Population Differentiation of Ontario Lake trout (*Salvelinus namaycush*) using the Major Histocompatibility Complex class II β gene

by

Steven Kuntz

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2014

©Steven Kuntz 2014

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The lake trout (*Salvelinus namaycush*) is a coldwater salmonid with an extensive native distribution across formerly glaciated regions of North America. Lake trout are known to be a glacial 'relict' species and evolved as a result of glaciation events during the Pleistocene era. Historic and recent human activities along with climate change have had a significant influence on the phylogeographic and genetic compositions of current lake trout populations. Human activities like overexploitation, habitat degradation and most importantly stocking have had detrimental effects on the diversity and variability of populations within Ontario. Prior genetic studies of lake trout have used multiple neutral genetic marker systems including allozymes, microsatellites and mitochondrial DNA (mtDNA) to distinguish multiple glacial refugia and phylogeographic lineages within North America. This project differentiated lake trout populations within Ontario using a non-neutral marker, the major histocompatibility (MH) class II beta gene. A total of fifty-seven unique MH class II β1 alleles were identified in 216 individual lake trout from ten lakes across Ontario and one in New York. Native, hatchery stocking sources and native but stocked lake trout populations were characterized for MH diversity and allelic states to assess their diversity and divergence. A geographic map of MH diversity within each lake was completed. Commonly stocked lakes within Ontario have shown a decreased amount of allelic diversity when compared to other populations. The introgression of MH class II beta alleles within the mixed-ancestry populations and the significance of adaptive genetic diversity within and among populations from different histories will be important for the management of lake trout facing future climatic and ecological conditions.

Acknowledgements

First I would like to thank my advisor Dr. Brian Dixon for giving me this opportunity to complete a master's degree under such a distinguished researcher in Canada. You let me work on my own schedule and pace and never made it feel like a "results-orientated" atmosphere more than a "learning-orientated" atmosphere. You allowed me to have the chance to supervise students under my guidance and as a result learned the importance of leadership, organization and communication. I do not know if I will continue on in academia or research but you have given me many tools to succeed in any career I choose.

I want to thank both of my committee members Dr. Chris Wilson and Dr. Jonathan Witt for their guidance and assistance throughout my degree. Dr. Wilson provided me with the samples and background information needed to fulfill this degree. Dr. Witt was always available to help provide any knowledge and insight into any problems encountered.

I would also like to thank the members of the Dixon lab past and present for their continued support, feedback and knowledge throughout my degree. Many of them knew how to take the worst situation and turn it into an exciting new challenge. I would specifically like to thank Terin Robinson for her continued support, knowledge and assistance. She helped me through the most difficult times of my project making sure I kept going. To the men who made the "man cave" possible, I will miss the entertaining and never dull discussions that occurred. Unfortunately, the cave is all but retired but will be remembered in our memories. To my co-op student, Louis who helped me with many minipreps and digestions, all your hard work was greatly appreciated.

Lastly, I'd like to thank my family and friends for their continued support and enthusiasm in my continued education. Through the up and down roller coaster ride that research is they always pushed me to continue on and do my best in any challenge that came my way.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Acknowledgements iv
Table of Contents v
List of Figures viii
List of Tables xii
List of Abbreviations xiii
Chapter 1 General Introduction 1
1.1 Glacial History of North America 1
1.2 Lake trout (*Salvelinus namaycush*) 2
1.3 History, Glacial Refugia, and Dispersal 3
1.4 Populations and Stocking 6
1.5 Phylogeography and Molecular Markers 7
1.6 The Major Histocompatibility Complex 8
1.7 Major Histocompatibility Complex Regions 10
1.8 MHC Polymorphism 13
1.9 Trans-Species Polymorphism 14
1.10 Sources of MHC Polymorphism and Diversity 16
1.11 MHC as a Population Marker 19
1.12 MH Gene Polymorphism in Fish Populations 20
1.13 Objectives of Proposed Research 22
Chapter 2 Materials and Methods 24
2.1 Fish Samples for Population Study 24
2.2 DNA Extraction for Refugia Samples 26
2.3 PCR and Cloning 27
2.3.1 Polymerase Chain Reaction 27
2.3.2 Gel Extraction, Ligation and Transformation 28
2.3.3 Minipreparation of Plasmid DNA 28
2.3.4 Restriction Enzyme Digestion 29

2.4 Sequencing ... 29
2.5 Data Analysis ... 30
2.6 Phylogeographic Analysis ... 30
2.7 Network Analysis ... 31
Chapter 3 Results ... 32
3.1 Allelic Sequences of Ontario/New York Populations ... 32
3.2 Refugia Sequences ... 40
3.3 Selection ... 43
3.4 Phylogenetic and Network Inferences ... 43
3.5 Native Lakes - Hogan, Crystal, Dickey and Macdonald ... 56
3.5.1 Hogan Lake ... 56
3.5.2 Crystal Lake ... 62
3.5.3 Dickey Lake ... 67
3.5.4 Macdonald Lake ... 72
3.6 Hatchey Stocking Sources - Michipicoten, Slate Islands, Manitou and Seneca ... 77
3.6.1 Michipicoten Lake ... 80
3.6.2 Slate Islands Lake ... 85
3.6.3 Manitou Lake ... 90
3.6.4 Seneca Lake ... 95
3.7 Native but Stocked - Simcoe, Opeongo and Kingscote ... 100
3.7.1 Simcoe Lake ... 102
3.7.2 Opeongo Lake ... 107
3.7.3 Kingscote Lake ... 112
3.8 Comparison of MH class II β1 alleles with Salmonid fishes ... 117
Chapter 4 Discussion ... 120
4.1 MH Polymorphism of the β1 domain ... 120
4.2 MH as a Phylogeographic Marker and Refugia Lineages ... 123
4.3 Does MH class II β gene have multiple loci? ... 126
4.4 Genetic Profiles of Stocking Sources ... 128
4.5 Genetic Profiles of Native Populations ... 131
4.6 Genetic Profiles of Stocked Populations ... 132

