more its overall productivity and evolutionary potential are likely to depend on maintaining its various components (Shaklee and Currens 2003). Stock identification can be accomplished through several means including population dynamics and life history, tagging studies, and genetics (Shaklee and Currens 2003). A strength of genetic stock identification (GSI) is that it allows direct statistical testing of hypotheses regarding stock structure (Shaklee and Currens 2003). The null hypothesis in all studies of population structure is panmixia; that is, all individuals are randomly interbreeding. A detailed genetic inventory and description of stocks is an essential prerequisite for meaningful, genetically based stock management. Therefore, GSI methods can and should be used when developing a stock inventory (Shaklee and Currens 2003).

Previous genetic studies of lake whitefish stock structure in Lake Michigan (Imhoff 1977; Leary 1979; Imhoff et al. 1980) found some evidence of stocks in Green Bay and northern Lake Michigan but were hampered by a lack of variation at the allozyme markers assayed. Leary (1979) and Imhoff (1977) looked at three isozyme markers: L-lactate dehydrogenase (LDH), malate dehydrogenase (MDH) and glycerol-3-phosphate dehydrogenase (G3PDH). These studies had a low level of observed polymorphism (mean allele frequency of the most common *LDH-B** allele was 96.1%) and had no significant heterogeneity at the *LDH-B** allele when the populations east of Seul Choix Point were removed. Also, the scoring of the genotypes for *MDH** was complicated by suspected polyploidy issues associated with the salmonid genome

duplication (Bailey et al. 1969). As such, the genetic diversity at this locus was assayed based solely on genotypic variation and not on allele frequencies resulting in low confidence in inferred genotypes and putative stock structure based on this locus. Given the low level of observed polymorphism in the data and the relatively low levels of heterozygosity observed in the studies, these studies suffered from a lack of power to statistically evaluate differentiation among putative stocks. The spatial and temporal components of the study by Leary (1979) were further complicated by the sampling of unripe fish. Because lake whitefish exhibit philopatry, the use of unripe fish may have resulted in some fish being assigned to the wrong spawning location. This caused extensive overlap in the inferred population range even during the spawning season. This led to an overlap in the data on different populations (stocks) spawning in the same area, depth, and time. The result was the actual sampling units (sites) did not represent the intended target (i.e., spawning aggregates).

Stock identification and delineation has been identified as a priority by both the Lake Michigan Lake Whitefish Task Group as well as the Lake Michigan Technical Committee (Casselman et al. 2001). If discrete genetic stocks of lake whitefish exist within Lake Michigan, accurate delineation of stocks will enable fisheries managers to manage the lake whitefish resource on an individual stock basis, which will help to maintain genetic diversity of the populations. The delineation of stocks coupled with demographic information would allow managers to identify if multiple stocks occur within a single management zone and subsequently manage for the smallest stock to prevent over harvest and a loss of genetic diversity. Additionally, stock identification would provide a foundation for comprehensive mixed stock analysis that could determine

proportional stock harvest in each of the contemporary commercial management zones (Figure 2) throughout the fishing season. The combination of comprehensive mixed stock analysis and statistical catch-at-age (SCAA) models would provide the basis for effective stock management of Lake Michigan's lake whitefish and ensure the long-term sustainability of the fishery, continued overall productivity in catch, and the maintenance of adaptive potential in Lake Michigan lake whitefish.

OBJECTIVES

The goal of this study was to determine if discrete genetic stocks of lake whitefish exist in Lake Michigan. My first objective was to determine if lake whitefish microsatellite and mitochondrial DNA genetic diversity within and among known lake whitefish spawning aggregates was sufficient for population/gene pool discrimination. My second objective was to determine if genetic population structure exists among the spawning aggregates and if this structure was sufficient to recognize lake whitefish genetic stocks.

MATERIALS AND METHODS

Study Design

There were two assumptions made in this study. Based on the philopatric nature of salmonids and evidence from tagging studies that Lake Michigan lake whitefish home to natal spawning grounds (Ebener 1980; Ebener and Copes 1985; Walker et al. 1993), I assumed spawning aggregates represented potential genetic units. Therefore, I attempted to identify and sample the known spawning reefs of lake whitefish throughout Lake Michigan. Where possible, multiple year samples were included to minimize any bias associated with a single year estimate. Temporal genetic differences exist within some spawning runs of salmonid fish (Wilmot and Burger 1985; Burger et al. 1997; Waples et al. 2004). Therefore, to assess the impact of potential temporal genetic differences across a lake whitefish spawn, two temporal samples (n = 75) were obtained annually from NMB and BBN, the two primary study sites (discussed below). If there were genetic differences between the samples, two putative stocks of lake whitefish could exist. If no differences existed, the runs were considered to represent a single genetic unit.

My second assumption was differences in genetic diversity within and among these spawning aggregates was consistent with the aggregates representing separate gene pools. The delineation of separate gene pools (i.e., stocks) is the central theme of the stock concept. Therefore, if spawning aggregates represent unique and discernible gene pools they could be considered and tested as genetic stocks.

Research Area

The primary research area was located in BBN and the NMB spawning reef

complex (Figure 5). This primary study area is the center of the commercial lake whitefish fishing industry (both tribal and state-licensed; Figure 3) and is thought to represent the primary spawning areas for Lake Michigan lake whitefish (P.J. Peeters, WDNR personal communication; P.J. Schneeberger, MDNR, personal communication). Secondary study sites were found at various locations throughout the lake known to contain smaller lake whitefish spawning aggregates including: Naubinway, Epoufette, Hog Island, Traverse Bay, Elk Rapids, Ludington, Muskegon, Saugatuck, and Cedar River in Michigan, and the Lower Menominee River, in Wisconsin (Figure 5).

Sample Collection

Sample collection of commercial fish occurred during late October through November of the 2005 and 2006 commercial fishing season. Primary study sites (NMB and BBN) and some secondary study sites (Cedar River, Naubinway, Epoufette, Ludington, Muskegon, and Saugatuck) were sampled during both 2005 and 2006. Secondary spawning sites were sampled more opportunistically based on the availability of cooperating commercial fishermen. As such, some sites were not able to be sampled in both study years. Sampling occurred during the months of October and November and only fish that could be positively identified as sexually mature (i.e., presence of gametes) were included in the sample to maximize the probability the samples were from the spawning aggregate associated with that location (Ebener and Copes 1985). Primary study sites (BBN and NMB) were sampled throughout the course of the spawn to determine if any genetic differences existed between early versus late run fish. Samples were acquired from commercial fishermen, in conjunction with agency (WDNR and

MDNR) sampling and commercial monitors, and from tribal harvest and tribal monitors (CORA). The majority of sampled fish were obtained from trap nets and a smaller proportion of samples were obtained from gill nets. A single sample was obtained from the Muskegon (2006) reef by angling. All samples consisted of un-sorted (i.e., no intentional size discrimination) commercial catch sampled dockside to eliminate potential size-related bias in genetic diversity estimates. Lake-wide samples from known spawning grounds in the secondary study sites were obtained in collaboration with an ongoing research project by Dr. Trent Sutton (Purdue University) and through the cooperation of the MDNR, CORA, GTBNR, LTBBOI, and the LRBOI. Individual fishing grids (Figure 4) were noted to assign sampled fish to the location where they were harvested and to identify potential fine-scale stock structure within a single spawning reef. By law, commercial operations have to document the individual fishing grids where their catch was taken. Samples were collected during the closed spawning season through the WDNR fall graded-mesh gill net assessment. Pelvic fin clips were taken from all fish, placed in individually labeled tubes with 95% ethanol, and transported to the Molecular Conservation Genetics Laboratory (MCGL) at the University of Wisconsin-Stevens Point for subsequent genetic analysis.

Determination of *a priori* power for genetic analysis is difficult due to variation in the number of alleles at each locus and the relative frequencies of the alleles. However, a minimum of 75 fish per site were sampled to approach a predicted statistical power equivalent to $1-\beta = 0.80$, reduce bias, and reduce confidence intervals when estimating population structure and genetic distance based on the recommendations of Ruzzante (1998) and Ryman et al (2006).

DNA Extraction

Total genomic DNA was isolated from the pelvic fin tissue using the Promega Wizard® Genomic DNA purification kit (Promega Corp., Madison, WI) following their recommended protocol. The extracted DNA was electrophoresed in a 1% agarose gel in the presence of ethidium bromide and visualized using UV-light. This allowed the relative quality of the DNA to be determined by comparing the relative molecular weight of the DNA compared to a commercially available standard. The DNA was then quantified using a Nanodrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE) and subsequently normalized to a standard concentration of 20 ng/µl to allow for more consistent amplification of the DNA during multi-locus genotyping.

Microsatellite Genotyping

Microsatellites are 2-8 base-pair, tandemly-repeating motifs of DNA that are thought to be randomly distributed throughout the nuclear genome (Tautz 1989). An example of a microsatellite would be: ACACACACACACACACACACACACAC, abbreviated AC_{14}. Microsatellites vary in length (number of repeats) at different loci and in different individuals. In fisheries biology, these markers are widely used to assess population structure and gene flow, stock identification, levels of inbreeding, effective population size of stocks, parentage, and other quantitative traits (Neff et al. 2000). Benefits of microsatellites versus other genetic markers, especially allozymes, include higher mutation rates, non-lethal sampling, high level of detectable variation, and efficient semi-automated analysis of a large number of samples (Shaklee and Currrens 2003).

Genetic diversity at microsatellite loci was assessed through the use of the polymerase chain reaction (PCR) and locus-specific primers. The PCR is a process used to make multiple copies of a single region of DNA (i.e., DNA amplification). Primers were selected from a suite of 31 lake whitefish microsatellite primer sets (Rogers et al. 2004) as well as primers designed for brook trout (*Salvelinus fontinalis*) (Angers et al. 1995; Hansen et al. 1999), and broad whitefish (*C. nasus*) (Patton et al. 1997; Hansen et al. 1999; Lu and Bernatchez 1999). I selected 20 loci for an initial survey of genetic diversity in Lake Michigan lake whitefish based primarily on the number of repeat motifs each microsatellite was expected to have in an attempt to maximize the likelihood of seeing allelic diversity. Selection criterion included microsatellites 8-40 repeat units to provide reasonable genetic variation (i.e., 2-20 alleles/locus) and accurately measure allele frequencies. I initially analyzed the microsatellites for length variation (i.e., genetic diversity) using unlabeled primers. PCR was conducted on samples (n = 10) from the two primary study populations to ensure the locus-specific PCR conditions (Angers et al. 1995; Patton et al. 1997; Hansen et al. 1999; Lu and Bernatchez 1999; Rogers et al. 2004) amplified under our laboratory conditions. The amplicons were electrophoresed on a 2% agarose gel, stained with ethidium bromide, and visualized under UV light. A locus was considered usable if a product (i.e., fuzzy band) consistently appeared on the gel with no non-specific banding.

Thirteen microsatellite loci were selected to effectively screen total population samples for genetic diversity (Table 3). The final number of loci was largely based on a combination of factors especially the number of alleles per locus, desired error level (α = 0.05) for population differentiation, desired statistical power, and available sample sizes.

In addition the final count of 13 loci was based on the polymorphism of the loci and the models created by Bernatchez and Duchesne (2000) who found that population assignment and differentiation success was a function of both the number and properties of loci. This study suggested that based on my sample sizes, ≥10 polymorphic microsatellite loci should be sufficient for population differentiation.

When conducting multi-locus genotyping, multiplex PCR is preferred over standard PCR because it allows the amplification and analysis of multiple loci at one time thereby reducing the cost of genetic analysis. I developed four multiplex reactions for this study (Table 4). Loci were surveyed by conducting PCR amplification with fluorescently labeled primers. In an attempt to standardize data for the Great Lakes Basin, initial PCR reaction conditions followed the Great Lakes Science Center standard operating procedure (W. Stott, USGS, personal communication), with optimization following each sequential run until primer optimization was completed. Initial annealing temperatures were specific for individual primers and were based on Rogers et al. (2004) and Lu et al. (2001). If the recommended annealing temperatures yielded non-specific products, a temperature gradient (e.g., 50-60°C in ½ degree increments) was established for that specific locus to establish an optimal annealing temperature (Table 4). Some loci failed to work in any multiplex and were amplified separately. Specific PCR reaction conditions used in this study are shown in Table 4. All PCR reactions were 10 μL total volume.

Thermocycler conditions followed one of two general approaches. Multiplex 1 and 2 followed a standard denaturation, annealing, extension program for 30 cycles with only the annealing temperature varying (Table 4). The remaining two multiplexes and

the single locus reactions used a touchdown PCR approach (Don et al. 1991) where initial annealing temperatures were set higher than those suggested by the literature (Angers et al. 1995; Hansen et al. 1999; Patton et al. 1997; Lu and Bernatchez 1999; Rogers et al. 2004) and reduced by 0.5-1° C every 5-8 cycles to improve annealing efficiency.

Following PCR, the resulting amplicons were separated by polyacrylamide gel-electrophoresis and visualized on an ABI Prism® 377XL DNA sequencer (Applied Biosystems Inc., Foster City, CA). Allele sizes were determined by comparison to an internal size standard (GeneFlo™ 625, Chimerx Inc., Milwaukee, WI) and estimated using GeneScan® software (Applied Biosystems Inc.). All allele calls were confirmed manually and the resulting data represented multi-locus genotype data.

Mitochondrial DNA Sequencing

In philopatric fish species analyses of mitochondrial DNA (mtDNA) genetic diversity often result in more resolution of genetic structure among populations compared to nuclear diversity (Billington and Hebert 1991). I surveyed the mtDNA genetic diversity of lake whitefish to determine the molecule's effectiveness as a potential genetic marker for stock delineation. Substantial mtDNA variation in many fish species has served as a basis for stock discrimination and incongruent patterns of genetic structuring in mtDNA versus nuclear DNA (microsatellites) are not uncommon (Billington and Hebert 1991; Billington 2003). Because mtDNA shows maternal inheritance the molecule is haploid resulting in mtDNA having 25% the effective size of nuclear DNA (Billington 2003). Mitochondrial DNA is not subject to recombination, and therefore, provides a set of completely linked, homologous markers that permit clear definition of

maternal genealogies. The mtDNA mutation rate is 5 to 10 times faster than the nuclear genome (non-microsatellite portion) thereby enhancing the resolution for population level studies. These characteristics make mtDNA more sensitive to genetic drift, population bottlenecks, and female mediated gene-flow compared to protein coding genes such as allozymes (Billington 2003).

I tested two sections of the mtDNA genome (ND5/6 and D-loop; Table 3; Lu et al. 2001; Billington and Hebert 1991) for PCR efficiency in a small subset of individuals (n = 8) from five populations throughout the lake. All PCR cocktails were 20 µL total volume with reagent concentrations as follows: 1.0X of Fisher brand10X buffer B (Thermo Fisher Scientific Inc., Waltham, MA.), 0.8mM of dNTPs, 1.75 mM of $MgCl_2$, 0.5µM of all primers, and 1 unit of Taq polymerase (New England Biolabs Inc., Ipswich, MA.). An identical thermocycler profile was used for both the ND 5/6 and D-loop. Thermocycler profile included an initial denaturation of 5 min at 94° C, followed by 35 cycles of denaturation, primer annealing, and elongation of 94° C for 45 s, 52° C for 45 s, and 72° C for 45 s, with a final elongation period of 40 min at 60° C. Resulting amplicons were visualized on a 2% TAE (Tris-Acetate-EDTA) agarose gel and visualized using UV-light to check for the presence of a single, sharp band (fragment). This acted as an initial screening process to determine the ability of the selected primers to amplify mtNDA from Lake Michigan lake whitefish.

Following confirmation of mtDNA amplification efficiency, I subsequently amplified 10 randomly chosen individuals from each of the nine 2005 sample year spawning aggregates. Based on the ¼ effective size of mtDNA, sample sizes of ~10 individuals per population have been successfully used in population differentiation

studies of lake whitefish (Bernatchez and Dodson 1991). Following amplification, the samples were purified of excess primers and dNTPs using the PerfectPrep® PCR Cleanup 96 kit (Eppendorf, New York, NY) with manufacturers suggested protocol. All samples were sequenced using the ABI BigDye® v.3.1 Terminator kit and manufacturers suggested protocol (Applied Biosystems Inc.). Sequencing was conducted on an ABI Prism® 377XL DNA sequencer (Applied Biosystems Inc.). The resulting sequences were verified by eye and aligned into a multiple sequence alignment using Geneious Pro v.2.5 (Drummond et al. 2006). Sequence data for the ND 5/6 region was translated using the vertebrate mitochondrial genetic code to ensure accuracy.

Analysis – Objective 1

Several genetic measures were used to determine the extent of lake whitefish microsatellite and mitochondrial DNA genetic diversity within and among putative lake whitefish spawning aggregates and to determine the utility of these markers to discriminate among spawning stocks of lake whitefish. These measures included allele frequencies, mean number of alleles, allelic richness, and heterozygosity for microsatellites, and number of haplotypes or haplotypic diversity, and the nucleotide diversity index for the mtDNA based on Ruzzante (1998) and Ryman et al. (2006).

Microsatellite genetic diversity measures.—Allele frequencies are the proportion of a particular allele present in a sampled population and represent the base unit of genetic diversity used in most subsequent analyses. Microsatellite allele frequencies were calculated from the observed genotype frequencies according to the following formula (Hedrick 2000):

$$p_i = \frac{2n_{ii} + \sum_{i=j} n_{ij}}{2n}, \sum_i p_i = 1,$$

where p_i is the frequency of allele i, n is the total number of individuals sampled, n_{ii} is the number of observed homozygotes, and n_{ij} is the number of observed heterozygotes with that specific allele (i).

Knowledge of the mean number of alleles within populations can provide a measure of genetic diversity to compare across populations and across loci. The mean number of alleles for each population was calculated by summing all alleles present and dividing the sum by the total number of loci analyzed. The formula for calculating the mean number of alleles for each population is as follows

$$\bar{a}_p = \frac{\sum a_1 + a_2 + \cdots + a_n}{L},$$

where $\bar{a}_p$ is the mean number of alleles for population p, a is one allele, and L is the total number of loci surveyed. Alternatively, knowing the locus to locus polymorphism provides a direct measure of marker usefulness for population genetic studies. The mean number of alleles per locus, $\bar{a}_l$, is an arithmetic mean of the number of alleles present in locus l at all sampled populations and is calculated as:

$$\bar{a}_l = \frac{\sum a_1 + a_2 + \cdots + a_n}{P_n},$$

where the numerator represents the sum of all alleles present across n populations for an individual locus (l) and the denominator is the number of populations sampled (P_n). Both $\bar{a}_p$ and $\bar{a}_l$ were calculated using Microsatellite Toolkit (Park 2001).

A problem with using simple allele counting and mean number of allele estimates

is the presence of unequal sample sizes which can result in different expected results. Therefore, I estimated allelic richness, the number of alleles in a sample, corrected for sample size using the rarefaction method of Goudet (1995) and Petit et al. (1998) as recommended by Leberg (2002). Allelic richness, $r(g)$ the number of different alleles in a sample of g genes was calculated using the following formula:

$$\hat{r}\left(g\right) = \sum_{i}\left[1 - \binom{N - N_i}{g} \Big/ \binom{N}{g}\right]$$

where N_i represents the number of occurrences of the i^{th} allele among the N sampled genes (Petit et al. 1998). Allelic richness estimates were conducted using HP-RARE v1.0 (Kalinowski 2005).

Heterozygosity is a widely used genetic diversity measure that estimates the fraction of individuals in a population that have two different alleles at a given locus. Unbiased heterozygosity estimates were calculated according to Nei (1987) and were averaged across all sampled loci to obtain a population-specific average expected heterozygosity value. Expected heterozygosity (H_e) for a multi-allelic system is calculated by the following formula:

$$H_e = 1 - \sum_{i=1}^{n} p_i^2$$

where p is the frequency of allele i over n alleles.

Hardy-Weinberg equilibrium and linkage disequilibrium.—The Hardy-Weinberg equilibrium (HWE) principle (Hardy 1908; Weinberg 1908) is a crucial law in population

genetics stating that in a large, randomly mating population with no selection, mutation, or migration, the allele frequencies and the genotype frequencies are constant from generation to generation (Guo and Thompson 1992). The HWE provides a baseline to determine if gene frequencies have changed in a population and if divergence or evolution has occurred.

Tests examining samples for conformance to HWE provide the first steps in assessing panmixia within populations, the genetic divergence between populations, and the utility of diploid markers for such studies (Shaklee and Currens 2003). The test examines the goodness of fit between observed genotypic counts versus the expected genotypic counts under the HWE expectations where the expected frequency of a homozygote genotype is p_i^2 where p_i is the frequency of allele i in a given population and the frequency of a heterozygous genotype is $2p_ip_j$ where p_i and p_j are allele frequencies for allele i and j, respectively, where $i \neq j$.

Tests for HWE of all loci within populations using a Fisher's exact test Markov Chain Monte Carlo (MCMC) method with 1000 batches and 1000 iterations each (Guo and Thompson 1992; Raymond and Rousset 1995) were conducted with the null hypothesis that expected numbers of each genotype are equal to the observed genotypes in our samples. Exact tests compare heterogeneity within and between samples and were also performed between sample years for the same locations to determine the combinability of samples. To minimize type I errors with multiple pairwise tests, levels of significance were adjusted using the sequential Bonferroni method (Rice 1989).

A common problem with highly polymorphic loci (i.e., microsatellites) that is not sufficiently resolved through the use of exact tests is the recovery of small expected

frequencies for rare genotypes (Pamilo and Varvio-Aho 1984). High numbers of rare alleles (and their corresponding genotypes) can lead to significant deviations from HWE based on traditional methods such as exact tests due to a cumulative effect of rare but non-zero expected genotypes. To alleviate this problem, all exact tests resulting in significant deviations from HWE expectations were re-tested following a modification of the suggestions of Hedrick (2000). All genotypes with an expected frequency of <1 were pooled into one observed and one expected frequency value. The pooling of genotypes with an expected frequency of <1 was conservative as Hedrick recommended the pooling of genotypes with an expected frequency of less than five (Crisp et al. 1978; Hedrick 2000). The new observed and expected genotype values were then tested using a chi-square goodness of fit test in Minitab v.14.20 (Minitab Inc., State College, PA). Significance was determined after the application of a sequential Bonferroni adjustment (Rice 1989).

Any locus or population that did not conform to HWE expectations was specifically tested for heterozygote excess/deficiency. Departures from HWE expectations due to heterozygote deficiency can indicate the sampling of multiple gene pools (i.e., multiple populations) due to the Wahlund effect (Wahlund 1928) suggesting multiple gene pools were present in the sample. This was determined using a U-test (score test; Raymond and Rousset 1995) where the rejection zone was focused strictly on if the direction of HWE (heterozygote excess or deficiency). All tests of heterozygote excess and deficiency were conducted in GENEPOP v3.4 (Raymond and Rousset 1995).

An important assumption of nearly all population differentiation measures is independence of sampled loci. Linkage disequilibrium is the condition in which the

gamete frequencies in a population deviate from the values they would have if the genes at each locus were combined randomly (i.e., independent assortment). When such deviation is not evident (i.e., linkage disequilibrium = 0), the population is said to be in linkage equilibrium and all sampled loci are effectively segregating independently. The test for linkage disequilibrium estimates the deviation from independent assortment of alleles at two loci versus that expected by random chance alone (Excoffier et al. 2005). Linkage disequilibrium between all pairs of loci was assessed using a likelihood ratio test of the data assuming linkage equilibrium versus the likelihood of the data not assuming linkage equilibrium with an empirical expected distribution estimated via permutation (Excoffier et al. 2005). The test was implemented in ARLEQUIN v.3.0 (Excoffier et al. 2005) using 10,000 permutations.

Mitochondrial DNA genetic diversity measures.—Several measures of polymorphism were assessed including the number of haplotypes, haplotypic diversity (h), and nucleotide diversity (π). Fish that have the same mtDNA sequence are classified as having the same mtDNA haplotypes and the number of haplotypes is simply a count of observed haplotypes. Once the different haplotypes in a sample were identified, their relative frequencies in each population were calculated. Nucleon diversity is the probability that two individuals will differ in mtDNA haplotype and nucleotide diversity is the average number of pairwise nucleotide changes per site.

Haplotypic diversity is the number of different haplotypes in a sampled population (Billington 2003). Once the different haplotypes in a sample were identified, their relative frequencies were determined. The haplotypic diversity index (h) was estimated by the following formula (Nei and Tajima 1981).

$$\hat{h} = \frac{n\left(1 - \sum_{i=1}^{l} x_i^2\right)}{n-1},$$

where x_i is the frequency of the i^{th} type of mtDNA in a sample of n individuals, and l is the number of haplotypes observed in the sample. The resulting value, $\hat{h}$, estimates the probability that two randomly chosen individuals from the population have different haplotypes (Billington 2003).