4.7 Trans-species Polymorphism 135
Chapter 5 General Conclusions 136
5.1 Conclusions 136
5.2 Future Work 136
References 139

List of Figures

Figure 1.1: Major Histocompatibility Class II Molecule 12

Figure 2.1: Eleven lake trout populations involved in this study, ten of which are located within Ontario and one in New York 24

Figure 2.2: Refugia lake trout sampling locations across Canada 26

Figure 3.1: Nucleotide sequences of MH class II β1 unique alleles in lake trout from 11 populations 33

Figure 3.2: Protein sequences of MH class II β1 unique alleles in lake trout from 11 populations 35

Figure 3.3: MH Class II β1 alleles in lake trout from Wisconsinan glacial refugia mtDNA lineages 41

Figure 3.4: Protein sequences of the MH class II β1 alleles in lake trout refugial samples 42

Figure 3.5: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 unique and refugia nucleotide sequences of lake trout within this study 44

Figure 3.6: Network of MH class II β1 study and refugia nucleotide sequences of lake trout 45

Figure 3.7: Unrooted maximum likelihood phylogenetic tree of the unique MH class II β1 nucleotide sequences of lake trout within this study 48

Figure 3.8a: Network of MH class II β1 study nucleotide sequences of lake trout 50

Figure 3.8b: Network of MH class II β1 study nucleotide sequences of lake trout 52

Figure 3.9a: Unrooted maximum likelihood phylogenetic tree of refugial MH class II β1 nucleotide sequences of lake trout within this study 54

Figure 3.9b: Unrooted maximum likelihood phylogenetic tree refugial most common MH class II β1 nucleotide sequences of lake trout within this study 55
Figure 3.10: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Hogan Lake 58
Figure 3.11: Network of MH class II β1 nucleotide sequences of lake trout from Hogan Lake 60
Figure 3.12: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Crystal Lake 63
Figure 3.13: Network of MH class II β1 nucleotide sequences of lake trout from Crystal Lake . 65
Figure 3.14: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Dickey Lake 68
Figure 3.15: Network of MH class II β1 nucleotide sequences of lake trout from Dickey Lake 70
Figure 3.16: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Macdonald Lake 73
Figure 3.17: Network of MH class II β1 nucleotide sequences of lake trout from Macdonald Lake. 75
Figure 3.18: Network of MH class II β1 nucleotide sequences of lake trout, coloured alleles show all stocked alleles in the lakes studied 78
Figure 3.19: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Michipicoten 81
Figure 3.20: Network of MH class II β1 nucleotide sequences of lake trout from the OMNR Michipicoten broodstock 83

Figure 3.21: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Slate Islands 86
Figure 3.22: Network of MH class II β1 nucleotide sequences of lake trout from the OMNR Slate Islands broodstock 88
Figure 3.23: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Manitou Lake 91
Figure 3.24: Network of MH class II β1 nucleotide sequences of lake trout from Manitou Lake 93
Figure 3.25: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Seneca Lake 96
Figure 3.26: Network of MH class II β1 nucleotide sequences of lake trout from Seneca Lake 98
Figure 3.27: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Lake Simcoe 103
Figure 3.28: Network of MH class II β1 nucleotide sequences of lake trout from Lake Simcoe. 105
Figure 3.29: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Opeongo Lake 108
Figure 3.30: Network of MH class II β1 nucleotide sequences of lake trout from Opeongo Lake 110
Figure 3.31: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Kingscote Lake 113

Figure 3.32: Network of MH class II β1 nucleotide sequences of lake trout from Kingscote Lake. 115

Figure 3.33: Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout (*Salvelinus namaycush*), arctic charr (*Salvelinus alpinus*), rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*). 118

List of Tables

Table 1.1: Previous studies that examined lake trout postglacial origins, analyses used and refugia presented 5

Table 2.1: Lake trout samples across Ontario distributed into native, hatchery stocking sources and native/historically stocked populations 25

Table 2.2: Lake trout refugia samples across Canada. 25

Table 3.1: Allelic diversity at Sana-DAB1 of lake trout populations studied 36

Table 3.2: Allelic richness (A_R), haplotype diversity (h), nucleotide diversity (π), non-synonymous (d_N) and synonymous (d_S) substitutions per population in lake trout MH class II β1 sequences 37

Table 3.3: Allelic frequencies of all 57 unique haplotypes per population and study. 38

Table 3.4: Refugia allelic diversity at Sana-DAB1 40

Table 3.5: Alleles from hatchery stocking sources in common with stocked populations in this study 101

List of Abbreviations

Antigen presenting cells	**APCs**
Cluster of differentiation 8	**CD8**
Deoxyribonucleic acid	**DNA**
Ethylenediaminetetraacetic acid	**EDTA**
Endoplasmic reticulum	**ER**
Human Leukocyte Antigen	**HLA**
Isopropyl β-D-1-thiogalactopyranoside	**IPTG**
Infectious salmon anaemia virus	**ISAV**
Luria-Bertani media	**LB**
Major histocompatibility	**MH**
Major histocompatibility complex	**MHC**
Mitochondrial DNA	**mtDNA**
Peptide binding region	**PBR**
Polymerase chain reaction	**PCR**
Pathogen mediated selection	**PMS**
Rough endoplasmic reticulum	**RER**
Sodium dodecyl sulfate	**SDS**
Transporter associated with antigen processing	**TAP**
Tumor necrosis factor	**TNF**
Microsatellite	**µsat**

Chapter 1 – General Introduction

1.1 Glacial History of North America

The physical environment of North America has drastically changed over the last 2,000,000 years through repeated glacial cycles of approximately 100,000 years (Pielou, 1991). Although more than 20 glacial cycles were believed to occur during the Pleistocene era in North America only the four major glaciations (Nebraskan, Kansan, Illinoian and Wisconsian) were named based on the location where evidence was originally obtained (Briggs, 1986; Flint, 1971; Pielou, 1991). For the duration of these Pleistocene glaciations, the formation of any inland lakes was prevented by glaciers covering the landscape. The challenges faced by aquatic species were enormous during this time, as the glaciers radically changed the distribution patterns of species causing profound evolutionary effects. As a result of the glaciations, aquatic species were forced to survive for thousands of years in peripheral habitats along the glaciers across North America which became known as refugia (Crossman & McAllister, 1986). These isolated refugia populations contained reduced gene pools losing genetic variation as a result of genetic drift (Briggs, 1986). This caused rapid evolutionary change in some species as populations were eliminated while others became only remnants of populations that once existed thousands of years ago (Briggs, 1986).

Continuous shifting of climates caused glaciers to retreat and then advance continuously throughout the Pleistocene glacial age, which produced scouring in the land providing basins for many lakes we see today. The warming climate caused ice sheets to melt, water bodies to move and sea levels to rise, resulting in many ecological changes (Pielou, 1991). Large proglacial

inland lakes formed from meltwater across North America, and provided new opportunities for aquatic species to disperse and extend their geographical ranges across areas that were previously covered in ice. Species using these avenues of dispersal had much larger ranges but lower diversities than taxa in non-glaciated regions (Bernatchez & Wilson, 1998). Present day aquatic species were significantly influenced by the Pleistocene glaciations through habitat destruction, displacement and population bottlenecks. A prime example of a species that underwent such processes is the lake trout, *Salvelinus namaycush.*

1.2 Lake trout (*Salvelinus namaycush*)

Lake trout (*Salvelinus namaycush*) are covered in large roundish spots and range in colour from silver to gray to green to brown to almost black but are usually greenish-gray (Martin & Olver, 1980). Their average adult size ranges from 15 – 20 inches (38 – 50 cm) in length (Scott, Scott, & Crossman, 1985). Lake trout have been known to exhibit limited sexual dimorphism with the males having slightly more pointed snouts than females. They are known primarily to spawn in the fall from late September to November on rocky shoals usually close to deep water (Martin & Olver, 1980; Wilson & Mandrak, 2004). Lake trout are omnivorous, eating plant material, annelids, crustaceans, insects, arachnids, molluscs, fishes and even small mammals (Martin & Olver, 1980).

Lake trout are known to be a glacial 'relict' species and evolved as a result of glaciation events during the Pleistocene era (Wilson & Mandrak, 2004). The species evolved approximately 3 million years ago from other species of *Salvelinus*. Their extensive native distribution across North America corresponds closely to deep freshwater oligotrophic habitats, which were created

by glacial scouring in formerly glaciated regions (Lindsey, 1964; Martin & Olver, 1980). Lake trout prefer cooler waters in the 4 - 9°C range but have been located in temperatures higher and lower (Martin & Olver, 1980). Although typically restricted to only freshwater surroundings, there have been some records of this species showing anadromous characteristics within the Arctic (Swanson et al., 2010). Lake trout are also economically important in commercial fisheries as a food fish and sport fisheries as a game fish throughout central and northern Canada.

1.3 History, Glacial Refugia and Dispersal

The complex history of aquatic species has made understanding freshwater fish like lake trout very difficult, resulting in many conflicting opinions when it comes to the number of glacial refugia they originated from and their zoogeography listed in Table 1.1 (Black, 1983a, 1983b; Crossman & McAllister, 1986; Grewe & Hebert, 1988; Ihssen, Casselrnan, & Martin, 1988; Khan & Qadri, 1971; Lindsey, 1964; Radforth, 1944; Walters, 1955; Wilson & Hebert, 1996; 1998). Radforth (1944) argued based upon distributional patterns that lake trout survived the glaciations in a single Beringian refuge north of the ice sheet and have spread south since. Walters (1955) disagreed with Radforth, stating rather that lake trout survived and spread from a single Mississippian refuge. Lindsey (1964) was the first to state that lake trout must have survived the glaciations in multiple refugia groups; some in the south while others were in the north. Based upon distribution patterns, it was thought there could be as many as four different refugia. Kahn and Qadri (1971) supported Lindsey with some morphological investigations on lake trout across its range with 4 different gene pools: Atlantic, Mississippian, Missourian and Beringian. A decade later Crossman and McAllister (1986), using morphological multivariate tests, revised the total number of refugia groups back down to two: Mississippian and Beringian.