After mtDNA haplotypic diversity was estimated, the amount of genetic divergence between haplotypes was measured using a nucleotide diversity estimate ($\hat{\pi}$). Nucleotide diversity considers both the frequency of occurrence of haplotypes in a population and the divergence among them (Billington 2003). An estimate of nucleotide diversity (π) was calculated using the following formula

$$\hat{\pi} = \frac{n}{n-1} \sum_{ij} \hat{x}_i \hat{x}_j \pi_{ij},$$

where x_i and x_j are the frequencies of the ith and j^{th} type of mtDNA in a sample and π_{ij} is the fraction of mismatched sites between two haplotypes (Nei and Tajima 1981).

Marker utility.—To assess the utility of the genetic markers to discriminate among spawning populations of Lake Michigan lake whitefish, I evaluated the genetic diversity measures discussed previously and compared the values to previously identified levels necessary for spatial genetic structure studies (Douglas et al. 1999; Turgeon et al. 1999; Lu et al. 2001; Douglas and Brunner 2002; Stott et al. 2004). A consensus of those studies indicated an observed level of heterozygosity of ~ 0.60 and a mean number of alleles/locus/population of ~7.0 was sufficient polymorphism for population differentiation if genetic differences existed. Loci with ≥2 and ≤32 alleles were

considered to have sufficient polymorphism to determine genetic differences among and between populations based on my sample sizes according to the simulation models of Bernatchez and Duschene (2000). Finally, all loci had to conform to Hardy-Weinberg equilibrium expectations (Hardy 1908; Weinberg 1908) across all populations and show no significant linkage disequilibrium with other loci.

The usefulness of the mtDNA data was determined by examining the haplotypic diversity, the frequency of different haplotypes, and the nucleotide diversity. For the data to be useful, it had to show ≥ two haplotypes with individual haplotypes occurring in more than a single representative. Furthermore, at least some haplotypes had to occur in multiple populations to allow analysis based on frequencies. Essentially, the mtDNA needed to have multiple haplotypes with some degree of shared haplotype variation among populations. Mitochondrial DNA statistics were assessed using ARLEQUIN v3.0 (Excoffier et al. 2005).

Analysis – Objective 2

The second objective of the project was to describe the genetic population structure of spawning lake whitefish aggregates in northern Lake Michigan and Green Bay in terms of genetic stock identification and degree of stock isolation. The analysis for this objective was accomplished through a process known as Genetic Stock Identification (GSI) that involves a series of nested statistical tests, where the findings from one test allow us to *a priori* group populations for subsequent tests. The first step in GSI is to test for conformance to HWE expectations within populations because differences in gene pools can be tested based on deviations from HWE if two separate

gene pools are tested as a composite population (i.e., Wahlund effect; Wahlund 1928). Tests for conformance to HWE were used to determine the combinability of samples collected in consecutive years from the same spawning aggregates. Individual sample year tests were conducted and then the two samples were combined to form a composite sample. Tests for HWE were conducted on this composite population. If there were no changes from the individual year tests, the two samples were considered to be representative samples from a homogeneous population and thus, were combined. An identical approach was taken with temporal samples within the same year (i.e., early versus late run samples) to assess for genetic differences associated with run-timing.

Following HWE testing, the second phase of GSI used a combination of tests of population differentiation, genetic distance measures, analysis of molecular variance (AMOVA), and estimates of F_{st}. This series of hierarchical tests to determine population structure is well published and widely accepted as a standard means of genetic stock identification (Grant et al. 1980; Waples et al. 1990; Shaklee and Currens 2003; Fraser et al. 2004, Beacham et al. 2005; Stepien et al. 2006). For all statistical testing, the assumptions of the infinite-allele model (IAM; Kimura and Crow 1964) were used unless otherwise noted.

Population differentiation based on genic differentiation.—Tests for genic differentiation compare allele frequency distributions between populations with the null hypothesis that allele frequencies are the same in each population. This test was based on the null hypothesis that all lake whitefish in Lake Michigan belonged to one large, panmictic unit. If genetic differences exist, the null hypothesis was rejected. Populations were tested using Fisher's exact test for allele frequency differences among populations

(genic differentiation option in GENEPOP v.3.4; Raymond and Rousset 1995). Significance was determined using a Markov chain of 100 batches with 1000 iterations each (Guo and Thompson 1992). First, a locus by locus analysis, which assumes statistical independence across loci, was performed. Then, a global test, which is the combination of tests across loci, was performed for each sample pair. Alpha was set at 0.05 for both tests and any p-value less than 0.05 was deemed significant.

Genetic distance.—When Lake Michigan's lake whitefish were shown to not be one panmictic unit, I performed cluster analysis on the groupings of populations based on their genetic distance. Genetic distance measures the amount of divergence or the extent of genomic differences that have occurred since two populations shared a common ancestor. Ideally, genetic distance values range linearly from 0 (when all alleles are shared between populations) to 1 (when no alleles are shared between populations) (Lowe et al. 2004). I used the chord distance of Cavalli-Sforza and Edwards (D_c; 1967) to estimate the genetic distance between all pairs of populations. This genetic distance measure has been shown to be the most efficient in obtaining correct tree topology under many different conditions (Takezaki and Nei 1996) including recent evolutionary divergence consistent with expectations of lake whitefish found in the Great Lakes (Bernatchez and Wilson 1998). The chord distance has also been recommended for use with microsatellite data (Takezaki and Nei 1996) and is widely published in recent literature on coregonids (Patton et al. 1997; Lu et al. 2001; Douglas and Brunner 2002; Stott et al. 2004).

The D_c uses allele frequencies to determine the distance between populations relative to their theoretical position on the surface of a hypersphere (Nei and Kumar

2000). The distance and surface of the hypershere are defined by the allele frequencies resulting in the genetic distance between two populations relying on the probability of randomly sampling the same allele in both populations. The D_c is estimated by

$$D_c = \frac{2\sqrt{2}}{\pi}\sqrt{1-\cos\theta}\ ,$$

where

$$\cos\theta = \sum_{i=1}^{m}\sqrt{P_i P_i'}$$

where P_i is the frequency of the *i*th allele in the first population and P_i' is the frequency of the same allele in the second population.

Relevant groupings of populations for subsequent analyses were determined by constructing an neighbor-joining tree (NJ tree; Saitou and Nei 1987) from the D_c pairwise matrix. Genetic distance measures were calculated for all population pairs using PowerMarker v3.25 (Liu and Muse 2005) and an unrooted neighbor-joining tree was constructed in the PHYLIP software package (Felsenstein 1993). Five-thousand bootstrap pseudoreplicates were conducted to measure confidence in nodal resolution with a majority rule consensus tree constructed using CONSENSE (Felsenstein 1993). A node was considered confidently resolved if bootstrap support was ≥80%. All trees were viewed using the program TREEVIEW (Page 1996).

Analysis of molecular variance.—To determine if the grouping of populations (putative stocks) based on node support from the NJ tree were biologically relevant, a hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992) was performed to compare the molecular variance within and between the putative stocks or populations to that which would be expected in one panmictic population. AMOVA is based on the calculation of a genetic distance matrix to determine the molecular variance

apportioned within and between predefined groups. Due to the absence of assumptions required, the test is both widely applicable and powerful in terms of assessing genetic structure (Lowe et al. 2004). The test is based on the premise that biologically cohesive groups of populations (e.g., stocks) should show minimal genetic variance when grouped together and significant between group variance when compared to other groups of populations.

Significance levels for AMOVA are computed by non-parametric permutation of the data set to develop a null distribution with which to test the recovered fixation index (Excoffier et al. 1992). The fixation indices in this case are Φ-statistics, which are correlation statistics directly analogous to *F*-statistics and derived from the variance components computed during AMOVA (Lowe et al. 2004). They express the correlation of a pair of individuals drawn at random from a particular subgroup of the dataset, relative to that of a pair of individuals drawn from a wider grouping, indicating the relative partitioning of diversity between the hierarchical levels being analyzed. A significant proportion of variance among groups but not within groups suggests biological justification for the group (i.e., potential stock). All AMOVA tests were performed using ARLEQUIN v3.0 (Excoffier et al. 2005) with 1,000 permutations.

Estimates of F_{st}.—The final statistical test I utilized for GSI was Wright's (1931) fixation index (F_{st}), a traditional measure of population subdivision, which measures the reduction of heterozygosity within a subdivided population versus the expected amount of heterozygosity if all subpopulations were acting as one large panmictic population (Hartl and Clark 1997). For multiple, high variation loci (such as microsatellites), an F_{st} estimator, θ, (Weir and Cockerham 1984) is preferred over the traditional F_{st}. Theta is

calculated based on the variance in allele frequencies under an infinite alleles model (IAM). I estimated θ for all population pairwise comparisons in ARLEQUIN v.3.0 (Excoffier et al. 2005) which uses the following formula:

$$\theta = \sigma_a^2 \Big/ \sigma^2,$$

where σ^2_a is the between population variance in allele frequencies and σ^2 is calculated as

$$\sigma^2 = \sigma_a^2 + \sigma_b^2 + \sigma_w^2,$$

where σ^2_b is the allele frequency variance among individuals in the population and σ^2_w is the allele frequency variation within an individual (Lowe et al. 2004). Using the estimated values of θ a population pairwise distance matrix is then created. If the comparison of the θ value between populations is significantly different from 0, significant gene flow between populations is inferred. Significance is calculated based on 1,000 permutations of the data.

Significant values for population pairwise comparisons of θ were evaluated within groups from the final AMOVA. Only differences between populations within groups were viewed to minimize the number of pairwise comparisons and allow us to further evaluate fine-scale genetic stock structure.

Additional genetic analyses.—The delineation of genetic stocks is only part of conserving and managing genetic diversity. It is also important to define and understand the interrelationships of stocks, quantify how genetic diversity is distributed among stocks within a species, and how this information can be useful for more effective management of lake whitefish. Three subsequent analyses were conducted to help resolve some of these issues. The effective population size (N_e) was estimated to help explain the amount and dynamics of genetic diversity within populations, a test for

isolation by distance (IBD) was conducted in an attempt to evaluate gene flow and migrational patterns, and a genetic mixture analysis was conducted to determine biological relevance of our population groupings (genetic stocks) and the degree of differentiation between genetic stocks.

Effective population size.—A critical parameter for genetic management of exploited populations is the effective population size (N_e). Effective population size can be defined as the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration (Wright 1931). A rough approximation of N_e is the number of successfully breeding individuals within a generation for a population. The effective population size is a primary measure determining how a population changes over time (Frankham 1996; Hallerman 2003). In a population with a large N_e, natural selection primarily accounts for the genetic dynamics and divergence of populations. Alternatively, when N_e is small, genetic drift will primarily account for the genetic dynamics and divergence of populations. Genetic drift is the random process of change in allele frequency as a result of random sampling of gametes (Lowe et al. 2004). The smaller the population size the more likely chance events are to change the allele frequency with the end result of genetic drift being a loss of genetic diversity. Effective population size estimates were determined from the genetic data using the point estimation technique based on linkage disequilibrium (N_eD; Hill 1981, Bartley et al. 1992) for all spawning aggregates using Ne Estimator (Peel et al. 2004).

As discussed previously, LD is the non-random association between alleles at

different loci in gametes (Wang 2005) and is measured by comparing the difference between the expected co-occurrence of two alleles at two loci given random mating, independent segregation, and no selection, and their actual instances of co-occurrence (Bartley et al. 1992). Effective population size based on linkage disequilibrium was calculated from the formula:

$$N_e(D) = \frac{1}{\left[3\left(r^2 - 1/S\right)\right]}$$

where r is the correlation among alleles and S is the sample size. The value for r can be calculated from the relation:

$$r = \frac{D}{\left[p(1-p) + q(1-q)\right]^{1/2}}$$

where p and q are frequencies of allele A at locus 1 and allele B at locus 2, respectively and D is Burrow's composite measure of disequilibrium (D* in Campton 1987).

Isolation by distance.—To determine if migration between gene pools was sufficient to maintain genetic diversity between populations but insufficient to arrest genetic divergence, I conducted a test of isolation by distance (IBD). Tests of IBD were performed using the program IBD (Bohonak 2002) to determine if genetic distances were correlated with geographic distances indicating potential barriers to gene-flow. Genetic distances D_c (Cavalli-Sforza and Edwards 1967) between population pairs were calculated as previously described. Geographic distances were calculated in ArcMap v.9.2 (Environmental Systems Research Institute, Inc., Redlands, CA) and were based on the shortest distance (km) across water from the center of the fishing grid (Figure 4) where population was sampled to the center of the fishing grid where the other population

was sampled.

Significance in the IBD test was determined using a Mantel test (Mantel 1967). Reduced Major Axis (RMA) regression was then applied to estimate the slope and intercept of the IBD relationship. RMA is recommended over ordinary least squares (OLS) methods for isolation by distance tests because OLS assumes that the independent variable is measured without any error (Hellberg 1994). Although this holds true for simulated data, it does not for actual data, where geographic distance approximates the number of steps between populations. Using OLS regression whenever error exists in the independent variable will underestimate the slope of the regression. In such instances, RMA regression should provide a better estimator of the relationship between the variables (Hellberg 1994).

Genetic mixture analysis.—An important reason why I am trying to delineate lake whitefish stocks was a concern over management of shared stocks between different management agencies (i.e., a mixed stock fishery). To determine the usefulness of our groupings of populations in terms of this issue, I examined the ability of the data to identify source locations (i.e., assignment testing) and relative contributions (i.e., mixed stock analysis) of sampled fish to their putative stocks of origin identified during the GSI portion of the study and their known sampling locality using Genetic Mixture Analysis (GMA; Kalinowski 2003).

First, I used GMA to assign individuals to their most likely population of origin. Assignment testing, which broadly speaking, determines how indicative an individual's genotype is of the population in which it was sampled, has become a powerful tool to evaluate the degree of genetic differentiation among populations (Dominguez et al. 2001).

I estimated assignment to the top three most probable populations based on the probability of the selected genotype coming from a random drawing of genotypes from each population using the methods of Rannala and Mountain (1997) as implemented in GMA (Kalinowski 2003). The probability of an individual coming from the *i*th baseline population is estimated with Bayes' rule using that population's estimated contribution to the mix as *a priori*. An assignment to the population of origin with a probability of ≥80% was considered a confident self-assignment.

Next, I used GMA to assign individuals to one of six delineated genetic stocks using the method of Rannala and Mountain (1997) for estimating the probability of observing a genotype in a population for both estimation and assignment, and Bayes' rule for estimating population contribution as described above. The same assumption was made for this assignment test as the population assignment test (i.e., high degree of self assignment indicates biological relevance). To assign fish to stocks, the mixture proportions or assignment probabilities for each population in a reporting unit (genetic stock) were summed to produce an estimate for the reporting unit as whole. The resulting estimate shows what proportion of the individuals in a sample came from each stock. If a high degree of self-assignment using the population or stock method is not observed this would indicate weak population structuring (i.e., low genetic divergence) among the Lake Michigan lake whitefish populations.

The final way I tested my data using a mixed-stock approach was to simulate mixed harvests or single stock harvests and assess the actual proportions of the simulated catch versus the predicted proportions of the simulated catch based on assignment from GMA (Kalinowski 2003). This was accomplished by taking various subsets of our data

(e.g., 10 fish from 10 sample locations), and treating these as a hypothetical commercial catch, and then using GMA to predict what proportion of the catch was attributable to each population/putative stock. During simulations I pre-defined the proportions of each stock present in a commercial catch and then used GMA to calculate stock proportions based on the calculated allele frequency distributions from the baseline data. All simulations were run using different proportions for each stock and population present in the commercial catch and were tested with the assumptions that if the genetic data collected strongly differentiated the populations or stocks then GMA would assign the simulated catch back to the correct population or stock of origin with a high probability. All six simulations were conducted by simulation of genotypes for 500 individuals that were based on allele frequencies present in my baseline data (actual data). The 500 simulated genotypes represented a single commercial catch. Then, using allele frequency distributions the probability of sampling a simulated multi-locus genotype from each baseline population was calculated using Bayes' rule. Because not every sampled population had ≥100 samples 100 multi-locus genotypes were simulated using baseline allele frequency data as a new baseline with which to compare the simulated multi-locus genotypes. To standardize sample size and simulate the fact that all fish in Lake Michigan could hypothetically be captured by any commercial fisherman at any location in the lake at any time, all simulations were run with an equal probability distribution of all 12 sampled populations and an equal proportion of fish (n =100) were made available from all populations. During simulations, the method of Rannala and Mountain (1997) was used to simulate the drawing of genotypes. When mixtures are being simulated, sampling of individuals is assumed to be independent and when simulated mixtures are

analyzed, the baseline is always re-shuffled.

To determine if any under or over-estimation of stock proportions was occurring in the simulated data versus the actual pre-defined proportions given during simulations this data was compared using residual plots. A negative residual represented an underestimation by GMA of the proportion of the stock present in the sample, whereas a positive residual represented an overestimation of the stocks contribution to the commercial catch.

RESULTS

A total of 1,624 fish were collected from 12 spawning aggregates (Figure 5). A total of 280 and 286 samples were collected from the primary study sites BBN and NMB, respectively. Samples were collected from all but two lake whitefish commercial/ statistical management zones (WFM-06, and WFM-09) known to contain spawning aggregates. Sample sizes for all but one of the spawning aggregates (MR) were ≥70. Eight of the twelve spawning sites were sampled for two years (2005 and 2006) resulting in ≥ 124 samples for these eight sites (Table 5). Samples from Elk Rapids were only collected in 2005, and samples from Traverse Bay, Hog Island, and the Lower Menominee River were only collected in 2006. All samples were analyzed at 13 microsatellite loci and 90 individuals (10 from all 9 populations sampled in 2005) were analyzed at two mitochondrial DNA genes. However, the accurate scoring of data from the locus C72 was suspect due to issues of potential polyploidy at this locus, so only the results from 12 loci will be presented.

Usefulness of Genetic Markers

A prerequisite to conducting genetic differentiation tests among lake whitefish spawning aggregates was the recovery of sufficient genetic diversity in the molecular markers I chose to use (Objective 1). All tests and estimates conducted on the microsatellite data showed sufficient genetic diversity and distribution to provide the discriminatory ability to delineate genetic structure among the sampled populations if significant structure exists.

Microsatellite genetic diversity measures.—An initial qualitative assessment of

the microsatellite genetic diversity showed markedly higher levels of polymorphism (Table 5; Appendix 1) compared to previous allozyme data from studies of Imhoff (1977), Leary (1979), and Imhoff et al. (1980). The observed numbers of alleles at a given locus ranged from a low of 2 (Menominee River; C18) to a high of 26 (North-Moonlight Bay; C52) while the mean number of alleles/locus ranged from 6.92 (Menominee River) to 11.92 (NMB). Allele frequencies represent the base unit of genetic diversity for nearly all tests of genetic differentiation. The allele frequencies for each locus varied considerably across populations (Appendix 1). For example, the distribution of allele frequencies at C18 (7 total alleles; Figure 6) showed a common allele (156) that varied by nearly 20% between the Muskegon sample (39.5%) and the Hog Island sample (57.9%). The distribution of allele frequencies at a more polymorphic loci C23 (Figure 7) showed an even greater amount of among population allele frequency variation. The final observation based genetic diversity estimate I considered was the observed heterozygosity (Table 5). Observed heterozygosity for each population ranged from 0.5869 (Elk Rapids) to 0.6549 (Traverse Bay) with a mean heterozygosity of 0.624.

The reliance on raw observations is complicated by issues of differing sample sizes in my data (Table 5). Sample sizes ranged from a low of 35 (Menominee River) to a high of 286 (NMB samples from 2005 and 2006 combined). To account for sample size effects on the measures of allelic diversity, I measured allelic richness and the number of private alleles using rarefaction. However, due to the large discrepancy in the size of the Menominee River sample (n = 35) and the next smallest sample size (n = 70; Hog Island), I chose to conduct the rarefaction method using the Hog Island sample as the minimum sample size and not unduly restrict the data with the small Menominee

River sample (Table 6).

Allelic richness estimates showed high levels of polymorphism consistent with that needed to assess genetic structure (Fraser et al. 2004; Nielsen et al. 2004). Allelic richness estimates showed a wide range of polymorphism with the largest discrepancy in an allelic richness estimates for each population being 8.10 between Traverse Bay (9.16) and Epoufette (17.26) at locus C49. Allelic richness estimates differed by 22-63% for populations at individual loci, with the largest discrepancy in percentage being at locus C18 between Epoufette (2.00) and Ludington (5.52).

Private alleles, alleles found in only one population, are good indicators of a lack of gene flow between populations because private alleles can only accumulate when gene flow is low (Lowe 2004). The estimates of private allelic richness based on rarefaction varied tremendously among loci with C49 exhibiting 8.69 rarefacted private alleles and C4 showing less than one total rarefacted private allele across all populations (Table 7). The overall population totals ranged from 0.000 to 4.955 at individual loci (Table 7), and from 0.326 (Saugatuck) to 7.955 (Elk Rapids) across loci (Table 7) suggesting restricted gene flow of Elk Rapids.

Nei's unbiased gene diversity (Nei 1987) is one of the most widely used measures of genetic variation. Unbiased heterozygosity values for the microsatellite loci ranged from 0.6357 (Menominee River) to 0.6637 (Ludington; Table 5). The mean unbiased heterozygosity for all populations was 0.6498. These values were above the value of 0.60 discussed in the methods as a necessary threshold for population differentiation studies (Douglas et al. 1999; Turgeon et al. 1999; Lu et al. 2001; Douglas and Brunner 2002; Stott et al. 2004). Observed heterozygosity is the mean percentage of individuals

heterozygous per locus (Avise 1994). Observed heterozygosity (microsatellites) for each population ranged from 0.5869 (Elk Rapids) to 0.6549 (Traverse Bay) (Table 5). Observed heterozygosity averaged 0.6240 across all populations. Furthermore, only one observed heterozygosity (0.5869; Elk Rapids) fell below the 0.60 threshold.

Hardy-Weinberg equilibrium.—The NMB and BBN populations did not deviate significantly from HWE when the early versus late run samples were combined indicating homogeneity among samples. Therefore, there was no evidence of genetic differences consistent with run time in these populations. Similarly, sample years (from a single site) did not differ from HWE expectations when combined suggesting the two samples represent independent samples from a single gene pool; therefore, only the combined year data will be reported. Initially, the combined year data showed 71 of 144 exact tests significantly deviated from HWE at a nominal α of 0.05. After pooling rare genotypes, 10 of 144 total comparisons were significant following sequential Bonferroni adjustment (Rice 1989). This number was only slightly higher than that expected by chance (6.9% vs. 5%). Six of twelve loci had at least one significant departure from HWE, no locus was significantly out of HWE for more than two populations, and only one population, Big Bay de Noc, had three departures from HWE (B1, C6, and C 4; the same three loci were out of HWE in the individual year tests). Because of the lack of discernible pattern to the significant locus/population deviations from HWE and because these findings were consistent with other studies involving population differentiation of salmonids (Brunner et al. 1998; Lu et al. 2001; Douglas and Brunner 2002; Stott et al. 2004), all sampled populations and loci were considered in HWE.

To determine if individual locus HWE deviations within populations were

consistent with multiple gene pools being sampled (i.e., heterozygote deficit; Wahlund 1928), all ten sequential Bonferroni corrected significant loci were assessed for heterozygote excess and deficiency. All deviations were consistent with heterozygote excess suggesting no inadvertent sampling of multiple gene pools and/or issues with systemic errors (i.e., null alleles or allele dropout; Navidi et al. 1992; Miller et al. 2002).

To ensure that all loci were independently segregating, linkage disequilibrium was then tested for. Linkage disequilibrium tests between all pairs of loci showed no significant deviations and all loci were considered to be independently segregating.

Mitochondrial DNA genetic diversity measures.—The mtDNA data was sufficient to assess the utility of the mitochondrial genome for discerning population genetic structure among the nine considered spawning aggregates (Objective 1). The multiple sequence alignment of the ND 5/6 amplicon was 574 base pairs in length and actually only encompassed sequence in the ND 5 gene region. A total of 80 usable sequences were obtained (others were eliminated due to poor sequence quality). There was a total of 6 haplotypes at the ND 5 region with the most common haplotype present in 90.1% of the individuals sampled (Table 8). Of the five other haplotypes, three were singletons, (only found in a single individual) and the two were only shared by two populations and no more than two haplotypes were shared between populations including the predominant haplotype (Table 9a). Haplotypic diversity, the probability that two randomly chosen individuals from the population have different haplotypes, was estimated to be 0.188 for the ND 5 region.

The multiple sequence alignment of the D-loop region was 538 base pairs in length. A total of 85 usable sequences were obtained with a total of 7 haplotypes present.

The most common haplotype was found in 85.9% of the individuals sampled (Table 8). Of the six other haplotypes, three were singletons and no more than two haplotypes were shared between populations including the predominant haplotype (Table 9b). Haplotypic diversity for the D-loop section was estimated to be 0.260.