Research by Black (1983a, 1983b) using two different parasitic nematodes *Cystidicola stigmatura* and *farionis,* both of which infect the swim bladders of lake trout, supported multiple refugia groups once again. Black's (1983a) study using *stigmatura* showed lake trout of Mississippian origin were parasitized by the nematode while other lake trout from the Atlantic and western refugia were not. It provided independent confirmation of dispersal of lake trout from a Mississippian origin, but absence of the nematode in the other refugia inferred a different origin but did not confirm it, as the parasite also required an alternate host (*Mysis relicta*) with limited dispersal abilities. The study with *Cystidicola farionis* (Black, 1983b) showed the opposite in that this nematode was restricted to lake trout that survived in a Beringian refuge while it was not present in the others. This suggested multiple refugia groups allowing lake trout to invade North America in three directions from the southeast and northwest establishing Mississippian, Bering, and other populations. Following these studies, molecular markers started to be applied to understand lake trout's history. Through the use of allozyme data, two distinct genetic groups of lake trout in Ontario and Manitoba were determined to be Mississippian and Atlantic in origin, and a third genetic group was inferred, possibly Beringian (Ihssen et al., 1988). Analysis of mitochondrial DNA (mtDNA) revealed three distinct lineages: Atlantic, Mississippian and Beringian origins (Grewe & Hebert, 1988). Finally, through more extensive examination of geographical distribution with mtDNA, it was concluded that lake trout survived in possibly six different refugia in the Pleistocene glaciations: two Mississippian, one Atlantic, one Nahanni, one Beringia and one Montana (Wilson & Hebert, 1996; 1998).

Table 1.1: Previous studies that examined lake trout postglacial origins, analyses used and refugia presented.

Study	Method	Number of Refugia Determined	Refugia Groups
Radforth 1944	Distribution	1	Beringian
Wynne-Edwards 1947	Distribution	1	Mississippian
Wynne-Edwards 1952	Distribution	2	Mississippian, Beringian
Walters 1955	Distribution	1	Mississippian
McPhail 1963	Distribution	1	Mississippian
Lindsey 1964	Distribution	4	Atlantic, Mississippian, Missouri, Beringian
Khan and Qadri 1971	Morphological	4	Atlantic, Mississippian, Missouri, Beringian
Black 1983a,b	Parasites	3	Atlantic, Mississippian, Beringian
Crossman and McAllister 1986	Morphological	2	Mississippian, Beringian
Stewart and Lindsey 1983	Distribution	1	Mississippian
Ihssen et al. 1988	Allozymes	2	Atlantic, Mississippian
Grewe and Hebert 1988	Mitochondrial DNA	3	Atlantic, Mississippian, Beringian
Wilson and Hebert 1996, 1998	Mitochondrial DNA	≥5	Atlantic, Mississippian, Beringian, Nahanni, Montana

1.4 Populations and Stocking

Historic and recent human activities, along with climate change, have had a significant influence on phylogeographic and genetic compositions of current lake trout populations. Overexploitation, habitat degradation and most importantly stocking have had major effects on the genetic makeup and fitness of lake trout within Ontario native populations (Wilson & Mandrak, 2004). The lake trout's characteristics of slow growth, late maturity, low reproductive potential and slow replacement rate make this species extremely vulnerable to overexploitation (Shuter et al., 1998). Long-term stocking programs have been widely practiced for most of the past century to compensate for the loss of naturally reproducing native fish available to restore ecological habitats and populations (Halbisen & Wilson, 2009). Stocking has raised many questions regarding its suitability for enhancing wild populations. There are many factors which influence stocking success in populations such as habitat suitability, age/size of fish when stocked, fish health and condition, handling/transport stress, stocking technique, interspecific competition, predation, genetic strain and forage availability (Kerr, 2001). Often it is quite a challenge to make stocking into an existing population a success. Unfortunately, only recently has stocking been seen to have extremely damaging effects, with the potential to cause native populations to disappear completely (Araki et al., 2007; Brannon et al., 2011; Evans & Willox, 1991). Other potential impacts stocking can have include hybridization/genetic introgression, displacement/loss of other species, competition for food, predation, introduction of diseases/parasites and increased angling pressure (Kerr, 2001).

As a consequence, supplemental stocking of naturally-reproducing lake trout populations was discontinued in 1985 (Evans & Willox, 1991). Now, inland lakes within Ontario are only

stocked for rehabilitation purposes or to support introduced fisheries that draw angling pressures away from sensitive indigenous lake trout populations (Halbisen & Wilson, 2009).

1.5 Phylogeography and Molecular Markers

Phylogeography looks at the distribution of genealogical lineages within species, which requires the use of different types of molecular markers. Previous genetic analysis of lake trout have used allozymes, microsatellites, and mitochondrial DNA (mtDNA). Nuclear DNA or genes are inherited biparentally and are often used as population markers also. Allozymes were one of the first markers used in molecular ecology to differentiate populations. Ihssen et al., (1988) used allozymes to differentiate lake trout populations in Ontario and Manitoba. Allozymes are functional proteins encoded by structural genes which make them very easy and simple to use. Unfortunately even though allozymes are very easy to assay, they often lack variability therefore are usually used in concert with other DNA markers to differentiate populations. One of the most common nuclear DNA markers is microsatellites. They consist of tandem repeats of nuclear DNA flanked by conserved regions. Microsatellites mutate quite rapidly, therefore they are only used to understand evolutionary events in the relatively recent past (Freeland, 2005). Mitochondrial DNA (mtDNA) consists entirely of maternally inherited DNA which means it is uniparentally inherited, haploid and functionally acts as a single locus marker (Freeland, 2005). Even though the arrangement of genes is conserved, mtDNA is often used due to its high mutation rate compared to nuclear DNA, which can be around ten times the rate of synonymous substitutions in protein-coding nuclear genes (Brown et al., 1979). Also the lack of recombination in mitochondrial genes results in an effectively clonal inheritance, allowing

lineages to be tracked over time and space with ease (Freeland, 2005). The problem with mtDNA is that using only a single locus to reconstruct populations is not ideal if the populations could have been subjected to selection or another process contributing to an unusual history (Freeland, 2005) or mixing between divergent groups. Usually mtDNA provides an additional genetic character when supplemented through comparison to nuclear DNA by microsatellites (Grewe et al., 1993). The use of mitochondrial DNA in population phylogeography and differentiation in lake trout has become a useful tool which has been used in many studies (Avise et al., 1987; Grewe & Hebert, 1988; Grewe et al., 1993; Moritz, 1987; Wilson et al., 1985; Wilson & Hebert, 1996; Wilson & Hebert, 1998) as well as in many other species.