Nucleotide diversity, a weighted diversity measure considering the distribution of variants, considers both the frequency of occurrence of haplotypes in a population and the divergence among them (Billington 2003). Nucleotide diversity for the ND 5 section was estimated to be 0.00016 and nucleotide diversity for the D-loop section was estimated to be 0.00025.

Genetic Stock Identification

Population differentiation based on genic differentiation.—Tests of genic differentiation showed the lake whitefish spawning aggregates included in this study do not represent one panmictic unit. Significant differences in allele frequency distributions were observed among all populations at 11 of 12 loci ($p < 0.001$). The only locus that was not statistically significant was C4 ($p = 0.1974$). The overall (global) test had a chi-square value of infinity (indicating the number exceeded the maximum capacity of the GENEPOP program to calculate), with 24 degrees of freedom and a significant probability value of $p \leq 0.00001$. The overall significance of this test indicates a lack of panmixia within Lake Michigan.

Genic differentiation tests for each population pair were conducted overall across all loci (Table 10). Sixty-six pairwise comparisons were made between pairs of populations. A total of 8 tests were non-significant ($\alpha = 0.05$) following sequential Bonferroni correction (66 comparisons). Non-significant values were observed between

population pairs Epoufette and Naubinway (p = 0.377), Epoufette and Menominee River (p = 0.528), Naubinway and Menominee River (p= 0.0529), Saugatuck and Ludington (p = 0.0648), Muskegon and Ludington (p = 0.157), Menominee River and Cedar River (p = 0.380), Menominee River and North-Moonlight Bay (p = 0.0863), and Menominee River and Big Bay de Noc (p = 0.0748). Non-significant values indicated similar allele frequency distributions and allowed me to *a priori* group populations for subsequent GSI analyses.

Genetic distance.—Pairwise genetic distances and an unrooted NJ tree were used to cluster and visualize putative population groupings (Figure 8). Pairwise D_c (Cavalli-Sforza and Edwards 1967) ranged from 0.0687 (Cedar River and North-Moonlight Bay) to 0.2546 (Elk Rapids and Menominee River) (Table 11). The original NJ tree indicated 6 groupings of populations. Group A (BBN, NMB, Cedar River), Group B (Naubinway, Epoufette), Group C (Ludington, Muskegon, Saugatuck), Group D (Elk Rapids), Group E (Traverse Bay, Hog Island), and Group F (Menominee River). All population groupings corresponded with geographic location except the Menominee River population. The Menominee River sample was not significantly different than 5 of the other 11 populations based on genic differentiation (allele frequencies) and grouped more closely with populations from the Northeast portion of Lake Michigan suggesting a potential mixed stock. The origins of the Menominee River population are also suspect due to a lack of spawning fish following the population crash of the mid 1900s and its recent recovery (P.J. Peeters; personal communication). Because Menominee River is a suspected mixed stock and admixture populations can dramatically influence the topology of a neighbor-joining tree because of the nature in which NJ trees are

constructed (see Saitou and Nei 1987; Nei and Kumar 2000) another NJ tree was created without the Menominee River sample (Figure 9). This NJ tree indicated between 4-6 groupings of populations. If populations were put into four groups the groupings would be Group A (BBN, NMB, Cedar River), Group B (Traverse Bay, Hog Island), Group C (Naubinway, Epoufette), and Group D (Ludington, Muskegon, Saugatuck, Elk Rapids). If populations are placed into five groups the groupings would be Group A (BBN, NMB, Cedar River), Group B (Naubinway, Epoufette), Group C (Ludington, Muskegon, Saugatuck), Group D (Elk Rapids), and Group E (Traverse Bay, Hog Island). The six population groups would be Group A (BBN), Group B (NMB, Cedar River), Group C (Naubinway, Epoufette), Group D (Ludington, Muskegon, Saugatuck), Group E (Elk Rapids), and Group F (Traverse Bay, Hog Island).

Analysis of molecular variance.—The genetic and biological relevance of the various groupings of populations from the genetic distance analysis was tested using AMOVA. I ran sequential AMOVA tests using the smallest number of supported clusters (four from the NJ tree; Figure 9) and continued to divide groups until a significant proportion of genetic variance existed among but not within the groups, suggesting biological justification for the groupings (i.e., potential stocks). Because of potential confounding effects of admixtures (discussed previously) the Menominee River sample was excluded from this portion of the analysis. In total three AMOVA analyses were conducted with the number of groups ranging from four to six (Table 12). In all tests, the majority of genetic variance was attributed to the differences among individuals within populations (97.86% - 98.05%). However, for the four-group (Table 12a) and the five-group (Table 12b), AMOVAs significant variance was attributed to among

population differences within groups suggesting the groupings were in fact heterogeneous and not of a stable, genetic group. The six-group AMOVA showed significant variance attributed to between group differences but showed no signal of heterogeneity within groups (among populations within group $p = 0.07$; Table 12c). This resulted in determination of six putative stocks of lake whitefish in Lake Michigan. The six putative genetic stocks identified by GSI are the North-Moonlight Bay stock (NMB) which includes the NMB, Cedar River, and Menominee River populations, the Big Bay de Noc Stock (BBN) which is comprised solely by the Big Bay de Noc population, the Northern stock (NOR) which includes the Naubinway and Epoufette populations, the Northeast stock (NOE) which includes the Traverse Bay and Hog Island populations, the Elk Rapids stocks (EKR) which is comprised solely of the Elk Rapids population and the Southeastern stock (SOE) which is comprised of the Ludington, Muskegon, and Saugatuck populations.

Estimates of F_{st}.—Pairwise F_{st} comparisons between population pairs were performed to determine the stability of the putative stocks from the AMOVA analysis. F_{st} values ranged from 0.038 to -0.005 with twelve pairwise comparisons being non-significant ($p \geq 0.05$) indicating significant gene flow between population pairs and 54 comparisons being significant indicating a lack of gene flow between population pairs (Table 13). Of the recovered multiple spawning groups (genetic stocks) identified from AMOVA, only the grouping of Hog Island and Traverse Bay showed significant differences.

Effective population sizes.—Effective population sizes and upper and lower 95% confidence intervals for all populations showed a wide range of values. Effective

population sizes ranged from 96 in Elk Rapids to 1321 in Big Bay de Noc (Table 14). The mean value across all populations was 503. It is important to note, however, that there are relatively wide confidence intervals associated with many of the estimates.

Isolation by distance.—Isolation by distance was tested to see if genetic distances were correlated with geographic distances for all population pairs. No significant barriers to migration were evident as three of the four IBD test variations were non-significant (Table 15; Figure 10). When all populations were included in the comparison only the log genetic versus log geographic comparison test resulted in a significant p-value ($p = 0.0470$; Table 15; Figure 10b). The potential admixture origin of the Menominee River sample could confound a test of IBD. Therefore, I conducted a test of IBD to determine if this potential admixture population was having any dramatic influence on the test (Table 15; Figure 11). When the Menominee River sample was excluded no test resulted in a significant p-value (Table 15). A similar approach was taken with the Elk Rapids population because of its large relative genetic distance versus all other populations ($\overline{x} = 0.2175$; Table 11). When the Elk Rapids population was excluded from the analysis, all 4 variations of the test were significant for isolation by distance (Table 15; Figure 12), suggesting Elk Rapids high divergence may be masking IBD effects between the other Lake Michigan populations.

Genetic mixture analysis.—Genetic mixture analysis allows managers to estimate the proportion of fish belonging to each population or stock present in a single commercial harvest. Overall, GMA showed the populations were distinct with >99% of fish assigning back to their sampled population and to their stock of origin (Table 16) with a vast majority >97% correctly assigned back to their sampled population and stock

of origin with at least 80% confidence (Table 16). The proportion of samples that assigned to their putative stock of origin was high (100%-93.6%; Table 17) further supporting the biological relevance of the stocks and illustrating the usefulness of the genetic differences between them.

Simulations of commercial harvests resulted in a high level of accuracy with ≥91.14% of the simulated individuals correctly assigned to their population of origin. The largest error in assignment of individuals to population of origin was in Simulation 1 (20% CR, 25% BBN, 5% MR, and 50% NMB) where there was a 5.89% underestimation of fish assigned to CR and a slight (~5%) overestimation of fish assigned to NMB (Table 18a). The majority of comparisons (72.2%) yielded an error rate of <1.00% (Table 18).

Nearly identical results were observed when simulated commercial harvests were assigned. Simulated commercial harvests assigned correctly to their stock of origin with at least 91.14% accuracy. The largest negative residual difference was in Simulation 3 with 6.23% too few fish assigned to the SOE stock (Figure 13). The largest positive residual was in Simulation 6 with 4.22% too many fish assigned to the BBN stock (Figure 13). Twenty-three out of 36 comparisons yielded an error rate of < 1.00% (Table 18; Figure 13) and the average error rate for the five simulations (30 comparisons) was 0.01% across all 36 comparisons.

DISCUSSION

Usefulness of Genetic Markers

Genetic diversity levels (allelic richness, private alleles, unbiased gene diversity and observed heterozygosity) at the microsatellite loci involved in this study were near that or exceeded the diversity levels of markers successfully used in other studies of salmonids (Douglas et al. 1999; Turgeon et al. 1999; Lu et al. 2001; Douglas and Brunner 2002; Palm et al. 2003; Stott et al. 2004; Fraser et al.2004; Beacham et al. 2005; Wofford et al. 2005). This was not surprising as microsatellites have become a common tool for assessing population structure primarily due to their higher levels of genetic diversity (Neff et al. 2000; Shaklee and Currrens 2003). Lake Michigan's lake whitefish populations exhibited higher levels of genetic variation at the microsatellite loci when compared with lake whitefish populations from Lake Ontario (A. Bernard, University of Guelph, personal communication), and two lakes on Isle Royale (Siskiwit and Desor; Stott et al. 2004) and similar levels of variation to those found in populations from Lake Superior and Lake Huron (Stott et al. 2004). Additionally, genetic diversity levels of the microsatellite loci met or exceeded the *a priori* levels I deemed necessary for population differentiation from a review of the literature. The combination of these facts shows that the microsatellite loci I selected for this study had appropriate diversity levels to assess population differentiation.

The microsatellite loci employed in this study were consistent with HWE expectations and showed populations with multiple year samples were homogeneous across the two sample years. An observed heterozygote excess in the data could have resulted from the presence of related individuals in an individual sample or from an

unexpected sampling bias in small effective size populations (Rasmussen 1979; Balloux 2004). This observed heterozygote excess nevertheless supported the homogeneity of multiyear samples. If multiple gene pools (i.e., populations) were inadvertently sampled at a single site, Wahlund's effect (Wahlund 1928) predicts a heterozygote deficit would have been present. The presence of a significant heterozygote excess rejected the idea of two-year samples representing different gene pools. Because the microsatellite loci in this study had sufficient levels of genetic diversity and conformed to HWE expectations they were considered suitable for subsequent analysis of lake whitefish spawning aggregates.

However, the diversity in the mtDNA was deemed insufficient to provide resolution of genetic structure. Other fish studies have shown nucleotide diversity measures orders of magnitude higher than the observed values (0.07 - 0.32; Lankford et al. 1999) and the near lack of shared haplotypes between sampled populations suggested a large number of sequences would be necessary if any genetic structure were to be resolved. The lack of mtDNA polymorphism in the Lake Michigan lake whitefish was similar to the levels present in lake whitefish mtDNA across its native geographic range ($\pi \leq 0.01$ in the Great Lakes samples; Bernatchez and Dodson 1991). The predominant haplotype for each gene region in my study was found in 85.6 and 90% of all samples. Bernatchez and Dodson (1991) found that 92% of all fish in the Mississippian assemblage (the entire Great Lakes basin and the majority of the lake whitefish native range) shared a single mtDNA haplotype. The low levels of observed diversity coupled with relatively few shared haplotypes violated the *a priori* conditions established for mtDNA utility in this study and the use of mtDNA for subsequent stock structure

delineation was abandoned.

Genetic Stock Identification of Lake Michigan Lake Whitefish

The initial phase of GSI showed Lake Michigan's lake whitefish populations did not constitute one panmictic group indicating that population structuring did indeed exist. These findings were not unanticipated based on the tendency of philopatric fish species to show intraspecific population differentiation (Stepien and Faber 1998; Markert et al. 1999; Duftner et al. 2006), previous population dynamics, tagging, and genetic studies conducted on lake whitefish throughout Lake Michigan (Imhoff 1977; Leary 1979; Leary et al. 1980; Ebener 1980; Rowe 1984; Ebener and Copes 1985; Scheerer and Taylor 1985; Walker et al. 1993), and the sheer geographic area considered. For comparison, Scheerer and Taylor (1985) showed that tag returns and vital statistics (e.g., mortality rates) indicated the existence of at least 3 discrete stocks of lake whitefish in northeastern Lake Michigan. Additionally, Imhoff (1977) suggested that two distinct lake whitefish populations existed in the Green Bay region of Lake Michigan. One in the Big Bay de Noc area of northern Green Bay and one on the lake side of the Door County (WI) peninsula encompassing the spawning samples from North and Moonlight Bay. Furthermore, the grouping of NMB and Cedar River fish into one stock was supported by tagging data that indicated NMB fish comprise a large majority of the spawning fish at Cedar River (Rowe 1984).

Cluster analyses based on D_c indicated strong clustering of geographically proximal populations (Figure 9). These geographic clusters were also supported by data from several tagging studies throughout the lake (Ebener 1980; Ebener and Copes 1985;

Scheerer and Taylor 1985; Walker et al. 1993). The six putative stocks showing no significant within group molecular variance according to AMOVA were also well supported by geographic location and the previous tagging, population dynamics, and genetic studies. Within-stock analysis (F_{st} analyses) indicated that all but one proposed stock was a stable grouping (Table 13). Significant F_{st} values (Table 13) between Hog Island and Traverse Bay indicated significant differences exist between these two populations. Moreover, pairwise F_{st} results suggest that gene flow occurs between the Hog Island population and the NOR stock but not between Hog Island and the Traverse Bay population. Scheerer and Taylor (1985) hypothesized that a shallow reef extending westward from the Waugoshance Point (Grid 319; Figure 4) through the northern Beaver Islands (Hog Island) acted as a barrier to the southward movement of the NOR populations. The AMOVA results support their findings of restricted gene flow (i.e., NOR and NOE populations are separate) but the F_{st} results suggested gene flow exists across this potential barrier because Hog Island was not significantly different than the two NOR populations. If this reef is an impediment but not a barrier to migration, significant gene flow between the NOR populations and Hog Island would be plausible. The rarity of this migration could be sufficient to explain no observed connectivity between the Traverse Bay and the NOR populations based on a stepping-stone migration model (Kimura and Weiss 1964). Despite the apparent gene flow between Hog Island and the NOR populations and the significant F_{st} value between Hog Island and Traverse Bay, the AMOVA and NJ (D_c) results support the grouping of Hog Island and Traverse Bay populations into a stock based on the lack of within group variance (Table 12; Figure 9).

The identity of six genetic units of lake whitefish in Lake Michigan does not preclude gene flow among the groups. However, the level of gene flow among stocks is not sufficient to consider the groups part of the same dynamic genetic pool. Several analyses indicated that if gene flow was occurring, it was predominately between neighboring populations, with little gene flow between more distant populations (Table 15; Figure 12). This scenario is consistent with the theory of isolation by distance proposed by Wright (1943) where interbreeding is restricted to small distances resulting in remote populations becoming isolated by distance and is exhibited by many species world-wide (Turgeon and Bernatchez 2001; Pogson et al. 2001; Planes and Fauvelot 2002; Primmer et al. 2006). Wright (1943) described that isolation by distance results from less mixing among individuals, or pairs of populations, which are situated further apart than among those which are separated by shorter distances. This suggests migrant lake whitefish are more likely to spawn on more geographically proximate spawning grounds than spawning grounds on the other side of the lake. The geographical clustering of the populations based on genic differentiation, genetic distance, AMOVA, and F_{st} all supported Wright's (1943) theory of isolation by distance.

Virtually all significant gene flow appeared to occur between populations within stocks, with relatively little gene flow occurring between stocks. This is illustrated by the single population stocks, Big Bay de Noc and Elk Rapids, being genetically isolated from other surrounding populations based on all genetic analyses conducted in this study. The BBN population has been shown to exhibit high spawning site fidelity (M.P. Ebener, CORA, personal communication) that, if coupled with little immigration from neighboring populations, would result in increased genetic divergence of BBN compared

to neighboring lake whitefish populations. The lack of gene flow between neighboring populations and the Elk Rapids population agrees with the findings of Walker et al. (1993) who found that little mixing, if any, occurred between fish from the west side of Grand Traverse Bay (Traverse Bay population) and fish from the east side of Grand Traverse Bay (Elk Rapids) due to a potential thermal barrier to gene flow. The outer bay (area directly north of Old Mission Point, the peninsula separating west bay from east bay) is divided by a deep trough (>75 m) that may thermally separate these bays (Walker et al. 1993). Thermal barriers have also been suggested to segregate stocks of lake whitefish in Lake Huron (Casselman et al. 1981). Additionally, depths in the Elk Rapids side of Grand Traverse Bay exceed 150 m compared to depths on the Traverse Bay side only reaching slightly over 100 m (National Oceanic and Atmospheric Administration 2007). These differences in depth may indicate two completely different habitat types leading to a reduction in the mixing of the stocks.

Population Specific Issues

Throughout the course of GSI the Menominee River population exhibited signs of potentially being of mixed-stock origins. Based on the test of panmixia, allele frequency distributions indicated that the Menominee River was similar to five other populations throughout the lake (Table 10). Cluster analysis also indicated that the Menominee River population may be of mixed-stock origins as it was the only population that failed to group geographically with its neighboring populations (Figure 8). This population may have been extirpated or nearly extirpated during the mid-1900s with a recent population increase (P.J. Peeters, WDNR, personal communication; P.J. Schneeberger, MDNR,

personal communication) corresponding to the recent increase in lake whitefish populations throughout the lake. If this is the case, a mixed-stock origin is plausible with the most likely sources of founders being NMB, Cedar River and BBN based on the significant isolation by distance findings. However, it is important to consider that the sample size from the Menominee River (n = 35) is only 50% that of the next smallest sample size (n = 70; Hog Island) and just over 10% of the largest sample size (n = 286; NMB). This coupled with only one year of sample data could have altered the outcomes of many of the statistical analyses. That is why the majority of the GSI tests where conducted without the Menominee River population. However, the cohesiveness and integrity of the Menominee River population should not be dismissed as a possibility. Sample sizes of ~35 have been shown to be successful in other population differentiation studies (Hansen et al. 1999; Stott et al. 2004; Stepien et al. 2006) and when the Menominee River population was included in GMA, the population had 100% self-assignment indicating it could be genetically distinguished from other populations. To more confidently determine the stock origins of the Menominee River sample, the genotyping of additional samples is required.

Throughout the GSI process, the Elk Rapids population consistently exhibited a higher degree of genetic differentiation from all other populations. This population was significantly different from all other populations in all statistical analyses, showed the highest number of private alleles, and was the only population not exhibiting an isolation by distance effect. In fact, the divergence of Elk Rapids compared to all other populations effectively masked the positive isolation by distance between the other populations (Table 15; Figures 10 and 11). As previously discussed, this population has

been shown to be highly philopatric (Walker et al. 1993) and to inhabit an area of the lake where water depths are much deeper than any of the surrounding areas of the lake. These findings coupled with other studies of philopatric fish species (Stepien and Faber 1997; Markert et al. 1999; Duftner et al. 2006) have led me to conclude that the Elk Rapids population represents a genetically divergent population reinforced by migratory restrictions or life history differences.

Effective Population Sizes of Lake Whitefish Populations

One way to gain insight into the functional genetic dynamics of a population is to estimate the effective population size. For most organisms, effective size to census size estimates should range between 0.25 and 0.75 (Turner et al. 1999). However, estimates of effective sizes for Lake Michigan lake whitefish populations exhibited a much lower ratio than described by Turner et al. (1999) and seem to more closely exhibit a ratio seen in recent empirical studies of species with high fecundity and high mortality at early life stages typical of lake whitefish (Hedgecock 1994; Vucetich et al. 1997).

Effective size can deviate from census size for several reasons (Frankham 1996). Two of these, large variance in female reproductive success (Hedgecock 1994) and fluctuating population sizes (Vucetich et al. 1997) are suspected to occur in Lake Michigan's lake whitefish populations. Because of the nature in which lake whitefish spawn, high variance in female reproductive success likely occurs, helping to explain the effective size estimates. Populations with the three largest effective population sizes (Epoufette, Big Bay de Noc, and North-Moonlight Bay populations) are also believed to be three of the largest spawning aggregates in Lake Michigan supporting the findings of

Frankham (1996) that genetic variation is positively correlated with population size.

Genetic Mixture Analysis

High levels of self assignment to the stock of origin in both the actual data and the simulated data indicate that our groupings (genetic stocks) are easily discernable from one another based on allele frequencies of the microsatellite loci. The high degree of self-assignment back to both the population and stock of origin based on GMA simulations indicate that not only are the stocks identified biologically relevant but also that they are genetically divergent enough that mixed-stock analysis is not only feasible, but can be accomplished with a high degree of accuracy. Through mixed stock analysis simulations I have shown the ability of the genetic data to accurately (≥91.14%) assign fish to their stock of origin. To truly manage on a stock basis, the proportions of each stock harvested in each management zone must be known throughout the course of the fishing season. As in every estimate there is bias, however, compared to the current management schemes, knowledge of the proportions of each stock being harvested in management zones throughout the lake if applied to SCAA models should yield a much better estimate of harvest on individual stocks. This would reduce some of the bias and eliminate some of the estimating that is necessary for an SCAA. The additional information mixed stock analysis could provide for the current SCAA models, could allow the lake whitefish commercial fishery of Lake Michigan to thrive where other commercial fisheries have vanished (e.g., Peruvian anchoveta, *Engraulis ringens*; Atlantic cod, *Gadus morhua*).

Management Implications

Genetic stock identification was first developed for use in Pacific salmonid commercial fisheries management specifically for the purpose of estimating the composition of each stock present in the commercial harvest (i.e., mixed stock analysis) in an attempt to reduce overharvest (Grant et al. 1980). Since its inception, GSI and mixed stock analysis have become powerful and widely used management tools in worldwide fisheries research (Waples et al. 1990; Angers et al. 1995; Ruzzante et al. 1998; Gatt et al. 2003; Beacham et al. 2005). Because sustainability of commercial fisheries is dependant upon both genetics and demographics, management based on a combination of these biological tools is logical. This combination of genetic and non-genetic data (e.g., age structure) has proven a powerful means of analyzing mixed-stock fisheries of sockeye salmon (*Oncorhynchus nerka*; Wood et al. 1989). Fortunately, a great deal of population dynamic and morphometric data is collected annually by all the management agencies involved with Lake Michigan's lake whitefish commercial fishery. If this demographic data were coupled with the genetic data from this study and a more comprehensive mixed-stock analysis, the current SCAA models used by managers will be better enabled to set quotas and harvest regulations to not only maximize yield, but also to maintain current genetic diversity within Lake Michigan.

Based on the genetic evidence collected in this study, I proposed between five and seven genetic stocks of lake whitefish within Lake Michigan; the majority of the evidence (i.e., NJ tree; AMOVA; GMA) suggested six genetic stocks. A review of previous studies examining lake whitefish stock structure using population dynamics, tagging data, and genetics (Imhoff 1977; Leary 1979; Imhoff et al. 1980; Ebener 1980;

Ebener and Copes 1985; Scheerer and Taylor 1985; Walker et al. 1993) shows supporting evidence for these six stocks. For example, Ebener and Copes (1985) found strong evidence for differentiation of the NMB and BBN populations into two separate stocks. Furthermore, they documented NMB fish movement in and out of Green Bay throughout the year allowing for gene flow in support of my inclusion of Menominee River and Cedar River into the NMB stock. Additionally, Walker et al. (1993) indicated relatively little mixing between the populations from the east (Elk Rapids) and west (Traverse Bay) basins of Grand Traverse Bay, supporting my findings of these populations being genetically differentiated.

Based on the genetic management zones (GMZs), established through a combination of GSI and contemporary management zones (Figure 14), the harvest from commercial and statistical management zones may need to be adjusted to properly manage on a genetic stock basis. Some GMZ boundaries correspond directly with the statistical management zones currently in place (e.g., GMZ 3 and WFM-03) indicating their initial founding was well justified. Other management zones may need to be split because of the presence of multiple genetic stocks present within one zone (e.g., WFM-05 has two genetic stocks present, EKR and NOE).

The goal of any commercial fishery should be optimum sustainable yield without disrupting the genetic characteristics of component stocks and Lake Michigan's lake whitefish commercial fishery is no different. Using a combination of GSI, mixed-stock analysis, and population dynamics and morphometrics will not only benefit Lake Michigan's lake whitefish commercial fishery and help ensure its sustainability, it will also benefit fish managers, fishermen, and local and regional economies. Additionally,

maintenance of the lake whitefish commercial fishery will preserve a long-standing tradition in the Great Lakes region.

FUTURE RESEARCH

Additional Research on Current Project

As with any research project, there are limitations on time, money, and effort which result in compromises. Although this data set is large and statistical power appears to be high, there are a few minor additions that would have improved the confidence in my findings. One thing that would improve this study would be to increase the sample size of all populations to at least 75-100 individuals. Another thing would be to get an additional sample from the Menominee River, Traverse Bay, Elk Rapids, and Hog Island populations. Based on the results from the 2005 data alone versus the 2005 and 2006 data combined, it seems that a second year of sampling was helpful in determining stable population structure and in reducing potential bias due to the nature of sampling the commercial harvest (e.g., sampling kin groups; Rasmussen 1979). Additionally, a sample from the Leland area (WFM-06) and the Seul Choix Point area (WFM-02) could improve resolution of GMZ boundary determination as there are spawning aggregates and suspected stocks there (Scheerer and Taylor 1985; M.P. Ebener, CORA, personal communication) and these management zones are presently unaccounted for. The combination of my current data and the minor additions would give the added confidence in identifying the genetic population structure of Lake Michigan's lake whitefish. Additionally, multi-year samples of juvenile or larval lake whitefish from known spawning aggregates could provide a more accurate estimate of N_e allowing managers to identify stocks with smaller effective sizes and manage for the smallest stocks present in a commercial zone.

Mixed Stock Analysis

The delineation of genetic stocks of Lake Michigan lake whitefish provides opportunities for several other questions and problems to be subsequently addressed. First, I would suggest comprehensive mixed stock analysis for the commercial lake whitefish fishery be conducted. Because of the high migration rate of lake whitefish (M. P. Ebener, CORA, unpublished data) and the inter-jurisdictional nature of the fishery, there is a high probability the commercial fishery is operating as a mixed stock fishery. Currently, managers use statistical catch at-age models for the majority of the statistical management zones. If managers knew what proportion of the fish captured in their management zones belonged to each genetic stock throughout the course of the fishing season, and coupled this information with growth and abundance estimates from their SCAA models, then each statistical management zone could be managed in an inter-jurisdictional manner on a true stock basis. The estimates and simulations run for this study indicate that mixed stock analysis is not only a possibility, but with slight refining of the markers used, could be highly useful with great statistical power and accuracy.

Historic Stock Structure Identification

A second study that should be considered is to determine if contemporary genetic population structure is representative of the historic stocks present in Lake Michigan. A comparison of historic versus contemporary stock structure will allow researchers and managers to understand the long-term stability of the contemporary structure. Additionally, knowledge of historic stock structure should help answer some of the questions left after this study of whether or not populations are diverging due to homing,

selection, or drift, or whether the populations are homogenizing due to migration. A historical survey could also indicate whether the amount of mtDNA diversity present in the contemporary study reflects that present in the historical stocks. Determination of the historic genetic diversity within Lake Michigan could also shed light on whether or not any Lake Michigan populations were historically bottlenecked, or extirpated. If these populations have been re-founded then these may be areas where stocks were more sensitive to environmental conditions, or where exotic species hit native stocks the hardest. Knowing this would allow managers to consider, for example, habitat rehabilitation in these areas, or potentially consider more sea lamprey control. Another thing this project could assess is how commercial harvest has affected the effective population size over time. Most importantly, a historic stock structure research project would allow managers to assess the stability of the current population structure over time. This would allow them to judge the importance of each stock within the lake relative to its long-term stability, contribution to the fishery, and overall genetic impacts on the lake as a whole.

Standardization of Genetic Data across the Great Lakes Region

Standardization of microsatellite genetic data across the Great Lakes region is essential for a variety of reasons. First, it allows researchers and managers to fill in gaps from individual research such as this project. For example, if the population structure of lake whitefish in Lake Huron was determined and the genetic data was standardized across the region, diversity levels could be compared between lakes. Additionally, managers could estimate migration rates between the lakes, and determine if any stocks

in Lake Michigan were founded from or receive significant gene flow from Lake Huron and vice-versa (i.e., Elk Rapids). Because there is no physical or apparent chemical barrier between these two lakes, lake whitefish from either lake could subsequently be harvested in the other lake throughout the course of fishing season. Knowledge of this would allow managers to better set quotas for each stock.

This data would also help to shed light on why populations exhibit certain genetic characteristics and may identify populations that may be imperiled or genetically bottlenecked. If lake whitefish populations ever declined in an area so much that supplemental stocking, or stocking to rehabilitate a population was required, knowledge of the genetic diversity available both within the lake and within surrounding lakes (genetic stock boundaries) is essential. Without this knowledge, our ability to make correct management decisions would be dramatically hampered. This could potentially result in fish being stocked across genetic boundaries and lead to outbreeding depression or introgression.

SUMMARY

Lake whitefish have comprised an important commercial fishery on Lake Michigan since the early 1800s. Previous studies have indicated potential stock structure, however, questions still exist regarding the number, distribution, and discreteness of lake whitefish stocks in Lake Michigan. Optimal long-term management and conservation of the lake whitefish commercial fishery depends upon knowing the number, distribution and characteristics of all component stocks and maintaining their integrity, diversity, and abundance.

This study aimed to identify the genetic stock structure of lake whitefish in Lake Michigan. The first objective was to determine the extent of lake whitefish microsatellite and mitochondrial DNA genetic diversity within and among putative lake whitefish spawning aggregates and determine the utility of these markers to discriminate among spawning stocks of lake whitefish and the second objective was to describe the genetic population structure of spawning lake whitefish aggregates in northern Lake Michigan and Green Bay in terms of genetic stock identification and degree of stock isolation. These objectives were then tested with two primary assumptions. Based on evidence of philopatry, I assumed that distinct spawning aggregates represented potential gene pools and I assumed that differences at genetic markers underlie population differentiation.

I assessed the level of polymorphism of microsatellites and mitochondrial DNA to determine a combination of genetic markers adequate to delimit potential stocks. These genetic markers were used to describe the genetic structure among spawning lake whitefish aggregates (N = 6) in northern Lake Michigan and Green Bay and several smaller spawning groups (N = 6) throughout Lake Michigan. Microsatellites exhibited

adequate levels of diversity for population differentiation, were considered to meet HWE expectations and therefore the utility of the markers was considered sufficient. Mitochondrial DNA genetic diversity was low within and between sampled populations and did not meet *a priori* levels of genetic diversity deemed necessary for population differentiation and further use of the molecule was abandoned.

Microsatellite genetic diversity measures rejected the null hypothesis of panmixia within Lake Michigan's lake whitefish populations. Cluster analysis based on genetic distance indicated 5-7 potential genetic stocks were present. The clustering of populations based on genetic distance also corresponded to geographic locations of the sampled populations. Analysis of molecular variance suggested six genetic stocks were present in the lake as a significant portion of variation was attributed to between groups of populations but not within groups of populations. Within stock analyses (pairwise F_{st}) suggested all but one stock delineated by AMOVA was a stable grouping with the exception being the Hog Island/Traverse Bay grouping (NOE stock). This putative stock showed significant differences between the two populations indicating significant gene flow between the NOR populations and Hog Island, but not between NOR and Traverse Bay. Elk Rapids was the most genetically divergent population and actually masked the isolation by distance that exists between all other populations within Lake Michigan.

Genetic mixture analysis simulations indicate that mixed-stock analysis is not only feasible but can be accomplished with a high degree of accuracy. The coupling of genetic data from this study with current demographic data and a comprehensive mixed-stock analysis would allow for more efficient and effective management of this economically and socially important resource.

LITERATURE CITED

Angers, B., L. Bernatchez, A. Angers, and L. Desgroseillers. 1995. Specific microsatellite loci for brook charr (*Salvelinus fontinalis*, Salmonidae) reveal strong population subdivision on a microgeographical scale. Journal of Fish Biology 47: 177-185.

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

Bailey, G.S., G.T. Cocks, and A.C. Wilson. 1969. Gene duplication in fishes: malate dehydrogenases of salmon and trout. Biochemistry and Biophysics. Research Communications 34: 605-611.

Bailey, R.M. and G.R. Smith. 1981. Origin and geography of the fish fauna of the Laurentian Great Lakes basin. Canadian Journal of Fisheries and Aquatic Sciences 38: 1539-1561.

Baldwin, N. A., R. W. Saalfeld, M. R. Dochoda, H. J. Buettner, and R.L. Eshenroder. 2002. Commercial Fish Production in the Great Lakes 1867-2000. Available: http://www.glfc.org/databases/commercial/commerc.php (June 2006)

Balloux, F. 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. Evolution 58: 1891-1900.

Bartley, D., M. Bagley, G. Gall, and B. Bentley. 1992 Use of linkage disequilibrium data to estimate effective size of hatchery and natural fish populations. Conservation Biology 6: 365-375.

Beacham, T.D., B. McIntosh, and C. MacConnachie. 2005. Population structure and stock identification of sockeye salmon (*Oncorhynchus nerka*) in coastal lakes in British Columbia, Canada. Canadian Journal of Zoology 83: 934-844.

Becker, G. C. 1983. Fishes of Wisconsin. The University of Wisconsin Press, Madison, Wisconsin.

Bernatchez, L., and J.J. Dodson. 1991. Phylogeographic structure in Mitochondrial DNA of the Lake Whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene Glaciations. Evolution 4: 1016-1035.

Bernatchez L, R. Guyomard, and F. Bonhomme. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. Molecular Ecology**,** 1: 161-173.

Bernatchez, L. and C.C. Wilson. 1998. Comparative phylogeography of nearctic and Palearctic fishes. Molecular Ecology 7: 431-452.

Bernatchez, L., and P. Duchesne. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? Canadian Journal of Fisheries and Aquatic Sciences 57: 1-12.

Billington, N. 2003. Mitochondrial DNA. Pages 59-100 in E.M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Billington, N., and P.D.N. Hebert. 1991. Mitochondrial DNA diversity in fishes and its implications for introductions. Canadian Journal of Fisheries and Aquatic Sciences 48 (Supplement 1): 80-94.

Bohonak, A.J. 2002. IBD (Isolation By Distance): a program for analyses of isolation by distance. Journal of Heredity 93: 153-154.

Burger, C.V., W.J. Spearman, and M. Cronin. 1997. Genetic differentiation of sockeye Salmon subpopulations from a geologically young Alaskan lake system. Transactions of the American Fisheries Society 126: 926-938.

Campton, D. 1987. Natural hybridization and introgression in fishes: methods of detection and genetic interpretations. Pages 161-192 in N. Ryman and F. Utter, editors. Population genetics and fishery management. University of Washington Press, Seattle, Washington.

Casselman, J.M., J.J. Collins, E.J. Crossman, P.E. Ihssen, and G.R. Spangler. 1981. Lake whitefish (*Coregonus clupeaformis*) stocks of the Ontario waters of Lake Huron. Canadian Journal of Fisheries and Aquatic Sciences 38: 1772-1789.

Casselman, J.M., K.T. Scribner, and G.R. Spangler. 2001. Review of the lake whitefish stocks in northern Lake Michigan, with special reference to stock structure and spawning site distribution in relation to Green Bay. Lake Michigan Technical Committee Meeting, July 17, 2001.

Cavalli-Sforza, L.L., and A.W.F. Edwards. 1967. Phylogenetic analysis: Models and estimation procedures. Evolution 21: 550-570.

Cronin, M.A., W.J. Spearman, R.L. Wilmot, J.C. Patton, and J.W. Bickham. 1993. Mitochondrial DNA variation in chinook (*Oncorhynchus tshawytcha*) and chum salmon (*O. keta*) detected by restriction enzyme analysis of polymerase chain reaction (PCR) products. Canadian Journal of Fisheries and Aquatic Sciences 50: 708-715.

Dominguez, E.V., D. Paetkau, N. Tucker, G. Hinten, and C. Moritz. 2001. Resolution of natural groups using iterative assignment tests: an example from two species of Australian native rats (*Rattus*). Molecular Ecology 10: 2069–2078

Don, R.H., P.T. Cox, B.J. Wainwright, K. Baker, and J.S. Mattick. 1991. 'Touchdown' PCR to Circumvent spurious priming during gene amplification. Nucleic Acids Research 19: 4008.

Douglas, M. R., P. C. Brunner. 2002. Biodiversity of central alpine *Coregonus* (Salmoniformes): impact of one-hundred years of management. Ecological Applications 12: 154-172.

Douglas, M. R., P. C. Brunner, and L. Bernatchez 1999. Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the Central Alpine region of Europe represent Species flocks? Molecular Ecology 8: 589-603.

Drummond, A.J., M. Kearse, J. Heled, R. Moir, T. Thiever, B.Ashton, and A. Wilson. 2006. Geneious v.2.5. Available from http://www.geneious.com.

Duftner, N., K.M.. Sefc, S. Koblmüller, B. Nevado, E. Verheyen, H. Phiri and C. Sturmbauer. 2006. Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. Molecular Ecology 15: 2381-2395.

Ebener, M.P. 1980. Population dynamics of lake whitefish, *Coregonus clupeaformis*, in Green Bay and Lake Michigan east of Door County, Wisconsin. M.S. thesis, University of Wisconsin-Stevens Point.

Ebener, M.P., and F.A. Copes. 1985. Population statistics, yield estimates, and Management considerations for two lake whitefish stocks in Lake Michigan. North American Journal of Fisheries Management 5: 435-448.

Ebener, M. P. 1997. Recovery of lake whitefish populations in the Great Lakes. Fisheries 22: 18-20.

Ebener, M.P., J.R. Bence, K.R. Newman, and P.J. Schneeberger. 2005. Application of statistical catch-at-age models to assess lake whitefish stocks in the 1836 Treaty Ceded waters of the upper Great Lakes. Proceedings of a workshop on the dynamics of lake whitefish (*Coregonus clupeaformis*) and the amphipod *Diporeia spp.* in the Great Lakes. Great Lakes Fishery Commission Technical Report number 66: 271-310

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.

Excoffier, L. G. Laval, and S. Schneider. 2005. Arlequin version. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Felsenstein, J. 1993. Phylogeny inference package (PHYLIP), version 3.5. University of Washington, Seattle, Washington.

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conservation Biology 10: 1500-1508.

Fraser, D.J., C. Lippe, and L. Bernatchez. 2004. Consequences of unequal population size, asymmetric gene flow and sex-biased dispersal on population structure in brook charr (*Salvelinus fontinalis*). Molecular Ecology 13: 67-80.

Gatt, M. H., T. L. McParland, L. C. Halyk, and M. M. Ferguson. 2003. Mitochondrial DNA variation and mixed-stock analysis of recreational and commercial walleye fisheries in eastern Lake Erie. North American Journal of Fisheries Management 23: 431-440.

Goudet, J. 1995. Fstat version 1.2: a computer program to calculate F-statistics. Journal of Heredity 86: 485-486.

Grant, W.S., G. B. Milner, P. Krasnowski, and F.M. Utter. 1980. Use of biogeochemical genetic variants for identification of sockeye salmon (*Oncorhynchus nerka)* stocks in Cook Inlet, Alaska. Canadian Journal of Fisheries and Aquatic Sciences 37: 1236-1247.

Great Lakes Fisheries Commission. 1961. Fishery statistical districts of the Great Lakes. Great Lakes Fisheries Commission Technical Report 2.

Guo, S.W., and E.A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48: 361-372.

Hallerman, E. M. 2003. Random genetic drift. Pages 197- 214 in E.M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Hansen, M.M., K-L.D. Mensberg, and S. Berg. 1999. Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus*) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. Molecular Ecology 8: 239-252.

Hardy, G. H. 1908. Mendelian proportions in a mixed population. Science 28: 49–50.

Hartl, D.L., and A.G. Clark. 1997. Principles of population genetics, 3rd edition. Sinauer Associates, Inc. Sunderland, Massachusetts.

Hedgecock, D. 1994. Does variance in reproductive success limit effective population sizes of marine organisms? Pages 122-134 in A.R. Beaumont, editor. Genetics and Evolution of Aquatic Organisms. Chapman and Hall, New York, New York.

Hedrick, P.W. 2000. Genetics of populations. Jones and Bartlett Publishers, Boston, Massachusetts.

Hellberg, M.E. 1994. Relationships between inferred levels of gene flow and geographic distance In a philopatric coral, *Balanophyllia elegans*. Evolution 48: 1829-1854.

Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetic Research 38: 209-216.

Imhoff, M.A. 1977. Population genetic structure of lake whitefish, *Coregonus clupeaformis*, in Green Bay, and northern Lake Michigan, as assessed by electrophoresis of lactate, glycerol-3-phosphate, and malate dehydrogenase isozymes. M.S. thesis, University of Wisconsin-Stevens Point.

Imhoff, M.A., R. F. Leary, and H.E. Booke. 1980. Population or stock structure of lake whitefish (*Coregonus clupeaformis*) in northern Lake Michigan as assessed by isozyme electrophoresis. Canadian Journal of Fisheries and Aquatic Sciences 37: 783-793.

Kalinowski, S.T. 2003. Genetic mixture analysis 1.0. Department of Ecology, Montana State University, Bozeman 59717. http://www.montana.edu/kalinowski.htm

Kalinowski, S.T. 2005. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes 5: 187-189.

Kimura, M. and J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725-738.

Kimura, M. and G.H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49: 561-576.

Kinnunen, R. E. 2003. Great Lakes Commercial Fisheries. Great Lakes Fisheries Leadership Institute Technology Report.

Lankford, T. E., T. E. Targett, and P.M. Gaffney. 1999. Mitochondrial DNA analysis of population structure in the Atlantic croaker, *Micropogonias undulates* (Perciformes: Sciaenidae). Fishery Bulletin 97: 884–890.

Leary, R. 1979. Population or stock structure of lake whitefish, *Coregonus clupeaformis* in northern Lake Michigan as assessed by isozyme electrophoresis. M.S. thesis, University of Wisconsin-Stevens Point.

Leberg, P.L. 2002. Estimating allelic richness: effects of sample size and bottlenecks. Molecular Ecology 11: 2445-2449.

Liu, K. and S.V. Muse. 2005. PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics 21: 2128-2129.

Lowe, A., S. Harris, and P. Ashton. 2004. Ecological genetics: design, analysis and application. Blackwell Publishing Company. Malden, Massachusetts.

Lu, G., D.J., Basley, and L. Bernatchez. 2001. Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. Molecular Ecology 10: 965-985.

Lu, G., and L. Bernatchez. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. Evolution 53: 1491-1505.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach. Cancer Research 27: 209-220.

Markert , J. A., M. E. Arnegard, P. D. Danley and T. D. Kocher. 1999. Biogeography and population genetics of the Lake Malawi Cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. Molecular Ecology 8: 1013–1026.

Miller, C.R., P. Joyce, and L. P. Waits. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160**:** 357–366.

Navidi, W., N. Arnheim and M. S. Waterman. 1992. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. American Journal of Human Genetics 50: 347–359.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145-163.

Nei, M. and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press. Oxford, New York.

Neff, B. D., P. Fu, and M. R. Gross. 2000. Microsatellite multiplexing in fish. Transactions of the American Fisheries Society 129: 584-593.

Nielsen, E. E., P. H.. Nielsen, D. Meldrup and M. M. Hansen. 2004. Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. Molecular Ecology 13: 585–595.

National Oceanic and Atmospheric Administration 2007. Bathymetry of Lake Michigan. Available: http://www.ngdc.noaa.gov/mgg/greatlakes/lakemich_cdrom/images/area7hi.gif. (March 2007).

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357-358.

Pogson, G.H., C.T. Taggart, K.A. Mesa, and R.G. Boutillier. 2001. Isolation by distance in the Atlantic cod, *Gadus morhua*, at large and small geographic scales. Evolution 55: 131-146.

Palm, S., J. Dannewitz, T. Jarvi, E. Petersson, T. Prestegaard, and N. Ryman. 2003. Lack of molecular genetic divergence between sea-ranched and wild sea trout (*Salmo trutta*) Molecular Ecology 12: 2057-2071.

Pamilo, P. and S. Varvio-Aho. 1984. Testing genotype frequencies and heterozygosities. Marine Biology 79: 99-100.

Park, S.D.E. 2001 Microsatellite Toolkit. Available: http://acer.gen.tcd.ie/~sdepark/ms toolkit/ (April 2006).

Patton, J.C., B.J. Gallaway, R.G. Fechhelm, and M.A. Cronin. 1997. Genetic variation of microsatellite and mitochondrial DNA genetic markers in broad whitefish (*Coregonus nasus*) in the Colville and Sagavanirktok rivers in northern Canada. Canadian Journal of Fisheries and Aquatic Sciences 54: 1548-1556.

Peel, D., J.R. Ovenden, and S.L. Peel. 2004. Ne Estimator: software for estimating effective population size. Version 1.3 Queensland Government, Department of Primary Industry and Fisheries.

Petit, R, A. El Mousadik, and O. Pons. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12: 844-855.

Planes, S. and C. Fauvelot. 2002. Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. Evolution; Journal of Organic Evolution 56: 378-399.

Primmer, C.R., A.J. Veselov, A. Zubechenko, A. Poututkin, I. Bakhmet, and M.T. Koskinen. 2006. Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. Molecular Ecology 15:653–666.

Ranala, B., and J.L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences 94: 9197-9201.

Raymond, M., and F. Rousset, 1995. An exact test for differentiation. Evolution 49:1280 1283.

Reckhan, J.A. 1995. A graphical paradigm for the sequential reduction and spectacular rehabilitation of the lake whitefish of Lake Huron. Pages 171-190 in M. Munawar, T.Edsall, and J. Leach, editors. The Lake Huron ecosystem: ecology, fisheries and management. Ecovision world monograph series, SPB Academic Publishing, Amsterdam, The Netherlands.

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

Rogers, S.M., M.H. Marchand, and L. Bernatchez. 2004. Isolation, characterization and cross-salmonid amplification of 31 microsatellite loci in lake whitefish (*Coregonus clupeaformis*, Mitchill). Molecular Ecology Notes 4: 89-92.

Rowe, M. 1984. Population dynamics of lake whitefish in the Big Bay de Noc, Bark Cedar Rivers, and Portage Bay Areas of Lake Michigan. M.S. thesis, University of Wisconsin-Stevens Point.

Ruzzante, D.E. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. Canadian Journal of Fisheries and Aquatic Sciences 55: 1-14.

Ruzzante, D.E., C.T. Taggart, and D. Cook. 1998. A nuclear DNA basis for shelf- and band-scale population structure in northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. Molecular Ecology 7: 1663-1680.

Ryman, N., S. Palm, C. Andre, G. R. Carvalho, T.G. Dahlgren, P.E. Jorde, L. Laikre, C. Larsson, A. Palme, and D.E. Ruzzante. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. Molecular Ecology 15: 2031-2045.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: A method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

Scheerer, P.D., and W.W. Taylor. 1985. Population dynamics and stock differentiation of lake whitefish in northeastern Lake Michigan with implications for their management. North American Journal of Fisheries Management 5: 526-536.

Schneeberger, P.J. 2004. Assessment of whitefish populations in the treaty area of Lake Michigan: study completion report. Michigan Department of Natural Resources.

Shaklee, J. B., and K. P. Currens. 2003. Genetic stock identification and risk assessment. Pages 291-328 in E.M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Spangler, G.R., A.H. Berst, and J.F. Koonce. 1981. Perspectives and policy recommendations on the relevance of the stock concept to fishery management. Canadian Journal of Fisheries and Aquatic Sciences 38: 1908-1914.

Spangler, G.R., and J.H. Peters. 1995. Fisheries of Lake Huron: an opportunity for stewardship. Pages 103-123 in M. Munawar, T. Edsall, and J. Leach, editors. The Lake Huron Ecosystem: ecology, fisheries and management. Ecovision world monograph series, SPB Academic Publishing, Amsterdam, The Netherlands.

Stepien, C. A., and J. E. Faber. 1998. Population genetic structure, phylogeography and spawning philopatry in walleye (*Stizostedion vitreum*) from mitochondrial DNA control region sequences. Molecular Ecology 7:1757–1769.

Stepien, C.A., D.J. Murphy, and R.M. Strange. 2006. Broad-to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography.

Takezaki, N. and M. Nei. 1996. Genetic distances and reconstruction of Phylogenetic trees from Microsatellite DNA. Genetics 144: 389-399.

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

Turgeon, J., A. Estoup, and L. Bernatchez. 1999. Species flock in the North American Great Lakes: molecular ecology of Lake Nipigon ciscoes (Teleostei: Coregonidae: *Coregonus*). Evolution 53:1857–1871.

Turgeon, J., and L. Bernatchez. 2001. Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei:Coregoninae). Evolution 55: 2274-2286.

Turner, T.F., L.R. Richardson, and J.R. Gold. 1999. Temporal genetic variation of mitochondrial DNA and female effective population size of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico. Molecular Ecology 8: 1223-1229.

Vucetich, J.A., T.A. Waite, L. Nunney. 1997. Fluctuating population size and the ratio of effective to census population size. Evolution 51: 2017-2021.