The choice of which molecular marker is used in studies to differentiate populations is a very important task. Each marker allows the researcher to look at populations in a different way. The difficulty with studying lake trout is that they are a glacial relict species meaning the present day populations evolved as a result of the glaciation events in the Pleistocene era, essentially in the last 10,000 – 20, 000 years. Keeping this in mind, there has not been much time for populations to evolve and differences to arise. This thesis will use a molecular marker that can differentiate stocks based on the selection by the environment. This marker changes more rapidly and thus provides a different view from past genetic markers.

1.6 The Major Histocompatibility Complex

The discovery and first description of the major histocompatibility complex (MHC) happened well over 60 years ago (Klein, 1987) but only within the last few decades have we truly started to understand the nature and function of the complex. The MHC was first

discovered in mice as a genetic region governing the ability of transplant success (Klein, 1986). The name MHC was given because the genes involved in graft rejection were clustered into a major chromosomal region (Danchin et al., 2004). MHC genes are present in all jawed vertebrates (gnathostomes) but not jawless vertebrates (agnathans) or any other metazoan phyla (Danchin et al., 2004; Klein & Sato, 1998). Since the discovery of MHC, it has been extensively researched in mammals, but it was only 20 years ago the first MHC genes were identified in carp (Hashimoto et al., 1990). Following Hashimoto et al. (1990) many MHC genes were isolated in other teleosts including Atlantic salmon (Fosse et al., 1991; Grimholt et al., 1993; Hordvik et al., 1993), rainbow trout (Glamann et al., 1991; Juul-Madsen et al., 1992), and carp (Ono et al., 1993). The idea of using MHC typing in conservation biology has been an important topic recently. MHC typing can be used to give more information than just genetic diversity, it has been associated with individual variation in parasite load, local adaptations, maternal-foetal interactions, life-time reproductive success, mate choice and used to plan captive breeding programs (Ujvari & Belov, 2011).

Structure differences and rearrangements in MHC have been seen in different mammals, birds and amphibians to various degrees when compared to humans (Dixon et al., 1995; Flajnik & Du Pasquier, 1990; Miller et al., 1994; Nakamura et al., 1986; Salter-cid, Kasahara, & Flajnikl, 1994). Even though variation was seen in each different species, most contained a chromosomal cluster where the classical class I and II genes were linked. The teleostean fish were the first discovered where this was not the case. Bingulac-Popovic and co-workers (1997) revealed that within zebrafish, *Danio rerio,* the two classes were on separate chromosomes and the class II loci were split between at least two different chromosomes. Following this study,

nonlinkage was seen in many other teleosts including carp (Wiegertjes et al., 1997), salmon (Stet et al., 1997), stickleback, guppy and cichlids (Sato et al., 2000). From this, it has been shown within the jawed vertebrates, the tetrapods and teleostean fishes have their MHC loci arranged differently. Following continuous studies that confirmed the nonlinkage of MHC genes in most teleosts, Stet and Dixon started referring to them as simply MH genes as the "complex" within the name was suggestive of the linkage of genes in tetrapods (Dixon & Stet, 2001; Dixon, 2008).

1.7 Major Histocompatibility Complex Regions

It has been nearly 15 years since the complete human MHC sequence and gene map was reported. The 3.6-Mbp DNA sequence consisted of 224 gene loci located on chromosome 6 (MHC Sequencing Consortium, 1999). In mammals, the MHC consists of three regions class I, II and III. Most notably, class I and II genes encode proteins involved in antigen presentation.

The MHC class I molecule is a cell-surface heterodimer containing a membrane bound alpha "heavy" chain and a light chain. The alpha 1 and 2 of the heavy chain fold together to form the peptide binding groove, which is known to be highly polymorphic allowing the ability to bind a variety of peptide antigens 8-10 amino acids in length (Engelhard, 1994). β2m is critical to the cell-surface stability of the receptor and without it the heavy chain is unable to bind peptides to present to CD8+ T-lymphocytes (Vitiello, Potter, & Sherman, 1990). Thus, the heavy alpha chain binds non-covalently to the non-polymorphic light chain β2m in the rough endoplasmic reticulum (RER) waiting for peptide antigens to be translocated into the endoplasmic reticulum (ER). Endogenous proteins representing intracellular pathogens such as viruses and bacteria are degraded to peptide antigens by the proteasome within the cytosol.

These peptide antigens are translocated to the ER by the transporter TAP (Murphy, Travers, & Walport, 2008). Once the peptide antigen binds to the MHC class I molecule to complete its folding, the MHC class I molecule is released from the ER and exported to the surface of all nucleated cells (York & Rock, 1996).