Walker, S.H., M.W. Prout, W.W. Taylor, and S.R. Winterstein. 1993. Population dynamics and management of lake whitefish stocks in Grand Traverse Bay, Lake Michigan. North American Journal of Fisheries and Aquatic Sciences 13: 73-85.

Wahlund, S. 1928. Zusammensetzung von Population und Korrelationserscheinung vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas 11:65–106

Waples, R.S., D.J. Teel, J.M. Myers, and A.R. Marshall. 2004. Life-history divergence in chinook salmon: historic contingency and parallel evolution. Evolution 58: 386 403.

Waples, R. S., G. A. Winans, F. M. Utter, and C. Mahnken. 1990. Genetic approaches to the Management of pacific salmon. Fisheries 15: 19-25.

Weinberg, W. 1908. Over the proof of the transmission with humans. Annual volume booklets of the association for patriotic natural history in Wuerttemberg 64: 368-382.

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

Wisconsin Sea Grant. 2007. Available: http://www.seagrant.wisc.edu/greatlakesfish/lakewhitefish.html

Wilmot, R.L. and C.V. Burger. 1985. Genetic differences among populations of Alaskan sockeye salmon. Transactions of the American Fisheries Society 114: 236-243.

Wood, C.C., D.T. Rutherford, and S. McKinnell. 1989. Identification of sockeye salmon (*Oncorhynchus nerka*) stocks in mixed-stock fisheries in British Columbia and southeast Alaska using biological markers. Canadian Journal of Fisheries and Aquatic Sciences 46: 2108-2120.

Wright, G. M, and M. P. Ebener. In press. Potential effects of dietary lipid reduction on growth and reproduction of lake whitefish in northern Lake Michigan. Pages xxx-xxx in M. Jankun, P. Brzuzan, P. Hliwa, and M. Luczynski, editors. Biology and management of coregonid fishes 2005. Archives of Hydrobiology Special Issues Advances in Limnology 61.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97-159.

Wright, S. 1943. Isolation by distance. Genetics 28: 114–138.

Table 1. Lake Michigan commercial fish harvest by kgs and value in the year 2000.

Species	Scientific Name	U.S. Kgs	U.S. Value	$/kg
Lake Whitefish	*Coregonus clupeaformis*	2,174,108	$5,152,409	$2.36
Chubs	*Coregonus hoyi*	695,137	$1,532,739	$2.20
Smelt	*Osmerus mordax*	695,137	$721,539	$4.10
Lake Trout	*Salvelinus namaycush*	282,349	$336,212	$1.19
Yellow Perch	*Perca flavescens*	26,308	$153,113	$5.82
Round Whitefish	*Prosopium cylindraceum*	12,030	$11,102	$0.93

Table 2. Great Lakes commercial fish harvest by kgs and value (U.S. Dollars) in the year 2000 (from Kinnunen 2003).

Species	Scientific Name	Total kgs	Total Value	$/Kg
Lake Whitefish	*Coregonus clupeaformis*	15,651,331	$18,635,839	$0.40
Yellow Perch	*Perca flavescens*	4,534,438	$10,921,975	$0.96
Walleye	*Sander vitreus*	7,279,383	$10,120,227	$0.63
Chubs	*Coregonus hoyi*	1,037,362	$1,890,406	$0.44
Smelt	*Osmerus mordax*	7,399,033	$1,859,772	$0.11
Lake Trout	*Salvelinus namaycush*	1,013,907	$761,560	$0.22
Channel Catfish	*Ictalurus punctatus*	261,103	$309,032	$0.26
Carp	*Cyprinus carpio*	788,503	$160,636	$0.05

Table 3. a) Microsatellite and b) Mitochondrial DNA primers, primer sequences, observed allele size range in base pairs (microsatellites), number of alleles observed (A; microsatellites), number of haplotypes (H; mtDNA) and references.

a)

Locus	Locus Abbre-viation	Primer Sequence (5'-3')	Allele size (bp)	A	Reference
Cocl-23	C23	gctgatgaggatagcattc gcattaggtcgttttgtg	250-278	15	Bernatchez 1996
Bwf-1	B1	gatcagagaaatacacacaacgcatcaa cagcggttccattactgagcac	169-243	23	Patton et al. 1997
Bwf-2	B2	gggatacatcggcaacctctg agacagtccccaatgagaaaa	140-166	13	Patton et al. 1997
Cocl-lav 18	C18	aacaaactaaaacatcccaagtc ttagattggggcctaccttg	148-160	7	Rogers et al. 2004
Cocl-lav 68	C68	gtgtgttacaagtggctatg gtgatggctttcagaggc	172-184	7	Rogers et al. 2004
Cocl-lav 4	C4	tggtgtaatggcttttcctg gggagcaacattggactctc	148-158	6	Rogers et al. 2004
Cocl-lav 49	C49	agccagttggaggctatttg agggctgctgttgaagtcat	163-243	37	Rogers et al. 2004
Cocl-lav 41	C41	aaacaaacagtggtggagtgg gccagcactctctcatgctttt	181-233	21	Rogers et al. 2004
Cocl-lav 6	C6	gccatcatcctcccaggaaac cagggaatctgcactggagc	148-158	19	Rogers et al. 2004
Cocl-lav 45	C45	gagtgacagcagggagcag ggctcggttgaaagttgaga	239-259	9	Rogers et al. 2004
Cocl-lav 28	C28	acaatagcaggccattcagg ccaatcttcaaagccatttca	168-182	8	Rogers et al. 2004
Cocl-lav 52	C52	ggcgattgggagagtgatta acagagccccagatggtaac	90-164	31	Rogers et al. 2004

b)

Gene Region	Locus Abbre-viation	Primer Sequence (5'-3')	Size (bp)	H	Reference
D-loop	na	accactagcacccaaagcta	~1,100	7	Bernatchez et al. 1992
	na	gtgttatgctttagttaagc			Bernatchez et al. 1992
ND 5/6	na	aatagtttatccrttggtcttagg	~2,400	6	Cronin et al. 1993
	na	ttacaacgatggtttttcatrtca			Cronin et al. 1993

Table 4. PCR reaction cocktail recipes, fluorescent labels and thermocycler temperature profiles for all developed multiplexes and singlet PCR reactions.

Locus	Multiplex	10X Buffer (Conc.)	MgCl2 (Conc.)	Primer F (Conc.)	Primer R (Conc.)	Label
B1[1]	1	1.4X	1.7mM	0.25µM	0.25µM	Ned
C72				0.4µM	0.4µM	6Fam
C6				0.08µM	0.08µM	Ned
C18[2]	2	2X	2.1mM	0.32µM	0.32µM	Hex
C68				0.25µM	0.25µM	Ned
C4				0.075µM	0.075µM	6Fam
B2[1]	3	1X	2mM	0.1µM	0.1µM	6Fam
C23				0.025µM	0.025µM	6Fam
C28[2]	4	2X	2mM	0.2µM	0.2µM	Hex
C45				0.18µM	0.18µM	Ned
C49[1]		1.5X	2mM	0.3µM	0.3µM	Ned
C52[3]		1.5X	1.75mM	0.04µM	0.04µM	6Fam
C41[4]		1.5X	1.5mM	0.2µM	0.2µM	Ned
ND5/6[5]		1.0X	1.75mM	0.5µM	0.5µM	n/a n/a
D-loop5		1.0X	1.75mM	0.5µM	0.5µM	n/a n/a

1	94° C for 3 min. 6 series of 5 cycles each at 94° C for 30 s, then 60, 59, 58, 57, 56, and 55° C annealing for 30 s. 72° C for 30 s then a final elongation of 72° C for 7 min.
2	94° C for 3 min. 7 series of 5 cycles each at 94° C for 30 s, then 62.5, 62, 61.5, 61, 60.5, 60, and 59° C annealing for 30 s. 72° C for 30 s then a final elongation of 72° C for 7 min.
3	94° C for 3 min. 2 series of 5 cycles each at 94° C for 30 s, then 63, and 62° C annealing for 30 s, then 72° C for 30 s. Then 2 series of 8 cycles each at 94° C for 30 s, then 61, and 60.5° C annealing for 30 s, then 72° C for 30 s.Then a final series of 5 cycles of 94° C for 30 s, then 60° C annealing for 30 s, then 72° C for 30 s and a final elongation of 72° C for 7 min.
4	94° C for 3 min. 7 series of 5 cycles each at 94° C for 30 s, then 60, 59.5, 59, 58.5, 58, 57.5 and 57° C annealing for 30 s. 72° C for 30 s then a final elongation of 72° C for 7 min.
5	94° C for 5 min. 35 cycles each at 94° C for 45 s, then 52° C annealing for 45 s., then 72° C for 45 s 2 with a final elongation of 60° C for 40 min.

Table 5. Population statistics for all 12 sampled populations.

Pop.	Sample size	Loci typed	Unbiased Hz	Unbiased Hz SD	Obs Hz	Obs Hz SD	No Alleles/locus	No Alleles SD
Epo	137	12	0.6452	0.0546	0.609	0.0125	10.25	7.1
Nau	132	12	0.6511	0.0516	0.624	0.0126	9.92	5.76
TB	74	12	0.648	0.046	0.655	0.016	7.83	5.22
ER	72	12	0.6467	0.0515	0.587	0.0178	8.5	5.14
HgIs	70	12	0.6481	0.0531	0.622	0.0169	8.08	5
Sau	132	12	0.6486	0.0518	0.622	0.0126	8.75	4.99
Mus	133	12	0.6526	0.0503	0.609	0.0125	9.25	5.64
Lud	124	12	0.6637	0.0472	0.654	0.0128	9.58	5.68
MR	35	12	0.6357	0.0544	0.618	0.0239	6.92	4.19
CR	149	12	0.6632	0.047	0.627	0.0118	10.25	5.77
NMB	286	12	0.6392	0.049	0.628	0.0084	11.92	7.25
BBN	280	12	0.6555	0.0518	0.636	0.0084	11.17	7.26

Table 6. Allelic richness estimates based on rarefaction (Kalinowski 2005) and mean values for each locus and each population. Rarefaction was based on second smallest sample size (N = 70; Hog Island).

	B1	C6	C18	C68	C4	B2	C23	C49	C52	C41	C28	C45	Mean
Epo	9.88	5.62	2.00	3.96	3.86	7.03	10.12	17.26	18.77	14.88	3.94	5.33	8.55
Nau	8.54	6.18	3.72	5.20	4.38	6.87	11.21	13.43	17.10	13.63	3.59	5.49	8.28
TB	9.72	4.90	2.90	3.00	3.69	5.59	8.88	9.16	18.91	12.27	4.36	5.00	7.37
ER	12.01	9.93	4.49	2.75	3.00	6.00	9.70	17.04	13.28	9.81	3.85	4.00	7.99
HgIs	11.21	5.50	2.00	3.71	3.00	5.64	10.68	10.36	16.97	13.28	4.00	5.75	7.68
Sau	5.62	5.03	3.93	4.41	3.62	6.02	9.64	14.30	14.79	12.74	3.35	4.95	7.37
Mus	6.85	5.01	4.33	5.22	3.92	5.43	10.39	14.72	17.3	11.57	4.65	5.66	7.92
Lud	4.99	5.72	5.52	5.21	4.38	6.35	9.37	15.69	17.26	13.09	4.53	5.81	8.16
MR*	6.00	5.00	2.00	3.00	4.00	4.00	10.00	12.00	14.00	13.00	4.00	6.00	6.92
CR	12.29	9.59	3.14	4.79	3.97	5.53	11.30	12.87	16.18	13.19	4.68	4.96	8.54
NMB	9.17	7.43	2.45	4.01	4.43	5.56	10.18	14.00	17.27	14.09	3.83	5.07	8.12
BBN	8.18	6.33	2.82	3.78	4.41	5.91	10.89	16.71	16.7	14.28	5.42	5.18	8.38
Mean	8.71	6.35	3.28	4.09	3.89	5.83	10.20	13.96	16.54	12.99	4.18	5.27	

* Menominee River allelic richness was based on its actual sample size (N=35).

Table 7. Private allelic richness estimates based on rarefaction (Kalinowski 2005) for the second smallest sample size (N=70; Hog Island) for each population at each locus and total private allelic richness values for each locus and for all loci.

	B1	C6	C18	C68	C4	B2	C23	C49	C52	C41	C28	C45	Total
Epo	0.425	0.022	0.000	0.000	0.000	0.085	0.342	0.826	0.651	0.726	0.000	0.001	2.653
Nau	0.004	0.002	0.000	0.044	0.098	0.301	0.174	0.135	0.372	0.237	0.000	0.505	1.870
TB	0.004	0.000	0.317	0.000	0.347	0.018	0.000	0.000	1.261	0.053	0.609	0.000	2.604
ER	3.473	1.777	0.037	0.000	0.000	0.567	0.019	4.955	0.600	0.000	0.000	0.000	7.955
HgIs	0.298	0.371	0.000	0.000	0.000	0.022	0.009	0.004	0.415	0.747	0.244	0.593	2.405
Sau	0.004	0.064	0.102	0.013	0.000	0.098	0.000	0.041	0.005	0.000	0.000	0.005	0.326
Mus	0.000	0.000	0.021	0.661	0.000	0.132	0.151	0.179	0.267	0.000	0.000	0.035	1.446
Lud	0.000	0.020	0.375	0.018	0.058	0.210	0.000	1.050	0.684	0.229	0.003	0.048	2.693
MR*	0.001	0.000	0.053	0.580	0.000	0.000	0.181	0.264	0.324	0.000	0.003	0.098	1.504
CR	2.438	1.812	0.016	0.005	0.000	0.003	0.188	0.043	0.006	0.000	0.600	0.000	2.674
NMB	0.326	0.572	0.010	0.004	0.127	0.477	0.151	0.339	0.329	0.560	0.012	0.190	2.768
BBN	0.003	0.264	0.057	0.000	0.180	0.104	0.231	0.752	0.302	0.492	0.063	0.001	2.447
Total	6.976	4.904	0.988	1.325	0.810	2.017	1.446	8.588	5.216	3.044	1.534	1.476	31.345

* Menominee River allelic richness was based on its actual sample size (N=35).

Table 8. Mitochondrial DNA haplotypes and their percent occurrence, haplotypic diversity and nucleotide diversity for the ND 5 and D-loop gene regions.

Region	N	Number of Haplotypes	Variable sites versus haplotype A	Haplotypes	% Occurrence	*h*	π
ND 5	80	6	--	A	90%	0.188	0.00016
			18	B	1.20%		
			104	C	1.20%		
			139	D	1.20%		
			462	E	3.75%		
			465	F	2.50%		
D-Loop	85	7	--	A	85.90%	0.26	0.00025
			9	B	1.20%		
			18	C	5.90%		
			22	D	2.40%		
			155	E	1.20%		
			373	F	1.20%		
			478	G	2.40%		

Table 9. a) ND 5 haplotype distribution; b) D-loop haplotype distribution.

a)

Population	# of Haplotypes	Haplotypes Present
Big Bay de Noc	2	A,B
North-Moonlight Bay	3	A,C,E
Cedar River	1	A
Muskegon	2	A,E
Elk Rapids	3	A,E,F
Ludington	2	A,F
Naubinway	1	A
Saugatuck	1	A
Epoufette	2	A,D

b)

Population	# of Haplotypes	Haplotypes Present
Big Bay de Noc	2	A,G
North-Moonlight Bay	2	A,D
Cedar River	2	A,C
Muskegon	2	A,B
Elk Rapids	3	A,C,G
Ludington	3	A,C,E
Naubinway	3	A,C,F
Saugatuck	1	A
Epoufette	1	A

Table 10. Population pairwise comparison of allele frequency distributions (genic differentiation) across all loci.

	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
Epo	-											
Nau	0.3417	-										
TB	0.0001	<0.0001	-									
ER	<0.0001	<0.0001	<0.0001	-								
HgIs	0.0001	<0.0001	0.0001	<0.0001	-							
Sau	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	-						
Mus	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0009	-					
Lud	<0.0001	0.0001	<0.0001	<0.0001	<0.0001	0.0741	0.1426	-				
MR	0.5259	0.0588	0.0009	<0.0001	0.0473	<0.0001	<0.0001	<0.0001	-			
CR	0.0001	0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.3989	-		
NMB	<0.0001	0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.1029	0.0271	-	
BBN	<0.0001	0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0924	<0.0001	<0.0001	-

Table 11. Pairwise genetic distance matrix based on Cavalli-Sforza and Edwards (1967) chord distance D_c.

	BBN	CR	Epo	ER	TB	HgIs	Lud	MR	Mus	Nau	NMB	Sau
BBN	0											
CR	0.0799	0										
Epo	0.0907	0.0993	0									
ER	0.1983	0.219	0.2015	0								
TB	0.1392	0.1382	0.1261	0.2345	0							
HgIs	0.1302	0.1392	0.1281	0.2421	0.1407	0						
Lud	0.1197	0.1294	0.1218	0.2028	0.1516	0.1605	0					
MR	0.1279	0.1325	0.1367	0.2546	0.1574	0.1589	0.1648	0				
Mus	0.1137	0.1169	0.1216	0.2134	0.1558	0.1621	0.0828	0.1628	0			
Nau	0.0943	0.0959	0.0843	0.2182	0.1356	0.1278	0.1158	0.1392	0.1128	0		
NMB	0.0736	0.0687	0.0872	0.1971	0.1268	0.124	0.11	0.1148	0.1026	0.0902	0	
Sau	0.1185	0.1257	0.1221	0.2106	0.158	0.167	0.0914	0.1626	0.0945	0.1195	0.1064	0

Table 12. Analysis of molecular variance groupings, sum of squares, percent of variation, and p-values.

a)

4 Group AMOVA		Source of variation	Sum of Squares	% of variation	p-value
Group 1	Epoufette Naubinway	Among Groups	55.265	0.89183	0.00001
		Among Populations within Groups	58.504	1.08854	0.00001
Group 2	Traverse Bay Hog Island	Within Populations	5013.903	98.01962	0.00001
Group 3	Elk Rapids Ludington Muskegon Saugatuck				
Group 4	North-Moonlight Bay Big Bay de Noc Cedar River				

Table 12. Continued

b)

5 Group AMOVA		Source of variation	Sum of Squares	% of variation	p-value
Group 1	Epoufette	Among Groups	82.85	1.77618	0.00001
	Naubinway				
		Among Populations within Groups	30.919	0.36618	0.00001
Group 2	Traverse Bay				
	Hog Island	Within Populations	5013.903	97.85765	0.00001
Group 3	Ludington				
	Muskegon				
	Saugatuck				
Group 4	North-Moonlight Bay				
	Big Bay de Noc				
	Cedar River				
Group 5	Elk Rapids				

Table 12. Continued

c)

6 Group AMOVA		Source of variation	Sum of Squares	% of variation	p-value
Group 1	Epoufette	Among Groups	94.022	1.81731	0.00001
	Naubinway				
		Among Populations within groups	19.747	0.13583	0.07429
Group 2	Traverse Bay				
	Hog Island	Within Populations	5013.903	98.04686	0.00001
Group 3	Ludington				
	Muskegon				
	Saugatuck				
Group 4	North-Moonlight Bay				
	Cedar River				
Group 5	Elk Rapids				
Group 6	Big Bay de Noc				

Table 13. Pairwise F_{st} values (above diagonal) and their corresponding p-values (below diagonal).

	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
Epo	*	-0.00474	0.00469	0.03125	0.00177	0.01195	0.00977	0.00716	-0.00027	0.00597	0.00374	0.00414
Nau	0.9990	*	0.00608	0.03643	-0.00031	0.01508	0.00538	0.004	0.00049	0.00061	0.00364	0.0042
TB	0.0039	0.0010	*	0.03269	0.01015	0.2243	0.0202	0.00997	0.00703	0.00896	0.01736	0.01769
ER	0.0000	0.0000	0.0000	*	0.03788	0.03053	0.03233	0.02261	0.03744	0.02847	0.03442	0.03169
HgIs	0.1143	0.5547	0.0000	0.0000	*	0.02013	0.01756	0.00933	0.00814	0.01037	0.01593	0.0142
Sau	0.0000	0.0000	0.0000	0.0000	0.0000	*	-0.00081	-0.00441	0.01817	0.0118	0.01302	0.01381
Mus	0.0000	0.0010	0.0000	0.0000	0.0000	0.7588	*	-0.00197	0.01161	0.00635	0.00705	0.00834
Lud	0.0000	0.0039	0.0000	0.0000	0.0000	0.9990	0.9775	*	0.0073	0.00726	0.00562	0.00741
MR	0.4951	0.3799	0.0088	0.0000	0.1660	0.0000	0.0010	0.0068	*	-0.00163	0.00504	0.00681
CR	0.0000	0.2373	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6963	*	-0.00131	-0.00177
NMB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0186	0.9863	*	0.00396
BBN	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0049	0.9971	0.0000	*

Table 14. Effective population size estimates based on linkage disequilibrium.

Population	N_e	Lower 95% CI	Upper 95% CI
Epoufette	953	495	7768
Naubinway	416	292	697
Traverse Bay	378	214	1318
Elk Rapids	96	76	125
Hog Island	493	245	9351
Saugatuck	220	172	296
Muskegon	426	293	744
Ludington	262	197	380
Menominee River	167	92	685
Cedar River	595	390	1194
North-Moonlight Bay	707	557	951
Big Bay de Noc	1321	868	2630

Table 15. Variations of tests for Isolation by distance with Z, R, r^2 and p-values estimated from reduced major axis regression.

Populations	Comparison	Z-value	R-value	r^2	p-value	Significant
All	Gen v. Geo	1817.766	0.1687	0.0285	0.172	ns
	Gen v. Lg Geo	20.3874	0.1876	0.0352	0.113	ns
	Lg Gen v. Geo	-11022.4	0.2389	0.0571	0.117	ns
	Lg Gen v. Lg Geo	-127.458	0.2569	0.066	0.047	*
No Menominee River	Gen v. Geo	1454.649	0.1278	0.0163	0.197	ns
	Gen v. Lg Geo	16.4785	0.1457	0.169	0.831	ns
	Lg Gen v. Geo	-9227.02	0.2057	0.0423	0.138	ns
	Lg Gen v. Lg Geo	-107.259	0.2217	0.0492	0.079	ns
No Elk Rapids	Gen v. Geo	1393.594	0.218	0.218	0.04	*
	Gen v. Lg Geo	15.1595	0.202	0.202	0.001	**
	Lg Gen v. Geo	-9710.3	0.225	0.225	0.002	**
	Lg Gen v. Lg Geo	-11.4189	0.213	0.213	0.002	**

Table 16. Individual assignment testing to both population and stock of origin estimated using Genetic Mixture Analysis (Kalinowksi 2003).

Pop.	Stock	N	# correctly assigned to pop	# correctly assigned to stock	% correctly assigned to pop	% correctly assigned to stock	# assigned with at least 80% confidence to pop	# assigned with at least 80% confidence to stock	% correctly assigned to pop 80% confidence	% correctly assigned to stock 80% confidence
NMB	NMB	286	285	285	99.60%	99.60%	281	281	98.50%	98.50%
CR	NMB	149	149	149	100.00%	100.00%	147	148	98.60%	99.30%
MR	NMB	35	35	35	100.00%	100.00%	35	35	100.00%	100.00%
BBN	BBN	280	277	277	98.90%	98.90%	259	265	92.50%	94.60%
Nau	NOR	132	132	132	100.00%	100.00%	131	131	99.20%	99.20%
Epo	NOR	137	135	135	98.50%	98.50%	129	129	94.20%	94.20%
HgIs	NOE	70	70	70	100.00%	100.00%	70	70	100.00%	100.00%
TB	NOE	74	74	74	100.00%	100.00%	74	74	100.00%	100.00%
ER	EKR	72	72	72	100.00%	100.00%	72	72	100.00%	100.00%
Lud	SOE	124	124	124	100.00%	100.00%	124	124	100.00%	100.00%
Mus	SOE	133	132	132	99.20%	99.20%	129	129	97.00%	97.00%
Sau	SOE	132	132	132	100.00%	100.00%	132	132	100.00%	100.00%
Totals		1624	1617	1617	99.60%	99.60%	1583	1590	97.50%	97.90%

Table 17. Proportion of fish sampled at each population/sample site that assigned to the suggested genetic stocks based on genetic mixture analysis (Kalinowski 2003).

Population	BBN	NMB	NOR	NOE	EKR	SOE
BBN	93.64	5.75	0	0.06	0.24	0.31
NMB	0	98.45	0	1.55	0	0
CR	0	99.21	0	0.79	0	0
MR	0	100	0	0	0	0
Nau	0	0	99.03	0.97	0	0
Epo	0	0.47	95.04	3.20	1.29	0
HgIs	0	0	0	100	0	0
GT	0	0	0	100	0	0
ER	0	0	0	0	100	0
Lud	0	0	0	0	0	100
Mus	0	0	0	1.60	0	98.40
Sau	0	0	0	0.04	0	99.96

Table 18. Proportional assignment of simulated commercial harvests using GMA (Kalinowski 2003).

a)

Simulation #1							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	0%	Epo	0.11%	BBN	25%	BBN	25.84%
Nau	0%	Nau	0.07%	NMB	75%	NMB	73.99%
TB	0%	TB	0.00%	NOR	0%	NOR	0.17%
ER	0%	ER	0.00%	NOE	0%	NOE	0.00%
HgIs	0%	HgIs	0.00%	EKR	0%	EKR	0.00%
Sau	0%	Sau	0.00%	SOE	0%	SOE	0.00%
Mus	0%	Mus	0.00%				
Lud	0%	Lud	0.00%				
MR	5%	MR	4.92%				
CR	20%	CR	14.11%				
NMB	50%	NMB	54.97%				
BBN	25%	BBN	25.84%				

b)

Simulation #2							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	10%	Epo	13.85%	BBN	10%	BBN	14.22%
Nau	10%	Nau	8.73%	NMB	40%	NMB	35.27%
TB	10%	TB	8.33%	NOR	20%	NOR	22.58%
ER	10%	ER	10.68%	NOE	20%	NOE	17.24%
HgIs	10%	HgIs	8.91%	EKR	10%	EKR	10.68%
Sau	0%	Sau	0.00%	SOE	0%	SOE	0.00%
Mus	0%	Mus	0.00%				
Lud	0%	Lud	0.00%				
MR	10%	MR	8.77%				
CR	10%	CR	10.64%				
NMB	20%	NMB	15.86%				
BBN	10%	BBN	14.22%				

Table 18. Continued.

c)

Simulation #3							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	10%	Epo	11.58%	BBN	0%	BBN	0.21%
Nau	10%	Nau	8.60%	NMB	0%	NMB	0.29%
TB	10%	TB	10.71%	NOR	20%	NOR	20.18%
ER	10%	ER	11.59%	NOE	20%	NOE	23.96%
HgIs	10%	HgIs	13.25%	EKR	10%	EKR	11.59%
Sau	10%	Sau	9.29%	SOE	50%	SOE	43.77%
Mus	10%	Mus	9.58%				
Lud	30%	Lud	24.90%				
MR	0%	MR	0.00%				
CR	0%	CR	0.00%				
NMB	0%	NMB	0.29%				
BBN	0%	BBN	0.21%				

d)

Simulation #4							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	60%	Epo	59.93%	BBN	0%	BBN	0.82%
Nau	10%	Nau	8.94%	NMB	0%	NMB	0.09%
TB	5%	TB	4.48%	NOR	70%	NOR	68.86%
ER	5%	ER	4.23%	NOE	25%	NOE	25.98%
HgIs	20%	HgIs	21.50%	EKR	5%	EKR	4.23%
Sau	0%	Sau	0.02%	SOE	0%	SOE	0.02%
Mus	0%	Mus	0.00%				
Lud	0%	Lud	0.00%				
MR	0%	MR	0.00%				
CR	0%	CR	0.09%				
NMB	0%	NMB	0.00%				
BBN	0%	BBN	0.82%				

Table 18. Continued.

e)

Simulation #5							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	0%	Epo	0.00%	BBN	0%	BBN	0.00%
Nau	0%	Nau	0.00%	NMB	0%	NMB	0.00%
TB	0%	TB	0.15%	NOR	0%	NOR	0.00%
ER	0%	ER	0.00%	NOE	0%	NOE	0.49%
HgIs	0%	HgIs	0.34%	EKR	0%	EKR	0.00%
Sau	0%	Sau	0.00%	SOE	100%	SOE	99.51%
Mus	100%	Mus	99.51%				
Lud	0%	Lud	0.00%				
MR	0%	MR	0.00%				
CR	0%	CR	0.00%				
NMB	0%	NMB	0.00%				
BBN	0%	BBN	0.00%				

f)

Simulation #6							
Pop.	% Catch	Pop.	% Assigned	Stock	% Catch	Stock	% Assigned
Epo	0%	Epo	0.27%	BBN	50%	BBN	51.61%
Nau	0%	Nau	0.00%	NMB	50%	NMB	45.57%
TB	0%	TB	1.28%	NOR	0%	NOR	0.27%
ER	0%	ER	0.00%	NOE	0%	NOE	1.28%
HgIs	0%	HgIs	0.00%	EKR	0%	EKR	0.00%
Sau	0%	Sau	1.27%	SOE	0%	SOE	1.27%
Mus	0%	Mus	0.00%				
Lud	0%	Lud	0.00%				
MR	0%	MR	0.00%				
CR	0%	CR	0.00%				
NMB	50%	NMB	45.57%				
BBN	50%	BBN	51.61%				

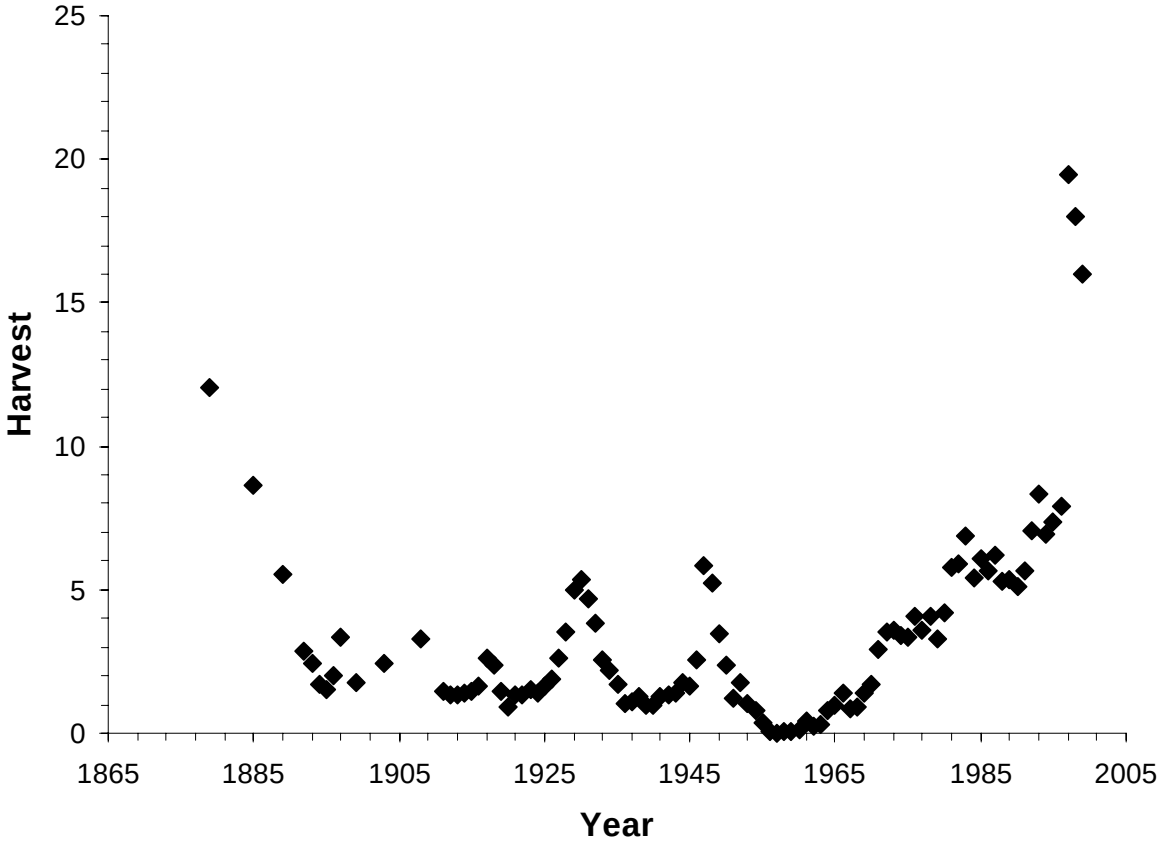


Figure 1. Historical Lake Michigan lake whitefish commercial harvest (Baldwin et al. 2002; from Great Lakes Fishery Commission database).

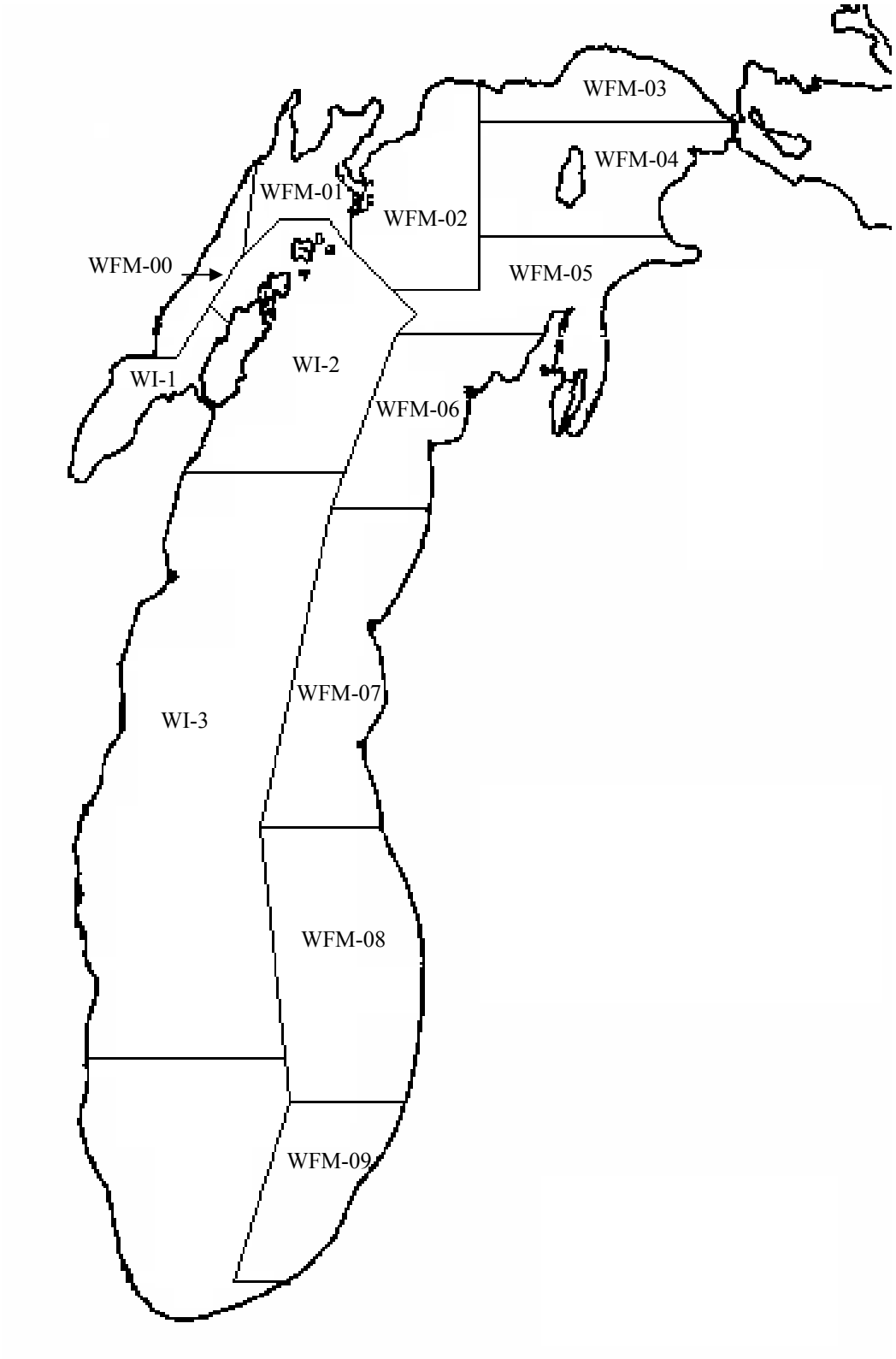


Figure 2. Lake whitefish commercial management zones. Zones in Wisconsin waters include WI-1, WI-2, and WI-3. All other zones are in Michigan waters (WFMs) and were originally established by the 1836 Consent Decree.

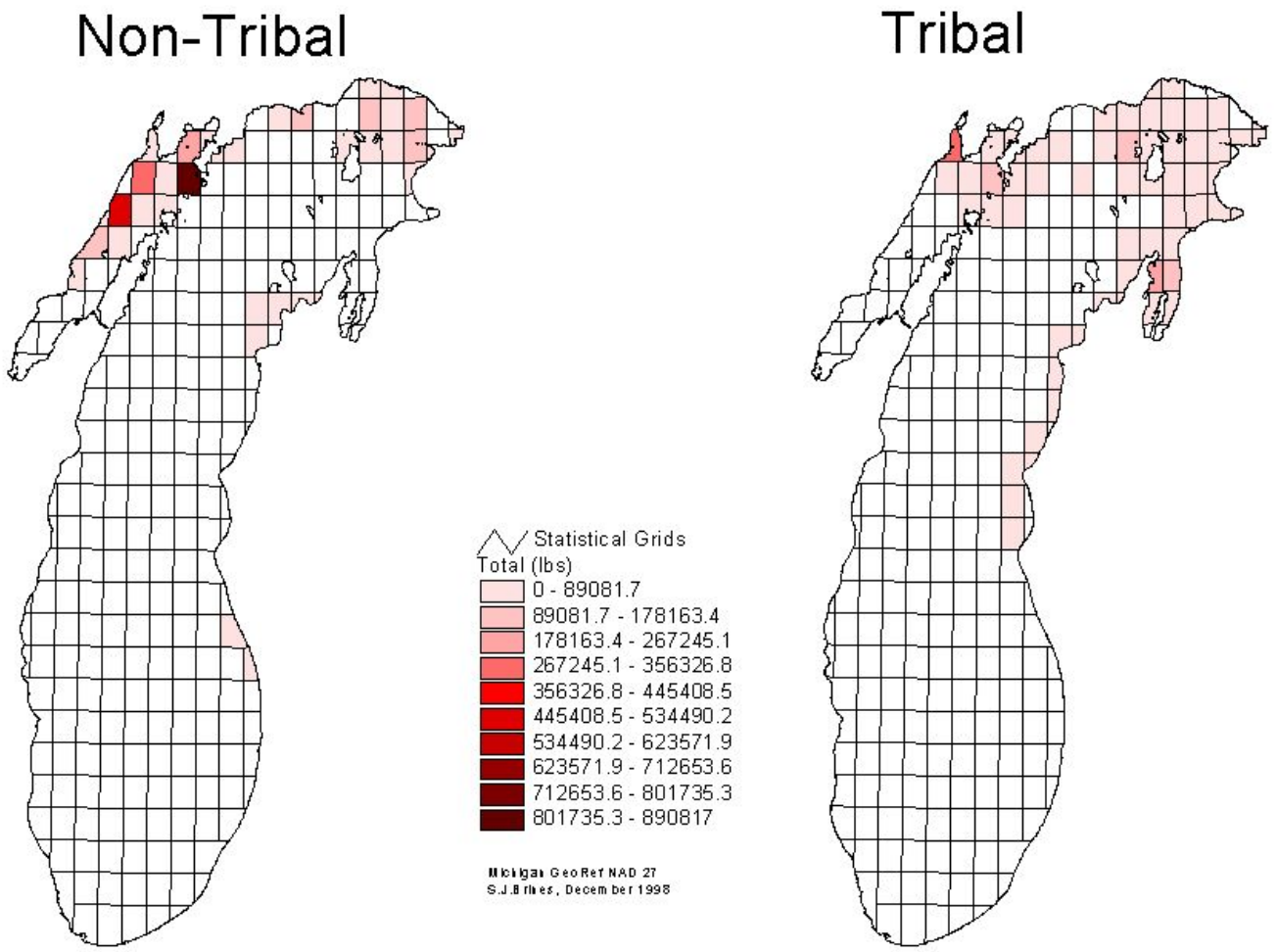


Figure 3. Non-tribal and Tribal commercial harvest of lake whitefish from Michigan waters, reported by statistical grids, Lake Michigan, 1985.

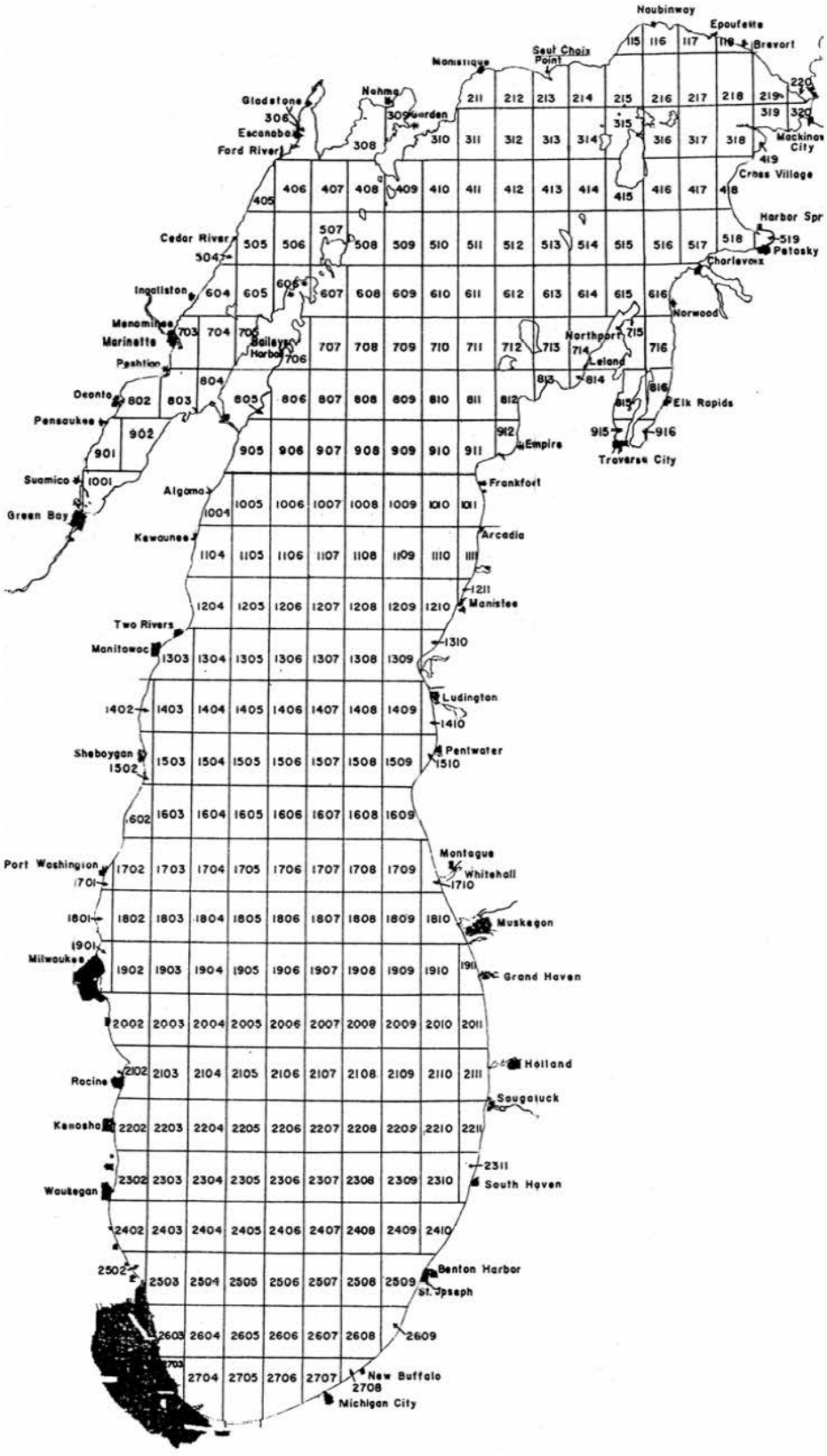


Figure 4. Lake Michigan commercial fishing grids.

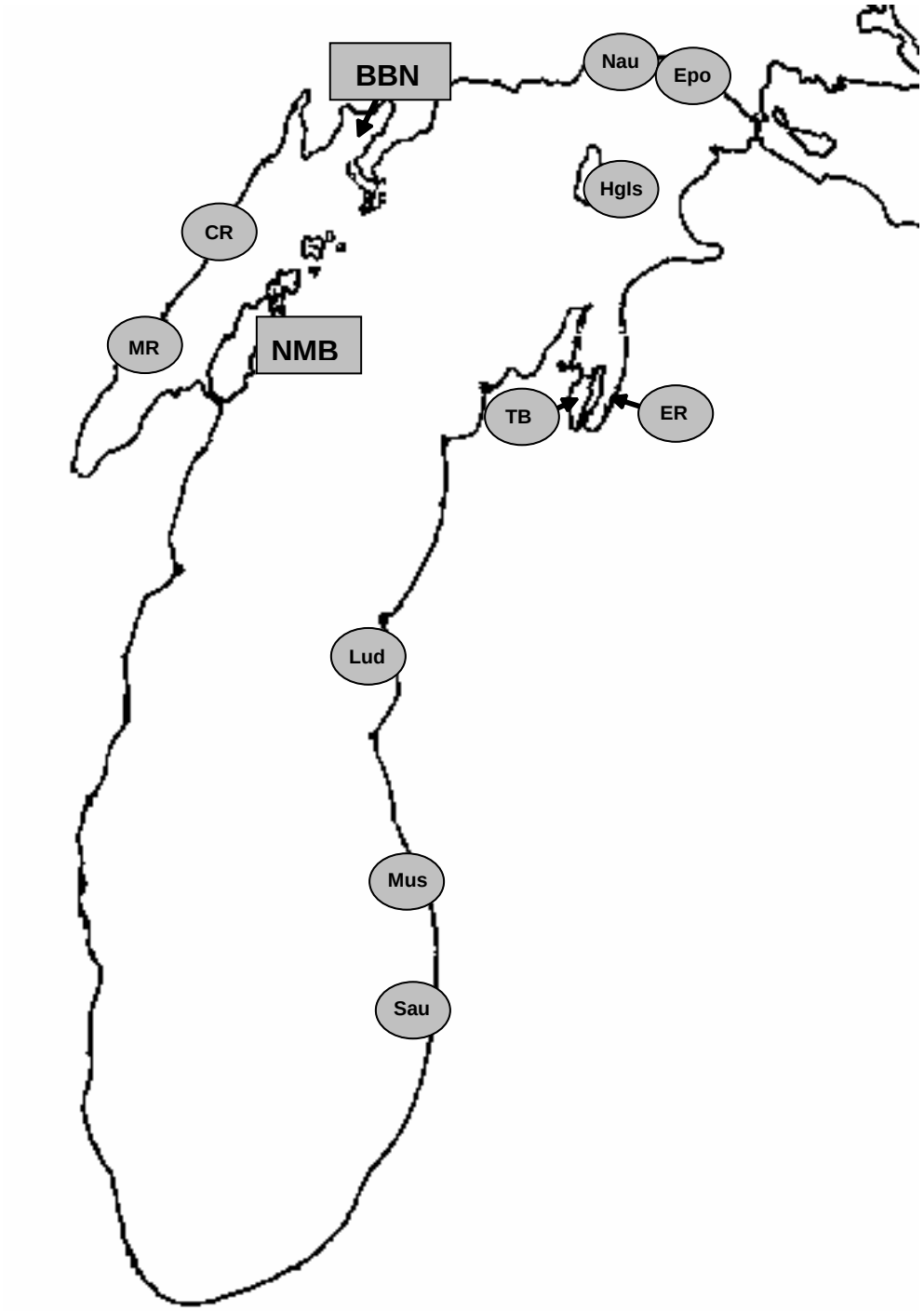


Figure 5. Primary and secondary study sites for Lake Michigan, lake whitefish study. Primary sites (denoted by gray boxes) include the lakeside of the Door County Peninsula (WI; NMB) and Big Bay de Noc (BBN) Michigan. Secondary study sites (denoted by gray circles) include Lower Menominee River (MR), Cedar River (CR), Naubinway, (Nau), Epoufette (Epo), Hog Island (HgIs), Traverse Bay (TB), Elk Rapids (ER), Ludington (Lud), Muskegon (Mus) and Saugatuck (Sau).