The MHC class II receptor (Figure 1.1) is also a heterodimeric molecule containing two membrane bound subunits, an alpha chain and beta chain. The two polypeptide chains are noncovalently associated and each contain a transmembrane, cytoplasmic and two structural external domains (Steinmetz & Hood, 1983). The alpha 1 and beta 1 domains also fold to form the highly polymorphic peptide binding groove, which is open at both ends compared to MHC class I molecules. This open peptide binding region allows for larger peptide antigens of 12-16 amino acids in length to be bound (Hemmer et al., 2000). While the peptide binding regions show high allelic variation the alpha and beta 2 domains are very conserved like the alpha 3 chain in MHC class I receptors. MHC class II molecules are expressed on the cell surface of antigen presenting cells (APCs) such as macrophages, B lymphocytes and dendritic cells which present exogenous extracellular peptides to CD4+ helper T lymphocytes (Watts, 1997). The MHC class II receptor is loaded with these extracellular peptide antigens using the lysosomal pathway. Vesicles containing degraded extracellular peptide fragments fuse with the MHC class II dimer, which was originally produced within the RER (Murphy et al., 2008). The MHC class II molecule cannot bind a peptide within the RER because the groove is blocked by the MHC class II associated invariant chain. The CLIP peptide from the invariant chain is released within the endocytic pathway and once released; the MHC class II molecule binds the exogenous

peptides degraded from the lysosomes and travels to the cell surface to be expressed on APCs (Kasaharaa, Flajnikb, & Natoric, 1995).

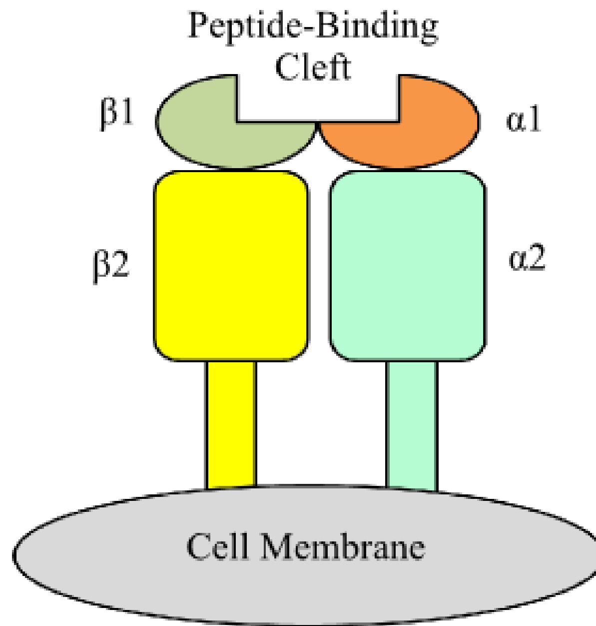


Figure 1.1. Major Histocompatibility Class II Molecule.

The MHC class II complex contains two membrane bound subunits, an alpha chain with two extracellular domains α_1 and α_2 and beta chain with two extracellular domains β_1 and β_2. The two polypeptide chains are noncovalently associated and each contain a transmembrane, cytoplasmic and two structural external domains (Steinmetz & Hood, 1983). The α_1 and β_1 domains fold to form the highly polymorphic peptide binding groove.

The third region of the MHC region called class III encodes genes for various immune functions such as the complement system, members of the tumor necrosis factor (TNF) family as well as other functionally unrelated proteins and cytokines (Danchin et al., 2004).

1.8 MHC Polymorphism

Even though MHC class I and II have very similar structures containing conserved amino acid sequences, there are domains within each protein known as the peptide binding regions that are highly polymorphic. The MHC is known to contain the most highly polymorphic nuclear encoded genes known (Nei & Hughes, 1991). Since 1989, alleles have been analyzed and named from the MHC genes HLA (Human leukocyte antigen) class I and II in humans. Currently, there are a total of 8,016 HLA alleles known: 6,292 are HLA class I alleles and 1,724 are HLA class II alleles (Robinson et al., 2011). Every individual human has three HLA class I genes and possibly four HLA class II genes on chromosome 6, thus a person can expresses up to six different HLA class I alleles and eight different HLA class II alleles on his or her cells (Murphy et al., 2008). Therefore, the number of possible combinations for class I and class II alleles in the human population is so vast that it is highly unlikely that any two individuals will ever share the same set of HLA class I and II alleles unless they are identical twins. This large pool of alleles allows for a wide range of protection against the various pathogens as well as different alleles will have different capacities for binding peptides.

The extent of MHC polymorphism seen in humans does not occur in every species, for example the cheetah. In a study by O'Brien and Yuhki, (1999) there was only 4 MHC class II beta alleles within 6 individual African cheetahs. Comparable to the cheetah, there were only 10

MHC class II beta alleles within 40 individual tigers. Based on the small populations and limited samples the cat family (*Felidae*) does show limited polymorphism. Although a small pool of alleles within a population can affect the ability of species to combat diseases and pathogens leaving them susceptible to new emerging infectious pathogens. This can contribute to the fact that many wild species of the cat family (*Felidae*) are endangered around the world. Pathogens have the potential to cause very intense selection pressures on particular MHC alleles and as a result the fitness difference conferred by alleles carried by individuals has been observed as high as 0.5 (Lohm et al., 2002).

The best examples of how a specific MHC allele can provide increased resistance to infectious diseases is detected in studies using model organisms where the species can be challenged and observed under controlled conditions. A great example of this occurs in the chicken, where varying B haplotypes have clearly demonstrated a range of resistance to Marek's disease (Bacon, 1987; Bacon, Hunt, & Cheng, 2000). Specifically, allele B21 was revealed to confer 7 times more resistance when compared to other alleles (Burgess, Basaran, & Davison, 2001). There are also many examples where certain haplotypes within humans can increase resistance or increase susceptibility to a disease.

1.9 Trans-Species Polymorphism

Jane Klein (1987) proposed the "Trans-Species Hypothesis" to explain the phenomenon of the origin and significance of MHC polymorphism. The trans-species hypothesis states that present differences between alleles of MHC loci within a current species were already established in extinct species in the ancestral evolutionary line (Klein, 1987; 1989). Thus, MHC

polymorphism usually predates the present species that maintain it, meaning identical, or nearly identical, sequences could occur in different species which have been separated for millions of years.