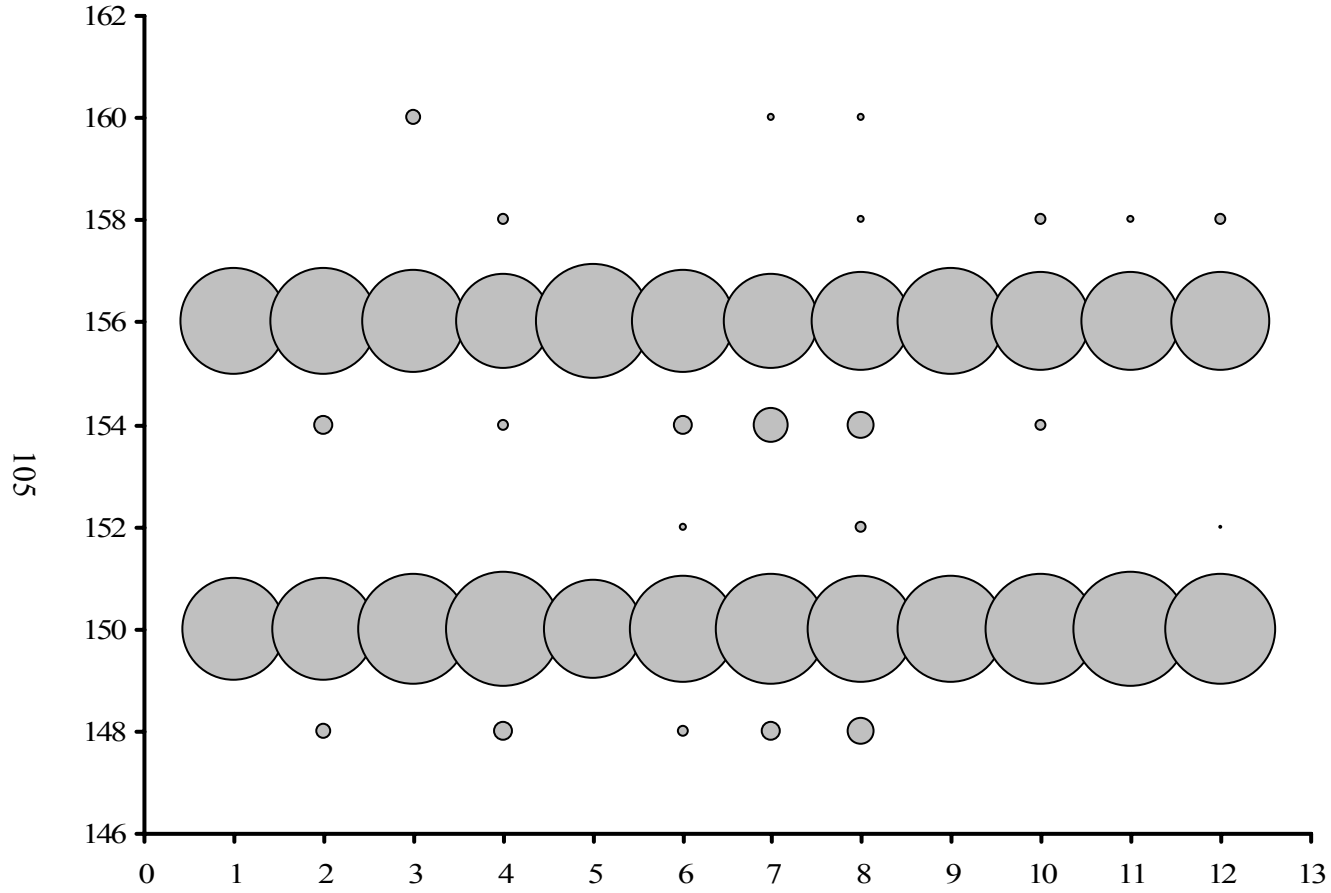


Figure 6. Allele frequency distribution for locus Cocl-lav 18. X-axis numbers correspond to populations (1 = Epoufette, 2 = Naubinway, 3 = Traverse Bay, 4 = Elk Rapids, 5 = Hog Island, 6 = Saugatuck, 7 = Muskegon, 8 = Ludington, 9 = Menominee River, 10 = Cedar River, 11 = NMB, 12 = BBN. Y-axis numbers correspond to allele size in base pairs. The size of each circle represents the frequency of that allele in each population.

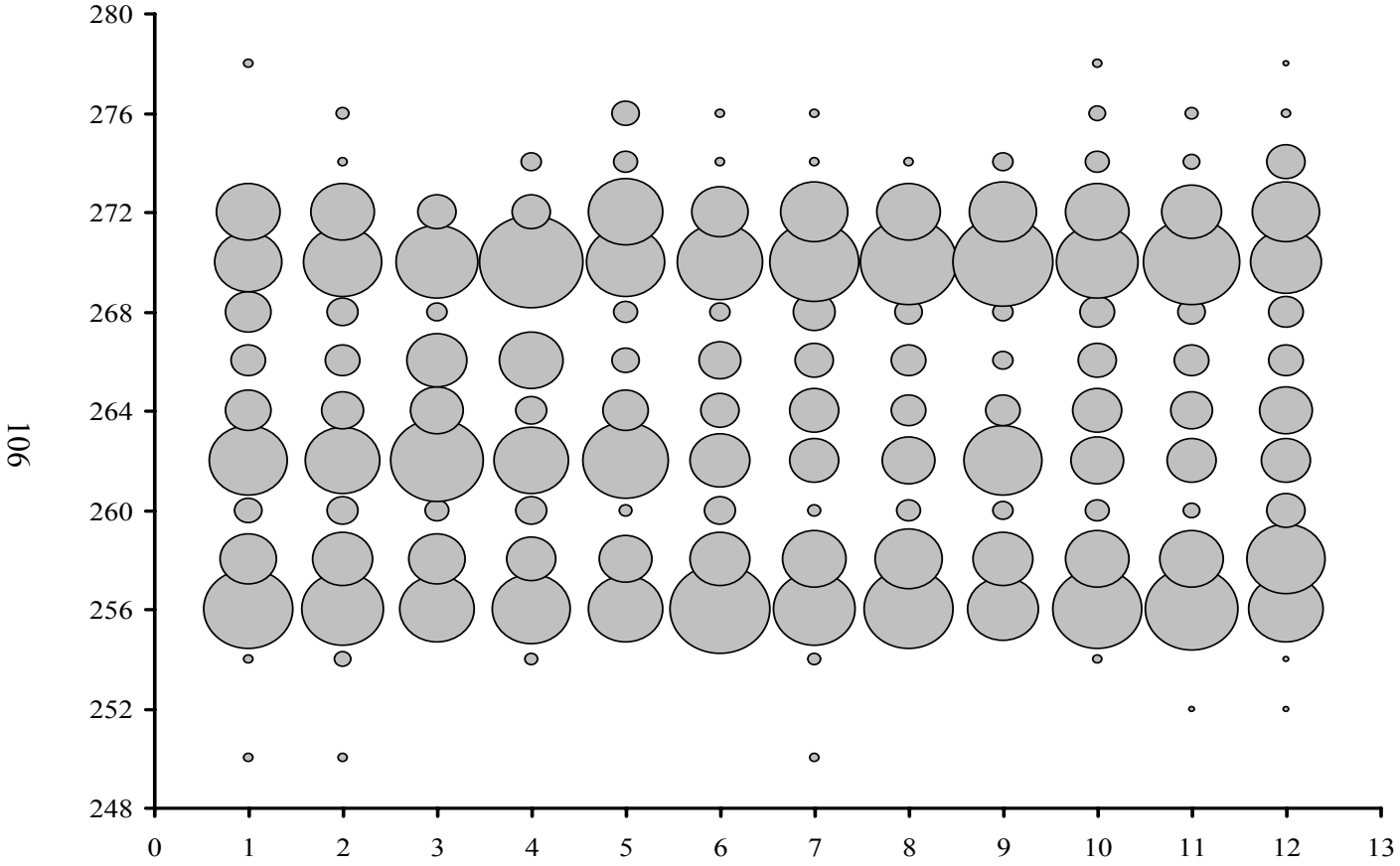


Figure 7. Allele frequency distribution for locus Cocl-23. X-axis numbers correspond to populations (1 = Epoufette, 2 = Naubinway, 3 = Traverse Bay, 4 = Elk Rapids, 5 = Hog Island, 6 = Saugatuck, 7 = Muskegon, 8 = Ludington, 9 = Menominee River, 10 = Cedar River, 11 = NMB, 12 = BBN. Y-axis numbers correspond to allele size in base pairs. The size of each circle represents the frequency of that allele in each population.

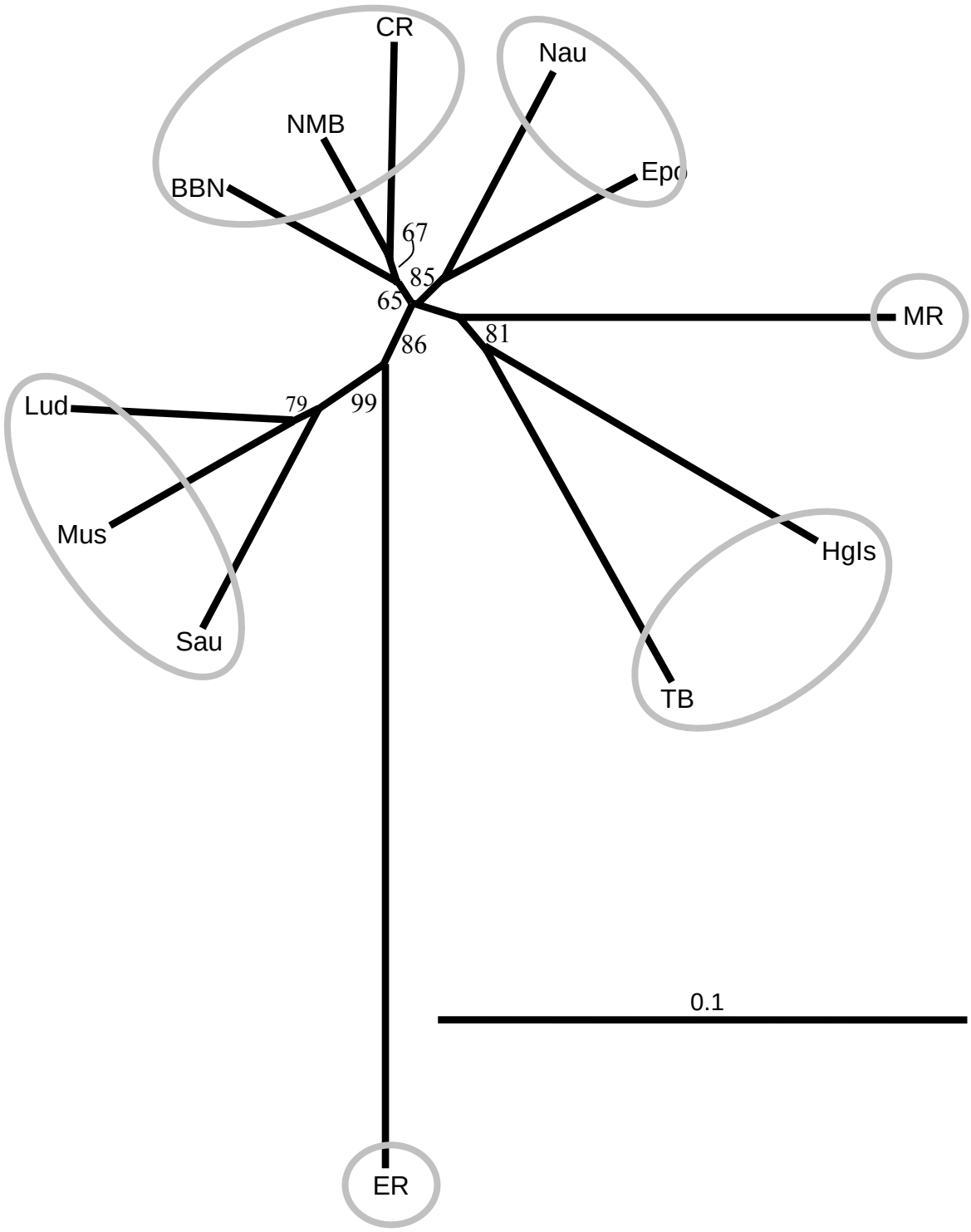


Figure 8. Neighbor-joining tree based on Cavalli-Sforza and Edwards (1967) chord distance (D_c). Branch support represents the percent recovery of that node based on 5,000 bootstrap pseudoreplicates. Support ≤50% is not shown. Most likely genetic groupings of populations are indicated by gray ovals.

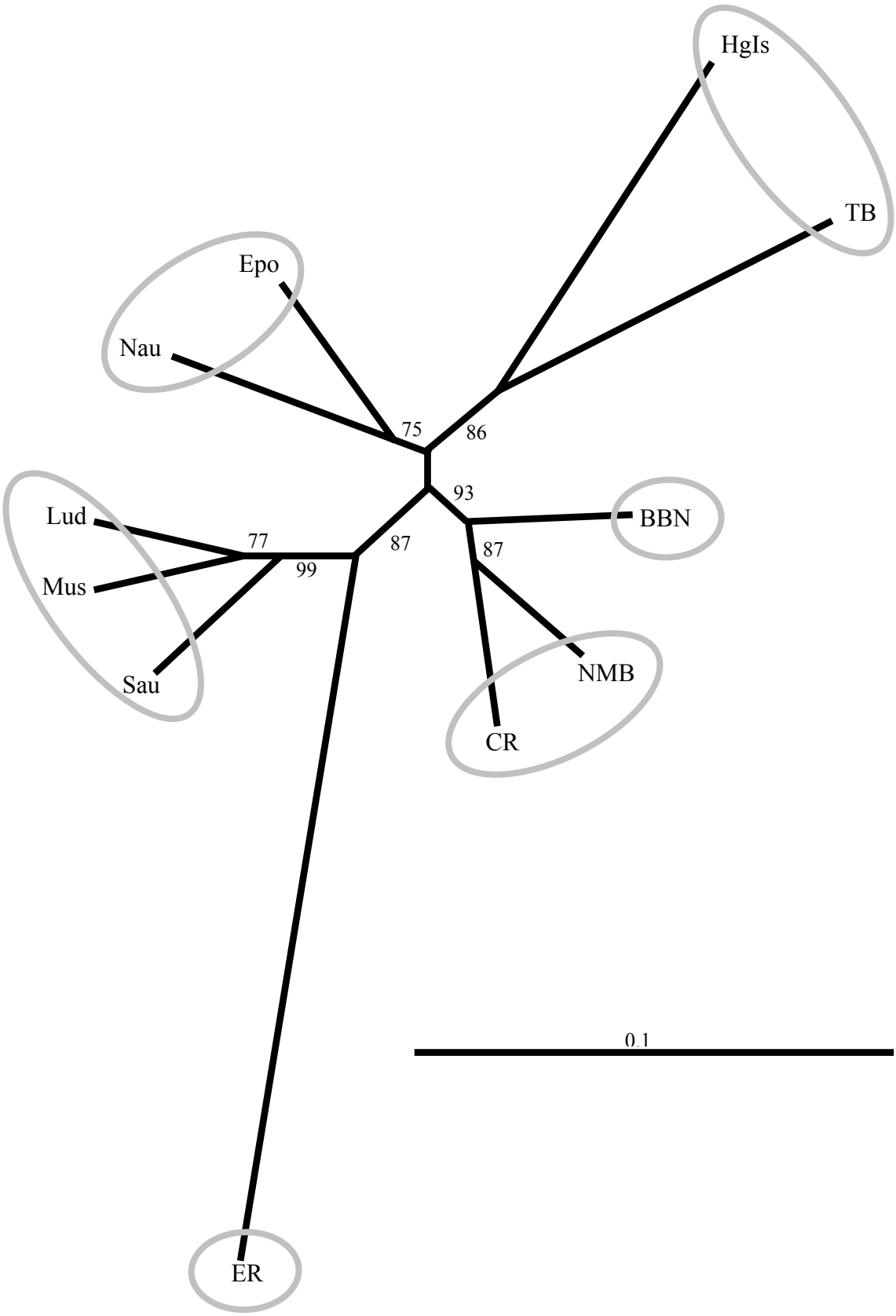
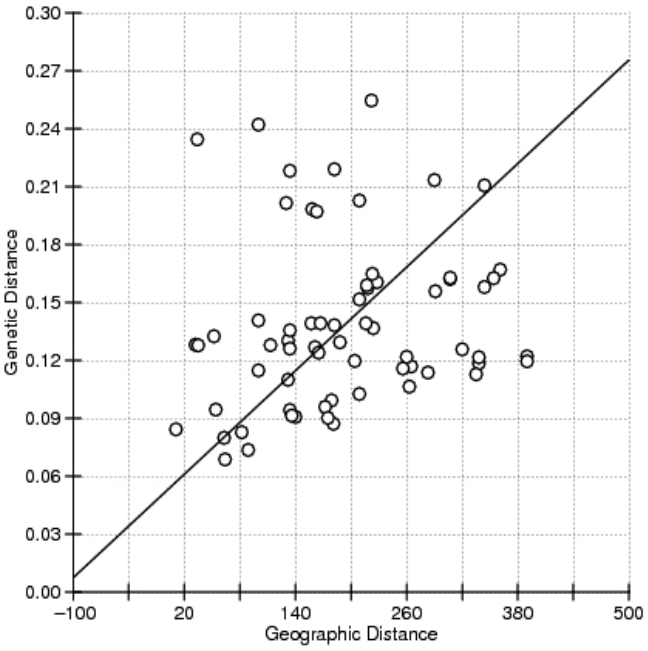


Figure 9. Neighbor-joining tree based on Cavalli-Sforza and Edwards (1967) chord distance (D_c) without the Menominee River sample. Branch support represents the percent recovery of that node based on 5,000 bootstrap pseudoreplicates. Support ≤50% is not shown. Most likely genetic groupings of populations are indicated by gray ovals.

a)



b)

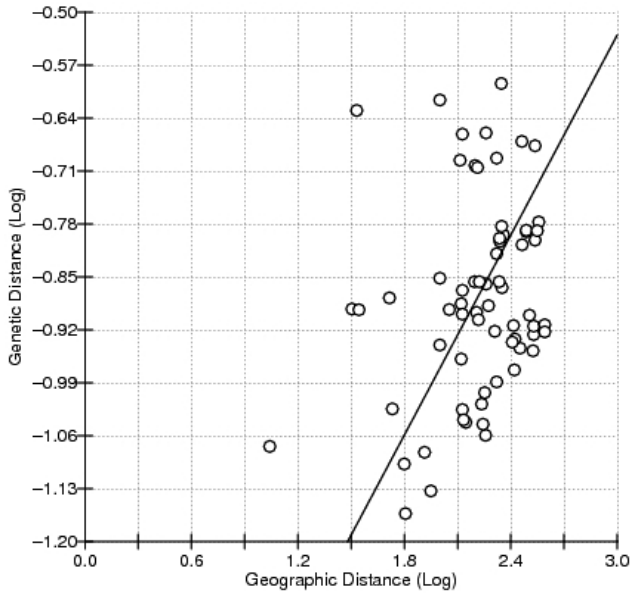
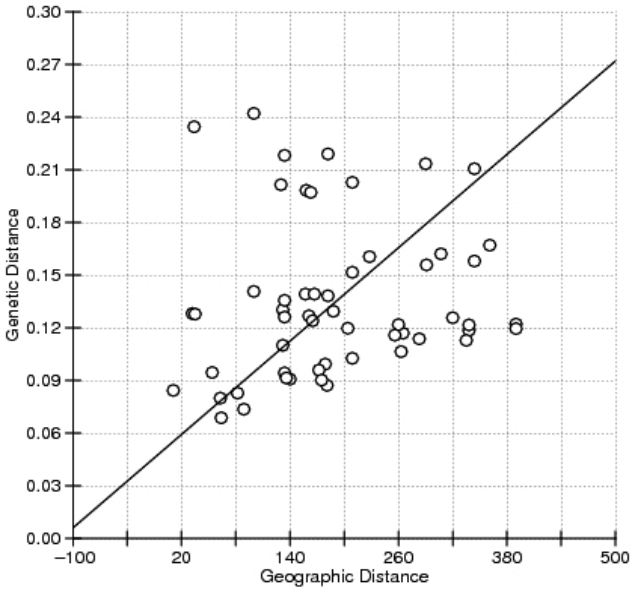


Figure 10. Isolation by distance for all populations. Figure 10a is genetic (D_c) distance against geographic distance (kms). Figure 10b is the log genetic distance versus log geographic distance. Regression line was calculated using reduced major axis regression.

a)



b)

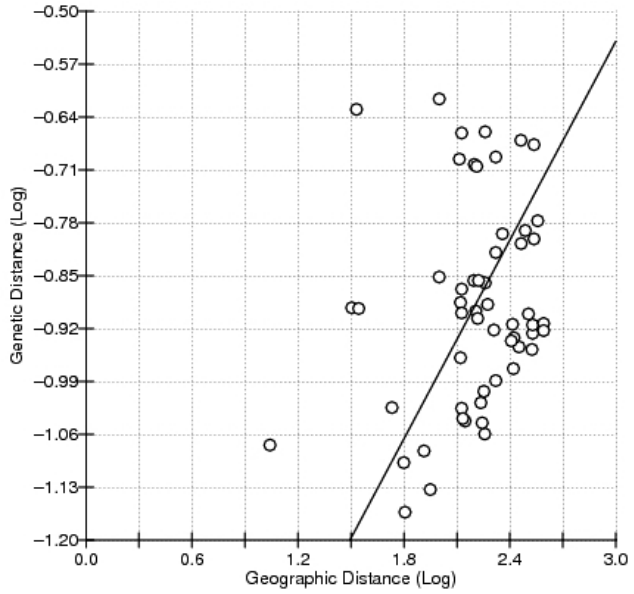
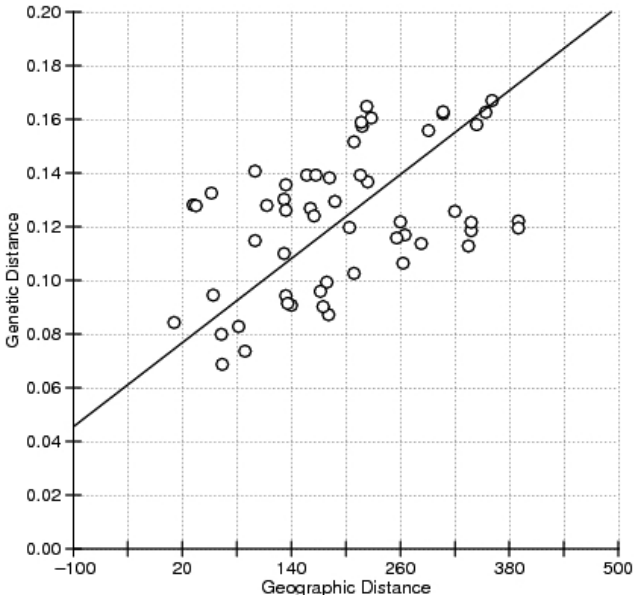


Figure 11. Isolation by distance without the Menominee River population to determine potential influences from an admixture population. Figure 11a is genetic (D_c) distance versus geographic (kms) distance. Figure 11b is log genetic versus log geographic distance. Regression line was calculated using reduced major axis regression.

a)



b)

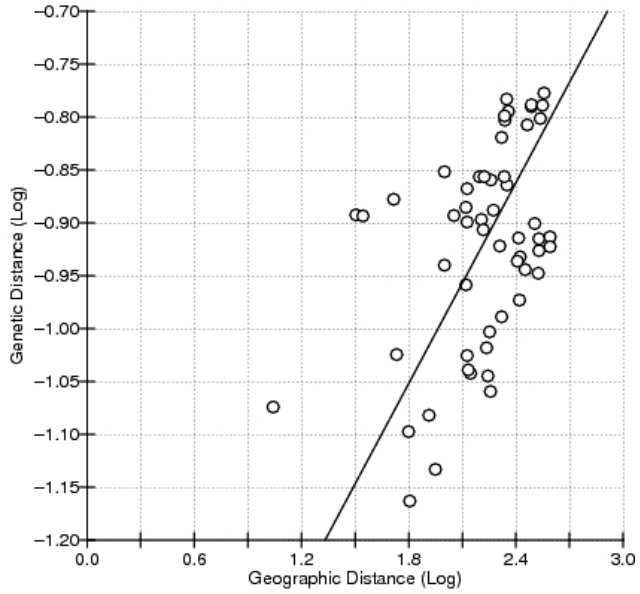


Figure 12. Isolation by distance without the Elk Rapids population. Figure 12a is genetic (D_c) distance versus geographic (kms) distance. Figure 12b is log genetic versus log geographic distance. Regression line was calculated using reduced major axis regression.