Following the emergence of this hypothesis, it was not long before support was published. Within a year a paper published comparing chimpanzee and human MHC class I alleles (Mayer et al., 1988), concluding that some human alleles are much more similar to chimpanzee alleles than to other human alleles. It was then seen in a comparison between the mouse and rat (McConnell, Talbot, McIndoe, & Wakeland, 1988). Research done on the cichlids of Lake Malawi showed shared polymorphic MHC lineages in a group of species that diverged 6000 to 2 million years ago (Klein, Satta, & O'hUigin, 1993). Following that, Graser et al. (1996) discovered allelic lineages that are related in three different *Danio* species (*Cyprinidae*) of the class II gene. More recently, Ottova et al. (2005) looked at 11 fish species within the Cyprinidae family and the evolution of exon 2 in MHC class II beta gene featured phylogenetic trees with clusters containing alleles from several different fish species (Ottova et al., 2005). Trans-species polymorphism was limited to within the subfamilies Cyprininae and Leuciscinae but not between them. A study on Atlantic salmon and brown trout by Stet et al. (2002) provided support, showing species with a short divergence time of about 4.5 million years have alleles with more similarity to another species than alleles of the same species. Carp and barbels shared a common ancestor over 30 million years ago and both shared equivalents of class II MHC genes with zebrafish, from which they diverged more than 50 million years ago (Dixon et al., 1996). These orthologous sequences have been passed down the evolutionary line showing that polymorphism does indeed evolve in a trans-species fashion as predicted by Klein in 1987.

There have been some species where very little support for trans-species polymorphism has been determined. A study of Pacific salmonids class II beta sequences by Miller and Withler (1996) proposed that there is little evidence for trans-species or shared ancestral polymorphism in the genus *Oncorhynchus.* They determined that very few alleles of *O. tshawystscha* and *O. kisutch* grouped or aligned with alleles of other species. Speciation in the genus *Oncorhynchus* is estimated to have occurred around 10 million year ago but the authors reasoned that there was little trans-species polymorphism due to the numerous founder effects resulting from the numerous glaciations that the *Oncorhynchus* species has dealt with. A few years later endangered Chinook salmon in a California lake were seen to carry alleles that descended from two of six major lineages that are common to all Pacific salmon (Garrigan & Hedrick, 2001). The most recent common ancestor of all the alleles probably existed approximately 15 million years ago, before the Pacific salmon species radiation. These studies endorsed the validity of trans-species hypothesis once again showing that it occurs within all species. The idea that founder effects from glaciations removed all differences in alleles from ancestral species is very unlikely. It is possible though that the evolution of trans-species polymorphism is related to speciation and diversification patterns of species (Stet et al., 2002). Some species will show much higher levels of trans-species polymorphism compared to others depending on their evolutionary history.

1.10 Sources of MHC Polymorphism and Diversity

It is thought that multiple factors contribute to the high levels of MHC gene polymorphism detected within a typical population. It could not just be generated through the

accumulation of single point mutations over time thus it is believed that the process may include recombination, gene conversion and pathogen-mediated selection (Spurgin & Richardson, 2010).

Intralocus or interallelic recombination is thought to be one of the processes contributing to polymorphism within the peptide binding regions of MHC genes. Recombination can occur in a small scale fashion or a large scale fashion. Small scale recombination occurs when only short stretches of nucleotides are transferred from one allele to another compared to large scale recombination which involves exchanging one or more full exons (Yeager & Hughes, 1999). Extremely complex hotspots of recombination have been identified in some species where the recombination rates have exceeded the rate of point mutations by up to ten times (Schaschl et al., 2005).

Pathogen-mediated selection (PMS) is believed to generate extraordinary levels of MHC diversity. PMS is thought to be driven by three mechanisms: heterozygote advantage, inverse frequency dependence and fluctuating selection. Heterozygote advantage suggests heterozygotes can respond to a wider range of pathogenic peptides than homozygotes at MHC loci, thus benefiting from increased resistance (Hughes & Nei, 1988). It can play a role when pathogen resistance is dominant meaning heterozygotes exhibit the same level of fitness as the fittest homozygote therefore the heterozygous genotype has higher levels of fitness than the average for all homozygotes (Spurgin & Richardson, 2010). If overdominant, the combined effect of two alleles at a locus will result in the heterozygote being fitter than the fittest homozygote (Spurgin & Richardson, 2010). Inverse frequency dependence says that there is strong selection on pathogens to overcome the resistance of the most common MHC alleles causing new low

frequency alleles to arise which provide greater pathogen protection, giving a selective advantage (Takahata & Nei, 1990). Finally fluctuating selection proposed that variation in the abundance of pathogens present may maintain diversity at MHC (Hill, 1991). It is believed that if the pathogens challenged by an organism fluctuate so will directional selection, leading to different MHC alleles selected upon at various times. This varying selection will maintain the genetic diversity across subpopulations (Spurgin & Richardson, 2010). It is thought that all three of these mechanisms work concurrently or separately in order to increase MHC diversity and polymorphism within populations. Some populations may have increased polymorphism as a result of increased pathogen loads, compared to some populations where there are low levels of pathogens located within their environments.