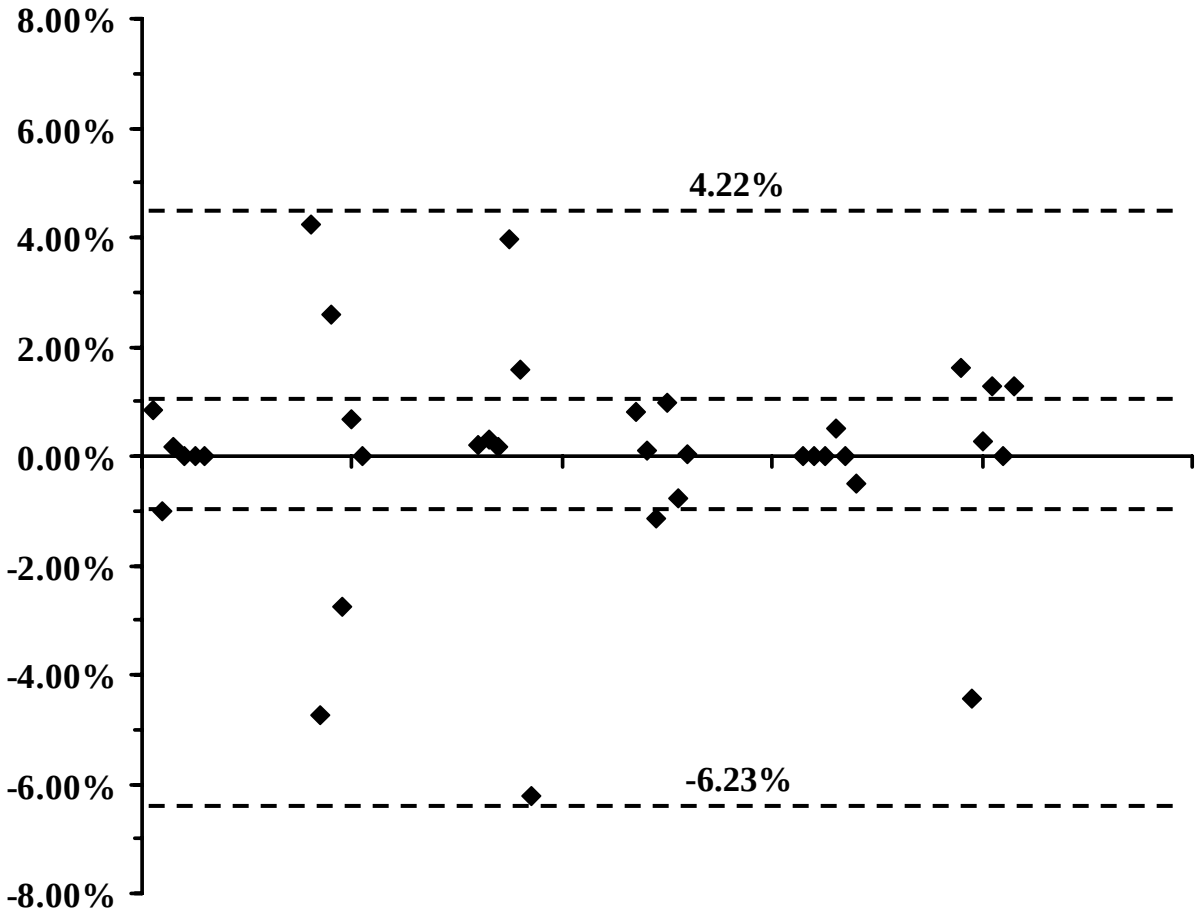


Figure 13. Plot of residuals from simulated genetic mixture analysis. Outside dashed lines represent the largest residual values. Inside dashed lines represent a 1% error mark. Twenty-two of 36 comparisons yielded an error of ≤ 1%.

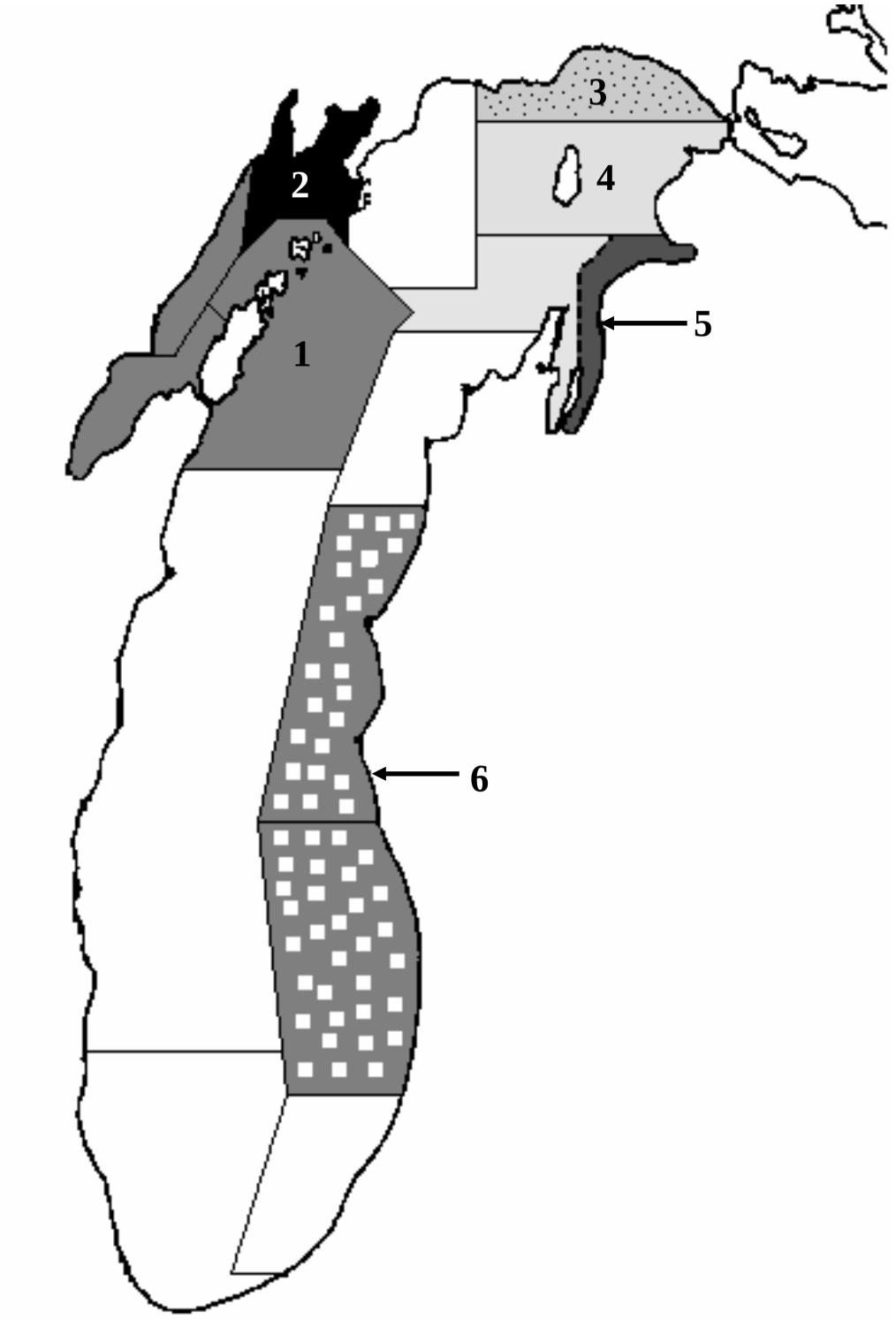


Figure 14. Six genetic management zones (GMZs) based on genetic stock identification and spawning site locations, over-layed on the statistical and commercial management zones currently in place for Lake Michigan.

Appendix 1. Allele frequencies for all loci for each population.

Locus/ Alleles	Populations											
BWF-1	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
169	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.00
179	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.00
181	0.00	0.00	0.00	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
183	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00
189	0.39	0.00	0.00	0.79	0.00	0.00	0.00	0.00	0.00	1.06	0.90	0.00
191	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.41	0.00	0.00
193	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06	0.54	0.00
201	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.00
203	0.78	0.39	0.00	0.00	0.75	0.40	0.00	0.00	0.00	0.70	0.36	0.36
205	7.03	6.98	2.11	3.17	8.21	6.05	7.52	6.10	2.86	4.23	3.07	3.47
207	0.39	2.33	3.52	3.17	2.24	0.00	0.00	0.00	0.00	0.00	0.18	0.18
209	0.39	0.00	2.11	0.79	0.75	0.40	0.75	0.00	0.00	0.70	1.08	1.09
211	2.73	1.55	1.41	0.00	5.22	1.21	1.50	0.00	2.86	2.11	1.81	3.10
213	6.25	6.98	4.93	7.14	8.96	8.06	4.51	8.13	8.57	9.51	10.83	9.31
215	42.97	36.82	31.69	42.86	35.82	51.21	43.61	54.07	31.43	35.92	44.95	37.77
217	31.64	39.53	47.18	30.16	27.61	32.66	37.59	28.05	50.00	34.15	33.39	34.12
219	0.78	2.33	1.41	1.59	4.48	0.00	0.38	0.00	4.29	3.87	1.44	6.39
221	5.08	2.71	4.23	4.76	3.73	0.00	4.14	3.66	0.00	3.17	0.72	4.20
223	0.78	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00
227	0.39	0.39	1.41	0.00	1.49	0.00	0.00	0.00	0.00	0.70	0.54	0.00
239	0.00	0.00	0.00	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
241	0.00	0.00	0.00	3.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
243	0.00	0.00	0.00	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 6	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
104	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00
108	0.00	0.00	0.00	1.49	0.00	0.00	0.00	0.00	0.00	1.35	0.70	0.00
110	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00	0.00
114	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.36	0.00	0.00
116	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00
128	0.00	0.00	0.00	1.49	0.00	0.00	0.00	0.81	0.00	0.34	0.18	0.00
130	0.37	0.00	0.00	0.00	0.00	0.78	0.00	0.00	0.00	0.34	0.18	1.07
132	75.37	75.95	61.27	55.22	75.00	71.71	75.76	64.11	71.43	59.12	69.89	67.68
134	2.57	1.15	2.11	8.21	0.00	4.26	7.58	4.84	4.29	3.72	3.52	3.75
136	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.18	0.36
138	12.50	12.60	15.49	16.42	16.18	15.12	10.61	14.52	20.00	18.58	15.85	20.54
140	1.84	1.91	0.00	0.75	0.00	0.39	0.00	0.00	1.43	0.68	1.76	1.43
142	0.37	1.53	0.00	0.75	0.74	0.00	0.38	0.40	0.00	2.70	0.88	0.36
144	0.00	0.76	0.00	0.00	2.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00
146	6.99	6.11	19.72	7.46	4.41	7.75	4.92	14.52	2.86	9.12	5.99	4.64
148	0.00	0.00	1.41	6.72	0.74	0.00	0.76	0.81	0.00	0.68	0.35	0.18
150	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.18	0.00
158	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
164	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 18	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
148	0.00	1.14	0.00	2.21	0.00	0.76	2.26	3.39	0.00	0.00	0.00	0.00
150	48.11	46.21	53.38	55.88	42.14	50.38	52.26	48.73	50.00	55.07	56.69	54.48
152	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.85	0.00	0.00	0.00	0.18
154	0.00	2.27	0.00	0.74	0.00	1.89	5.64	3.39	0.00	0.68	0.00	0.00
156	51.89	50.38	45.27	40.44	57.86	46.59	39.47	42.80	50.00	43.58	42.78	44.44
158	0.00	0.00	0.00	0.74	0.00	0.00	0.00	0.42	0.00	0.68	0.53	0.90
160	0.00	0.00	1.35	0.00	0.00	0.00	0.38	0.42	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 68	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
172	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00
174	2.69	2.65	6.08	0.74	1.49	5.34	4.89	4.35	1.47	4.36	2.82	3.26
176	17.69	15.15	18.92	19.85	17.16	35.11	30.08	28.70	17.65	18.12	18.66	20.65
178	0.00	0.38	0.00	0.00	0.00	0.38	3.01	0.87	0.00	0.67	0.18	0.00
180	76.15	77.27	75.00	79.41	80.60	58.02	60.53	63.91	80.88	74.83	77.29	74.64
182	3.46	3.03	0.00	0.00	0.75	0.38	0.75	1.30	0.00	1.68	0.70	1.45
184	0.00	1.52	0.00	0.00	0.00	0.76	0.00	0.87	0.00	0.34	0.35	0.00

Locus/ Alleles	Populations											
Cocl-lav 4	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
148	1.52	3.41	0.00	0.00	0.00	0.76	1.92	2.12	1.43	2.72	1.06	1.08
150	46.97	45.08	42.57	38.41	45.00	34.09	37.31	41.10	55.71	42.52	45.05	40.14
152	31.44	33.71	29.05	28.99	33.57	36.36	40.00	36.44	25.71	34.01	32.16	35.66
154	20.08	17.42	27.70	32.61	21.43	28.79	20.77	19.92	17.14	20.75	20.85	22.04
156	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.71	1.08
158	0.00	0.38	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00

Locus/ Alleles	Populations											
BWF-2	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
140	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.18
142	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37	0.00
146	7.20	8.27	8.78	5.88	7.14	8.27	5.14	7.14	12.86	7.48	9.74	6.86
148	1.14	1.57	3.38	0.00	2.14	2.76	3.74	1.26	0.00	3.40	2.02	0.72
150	0.76	0.00	1.35	0.00	1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.36
152	0.38	0.39	0.00	4.90	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00
154	75.00	65.35	72.30	76.47	62.86	69.69	72.43	72.27	71.43	69.73	68.20	71.84
156	4.17	5.91	0.00	5.88	0.71	11.02	2.34	6.72	0.00	1.02	0.74	1.62
158	0.76	1.18	0.68	0.00	0.00	0.00	0.00	0.00	0.00	1.36	0.18	1.44
160	0.00	0.00	0.00	0.98	0.00	0.79	0.00	0.00	0.00	0.00	0.00	0.00
162	9.85	16.54	13.51	5.88	25.71	7.09	15.89	10.92	14.29	17.01	18.20	16.79
164	0.76	0.39	0.00	0.00	0.00	0.00	0.00	0.84	0.00	0.00	0.18	0.18
166	0.00	0.00	0.00	0.00	0.00	0.39	0.47	0.42	1.43	0.00	0.18	0.00

Locus/ Alleles	Populations											
Cocl-23	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
250	0.38	0.39	0.00	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-23	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
252	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.19
254	0.38	1.18	0.00	0.75	0.00	0.00	0.75	0.00	0.00	0.36	0.00	0.19
256	23.11	20.08	16.22	18.66	17.14	29.23	20.68	24.07	15.71	23.36	25.18	17.11
258	9.85	10.63	10.14	7.46	8.57	11.15	13.16	14.35	11.43	13.14	12.41	17.86
260	2.65	3.15	2.03	2.99	0.71	3.46	0.75	1.85	1.43	1.82	1.28	4.70
262	19.32	17.72	25.00	17.16	22.86	11.15	7.52	8.33	18.57	8.76	8.03	8.27
264	7.20	5.51	8.78	2.99	6.43	5.00	7.52	3.70	4.29	7.30	5.84	8.65
266	3.79	4.33	10.81	12.69	2.86	6.15	4.51	4.17	1.43	4.74	3.83	3.76
268	6.44	3.54	1.35	0.00	2.14	1.54	5.64	2.78	1.43	4.01	2.92	3.95
270	14.02	19.29	20.95	31.34	17.86	21.54	23.68	27.31	30.00	21.17	26.82	16.17
272	12.50	12.99	4.73	4.48	16.43	10.00	14.66	12.96	14.29	12.04	11.68	13.91
274	0.00	0.39	0.00	1.49	2.14	0.38	0.38	0.46	1.43	1.82	1.09	4.70
276	0.00	0.79	0.00	0.00	2.86	0.38	0.38	0.00	0.00	1.09	0.73	0.38
278	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.00	0.19

Locus/ Alleles	Populations											
Cocl-lav 49	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
163	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37
167	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.19
169	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19
171	0.92	1.92	0.00	0.00	1.43	0.00	0.00	0.00	0.00	2.61	1.44	2.04
173	1.83	0.77	0.00	0.00	0.00	0.00	1.55	0.42	0.00	0.87	0.36	1.12
175	36.24	33.08	43.92	0.00	41.43	25.20	31.40	29.41	34.29	39.13	38.85	37.92
177	0.00	0.00	0.00	0.00	0.00	1.60	1.55	1.26	0.00	1.30	0.72	0.37
179	3.67	3.08	2.03	0.00	0.00	3.60	2.33	1.26	7.14	2.61	2.70	3.53
181	0.00	0.38	0.00	0.00	0.00	0.40	0.39	0.00	1.43	0.43	0.36	0.37
183	0.92	3.08	0.68	0.00	4.29	0.40	0.39	0.84	4.29	3.04	1.44	3.16
185	0.00	0.00	0.00	0.78	0.00	0.00	0.00	0.84	0.00	0.00	0.54	0.93
187	10.55	9.62	14.19	0.00	3.57	16.80	11.63	16.39	15.71	15.22	12.77	10.59
189	0.46	1.15	0.00	0.78	0.00	0.00	0.39	0.00	2.86	0.00	0.00	0.93
191	1.38	0.00	8.11	0.00	7.86	0.80	1.55	0.00	1.43	3.04	1.98	1.67
193	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.26	0.00	0.00	0.18	0.19
195	0.00	0.00	1.35	0.78	0.00	4.40	2.71	4.20	0.00	0.00	0.36	0.19
197	0.46	0.38	0.00	1.56	0.00	0.40	0.00	0.42	0.00	0.00	0.00	0.19
199	2.75	0.00	0.00	42.19	0.00	1.20	0.39	1.68	0.00	0.00	0.72	1.67
201	0.92	3.08	1.35	1.56	5.71	4.80	8.53	8.40	5.71	1.30	1.44	1.30
203	0.46	0.38	0.68	2.34	2.14	0.80	0.00	0.00	1.43	1.30	1.08	1.30
205	0.46	1.15	0.00	0.78	0.00	2.00	1.55	1.26	0.00	0.43	0.90	0.74
207	9.63	10.00	8.11	0.00	7.86	6.80	6.59	7.98	5.71	8.70	12.23	9.85
209	20.18	29.62	19.59	0.00	24.29	26.00	23.26	18.49	18.57	19.13	20.14	15.24
211	2.29	0.38	0.00	14.06	0.00	3.20	2.33	2.94	0.00	0.43	0.36	4.65
213	1.38	0.77	0.00	0.78	0.71	1.20	1.55	0.84	1.43	0.43	0.36	0.56
215	0.92	0.38	0.00	0.78	0.71	0.40	0.00	0.00	0.00	0.00	0.54	0.74
217	0.00	0.00	0.00	0.00	0.00	0.00	1.94	1.26	0.00	0.00	0.18	0.00
223	0.00	0.00	0.00	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
225	0.00	0.00	0.00	3.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
227	0.00	0.00	0.00	3.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 49	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
229	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
231	0.00	0.38	0.00	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
233	1.83	0.38	0.00	8.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
235	1.38	0.00	0.00	14.06	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00
237	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00
241	0.46	0.00	0.00	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
243	0.00	0.00	0.00	1.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 52	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
90	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
92	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00
94	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.18
102	9.70	8.19	9.46	0.00	15.94	10.71	5.12	11.79	14.29	8.28	6.91	8.24
104	0.00	0.00	0.00	0.00	0.00	0.00	0.79	1.22	0.00	0.00	0.00	0.00
106	0.00	0.00	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
108	0.75	0.86	0.68	0.00	0.72	0.00	0.39	0.00	0.00	0.69	0.18	0.00
110	0.00	0.00	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00
112	1.87	1.72	2.70	0.71	3.62	0.40	0.79	0.41	2.86	0.69	1.45	0.92
114	2.99	2.59	2.70	1.43	0.72	1.19	1.97	3.25	2.86	0.69	0.73	2.38
116	2.61	1.72	1.35	0.00	2.90	2.38	1.57	1.63	2.86	1.38	3.27	2.75
118	12.69	9.91	18.92	11.43	15.22	10.71	10.63	13.82	8.57	5.86	8.00	7.69
120	23.13	27.16	22.30	21.43	29.71	20.63	20.87	22.76	27.14	27.59	24.00	27.66
122	2.61	0.86	2.03	5.00	2.17	4.76	5.91	5.28	2.86	2.07	1.64	2.75
124	2.99	3.88	3.38	0.00	2.90	0.40	0.39	0.81	0.00	2.76	0.73	1.65
126	1.49	1.29	2.70	0.00	1.45	1.59	1.97	1.63	0.00	2.07	2.73	1.47
128	6.72	6.47	4.73	6.43	3.62	9.92	7.48	2.44	18.57	18.62	10.91	10.44
130	1.12	0.43	0.00	2.86	4.35	2.38	2.36	2.44	1.43	2.07	3.09	2.20
132	2.99	2.59	1.35	2.14	2.17	1.59	1.97	0.81	0.00	2.07	1.45	1.47
134	10.45	19.40	0.68	26.43	1.45	25.00	29.13	21.14	0.00	16.21	21.27	18.50
136	1.12	0.86	0.00	0.00	0.00	0.00	0.00	0.41	0.00	0.34	0.91	0.73
138	1.12	0.86	1.35	0.00	0.00	0.00	0.79	0.00	1.43	1.38	2.18	0.55
140	1.49	1.72	1.35	2.86	0.72	0.79	1.18	1.22	0.00	0.69	0.73	1.47
142	9.33	4.74	13.51	10.00	5.07	2.78	1.97	2.44	11.43	4.83	6.36	6.04
144	3.36	4.31	7.43	7.14	5.80	1.98	2.76	4.07	2.86	1.38	1.82	2.01
146	0.75	0.00	1.35	1.43	0.00	0.00	0.00	0.41	0.00	0.34	0.36	0.18
148	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.00	1.43	0.00	0.18	0.00
150	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18
158	0.00	0.00	0.00	0.00	1.45	0.00	0.00	0.00	0.00	0.00	0.36	0.37
160	0.00	0.00	0.68	0.00	0.00	2.78	1.57	1.22	1.43	0.00	0.18	0.18
164	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.81	0.00	0.00	0.00	0.00

Locus/ Alleles	Populations											
Cocl-lav 41	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
181	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
185	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.00
195	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62

Locus/ Alleles	Populations											
Cocl-lav 41	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
197	0.94	0.43	1.35	1.64	0.72	0.46	0.48	1.35	1.43	1.28	2.16	1.65
199	1.89	1.74	0.00	0.00	1.45	0.93	0.48	0.68	0.00	0.43	0.59	1.44
201	0.94	3.48	2.70	0.00	3.62	5.09	1.43	1.35	5.71	2.56	2.35	2.67
203	2.36	4.35	6.76	17.21	7.97	12.04	8.10	10.14	4.29	2.56	2.55	7.41
205	19.34	20.87	25.00	24.59	28.99	21.30	19.52	18.24	17.14	16.67	13.92	16.46
207	23.58	22.17	33.78	23.77	19.57	15.74	21.90	21.62	25.71	21.37	25.10	14.61
209	5.66	7.83	3.38	0.00	3.62	7.41	4.76	7.43	5.71	5.98	5.10	4.12
211	3.30	3.48	6.08	3.28	5.07	2.78	4.29	3.38	1.43	3.85	1.76	4.53
213	2.83	4.78	4.05	8.20	1.45	5.56	10.95	8.78	5.71	4.70	6.08	6.79
215	3.30	0.43	0.68	9.02	0.72	1.39	0.00	0.00	1.43	4.70	2.55	5.35
217	15.57	9.57	7.43	8.20	16.67	11.11	10.48	10.81	12.86	10.26	19.02	12.14
219	8.96	9.57	4.05	3.28	4.35	12.04	10.48	9.46	8.57	11.11	8.82	11.52
221	8.02	7.83	4.05	0.00	4.35	3.24	6.19	4.73	5.71	9.40	6.27	7.61
223	0.94	2.61	0.00	0.82	0.00	0.93	0.95	1.35	4.29	5.13	2.55	2.88
225	0.00	0.00	0.00	0.00	1.45	0.00	0.00	0.00	0.00	0.00	0.20	0.00
227	1.42	0.43	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
231	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.00	0.00	0.39	0.00
233	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.21

Locus/ Alleles	Populations											
Cocl-lav 28	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
168	2.69	1.42	0.00	4.17	6.52	0.00	1.14	0.84	3.03	3.19	4.89	4.29
170	57.69	61.79	60.56	70.83	61.59	79.13	73.11	72.27	65.15	63.48	63.59	64.37
172	0.00	0.00	0.70	0.00	0.00	0.43	2.65	4.62	0.00	0.00	0.18	0.75
174	2.69	0.94	0.70	0.83	0.00	1.74	1.89	0.42	4.55	2.48	0.54	2.05
176	36.92	35.85	36.62	24.17	28.99	18.70	21.21	21.43	27.27	29.79	30.62	27.43
178	0.00	0.00	0.00	0.00	2.90	0.00	0.00	0.42	0.00	0.00	0.00	0.75
180	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06	0.00	0.19
182	0.00	0.00	1.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.19

Locus/ Alleles	Populations											
Cocl-lav 45	Epo	Nau	TB	ER	HgIs	Sau	Mus	Lud	MR	CR	NMB	BBN
239	1.98	2.97	3.42	0.00	0.74	0.47	1.18	1.74	1.56	2.52	1.86	3.79
241	13.89	10.89	14.38	20.69	14.71	23.83	22.83	20.87	20.31	11.87	15.61	6.82
243	0.40	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00
245	13.89	20.79	17.12	15.52	23.53	11.68	11.81	16.09	15.63	12.59	11.71	18.75
247	0.00	0.00	0.00	0.00	0.00	0.47	1.57	1.74	0.00	0.00	0.00	0.19
249	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
253	38.49	35.15	26.71	25.00	24.26	32.24	37.01	28.26	39.06	48.56	48.88	49.24
255	31.35	29.70	38.36	38.79	33.82	31.31	25.59	31.30	20.31	24.46	21.75	21.21
259	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.13	0.00	0.19	0.00