A great example of how selection by pathogens can shape MH alleles frequencies was seen in wild Atlantic salmon populations. Eggs with known haplotype frequencies from parental crosses were released into a river and six months later the surviving individuals were analyzed (De Eyto et al., 2007). Selection resulting from disease would be detectable at immunogenic loci like MH while forces like genetic drift, migration and mutation would be detected at both the control and immunogenic loci. Both MH class I and class II loci along with eight control markers were tracked in the study but the MH class I along with the controls did not diverge from the parental ratios. The MH class II frequencies diverged significantly, showing directly that selection by wild pathogens shaped the allele frequencies. Two other studies on Atlantic salmon showed the effect selection can have on MH polymorphism. Langefors et al. (2001) showed the impact of the bacterium, *Aeromonas salmonicida* on MH class II beta alleles and how pathogen-driven selection maintained MH polymorphism. Following this, Grimholt et al. (2003) exhibited

the association between MH class I and class II polymorphism with disease resistance to infectious salmon anaemia virus (ISAV) and the bacterium *Aeromonas salmonicida*.

1.11 MHC as a Population Marker

Any genomic area that contains high levels of variability may be used as a molecular marker for population differentiation. The MHC is known to contain the most highly polymorphic nuclear genes known, which play a crucial role in the immune response to foreign antigens (Nei & Hughes, 1991). For this reason, the MHC is a molecular marker not only for the genetic variability of individuals within populations, but also for their interactions with the environment due to the presence of new pathogens. MHC allele polymorphism is vital to most species as it allows a wide range of protection against a large variety of pathogens. Since these markers are subject to rapid selection, they provide a potential method to differentiate populations that have been separated for only short periods of evolutionary time.

Many population studies have been performed to assess the use of MHC as a population marker, specifically comparing the patterns of MHC diversity to that of neutral markers. Two studies on salmonid fishes reported differential patterns of population structure between the MH and neutral loci markers. A study on wild Atlantic salmon (Landry & Bernatchez, 2001) compared the MH class II beta gene and microsatellites as population markers to see the genetic differentiation within populations. Landry and Bernatchez concluded that local adaptation may have promoted the maintenance of different subsets of alleles among different populations in Atlantic salmon as higher population differentiation was seen using MH markers. Following this, a study by Miller et al. (2001) on sockeye salmon reported a higher level of population

differentiation using MH markers compared to microsatellites. They concluded that the high degree of population variation at MH may have been due to spatial and/or temporal heterogeneity in pathogen driven directional selection. Both of these studies in salmonids showed excellent examples of higher levels of population differentiation seen using MH markers than were detectable with microsatellites (Bernatchez & Landry, 2003). MH class II beta differentiation in Steelhead populations in California was also higher than that seen with microsatellites (Aguilar & Garza, 2006). MH loci were determined to be particularly useful once the populations were separated into three different geographic regions.

MH genes have also been shown to be a useful marker system without the use of neutral markers. Genotypic studies by Dorschner et al. (2000) and Noakes et al. (2003) on lake trout showed high levels of diversity along with the ability to possibly differentiate populations. Conejeros et al. (2012) in a study on Arctic charr (*Salvelinus alpinus*) showed that MH class II genes were successful in differentiating populations, particularly the β subunit.

1.12 MH Gene Polymorphism in Fish Populations

A high degree of MH gene polymorphism has been seen in many fish species and has also been used to differentiate morphotypes, populations and even stocks. Lake Tana in Ethiopia contains 14 barbel morphotypes, each of which inhabits different ecological niches within the lake. A study by Dixon et al. (1996) followed by Kruiswijk et al. (2004, 2005) differentiated four sympatric barbel morphotypes using the MH class II beta gene discovering more polymorphism than is seen in many mammals. A total of 57 alleles in only 17 individuals were discovered which segregated into distinct allelic lineages (Dixon et al., 1996).

MH has been used as a population marker in many salmonid studies to differentiate populations. In Pacific salmon, Miller and Withler (1996) identified 29 different MH class II alleles by sequencing the beta 1 and beta 2 domains from 35 individual fish. Following this, Miller identified more MH alleles but this time in Chinook salmon. The study revealed 21 MH class I alpha 1 alleles from 36 fish but only 3 class II beta 1 alleles were identified within 47 fish (Miller, Withler, & Beacham, 1997). Thus the allelic diversity of class II beta 1 compared to class I was much lower, although the class II beta locus provided almost a complete separation of the two populations studied. A final study was done in Sockeye salmon showing how variation at the MH class II beta locus throughout 31 populations was greater than a neutral genetic marker. Miller et al. (2001) detected 25% variation at the MH locus compared to only 5% variation at neutral loci.

In Atlantic salmon, variation of MH class II beta locus was compared among 14 samples from seven different rivers and seven subpopulations selected to cover a variety of habitats and geographical scales. A total of 40 nucleotide positions out of 254 were determined to be polymorphic within the 18 MH class II beta alleles (Landry & Bernatchez, 2001). A different study by Langefors et al. (2001) supported MH polymorphism through pathogen-driven selection demonstrated similar amounts of polymorphism within Atlantic salmon with 53 variable nucleotide positions among 257 nucleotides for a total allele count of 17.

High amounts of polymorphism were also seen in Arctic charr in two separate global studies, one using the MH class II alpha and the other using MH class II beta as a marker. The

MH class II alpha gene as a marker had high amounts of polymorphism with 27 different alleles within 66 fish, 33 of which had two different alleles (Conejeros et al., 2008). The study using MH class II beta gene as a marker located 21 unique alleles that varied in length by 3-6 bp with a mean sequence divergence of 7% within 65 individuals (Conejeros et al., 2012). Throughout many of these studies the levels of polymorphism and diversity were high if compared to other genes. Any study that detected populations with low amounts of genetic diversity and polymorphism may have been the result of small sample sizes. High levels of allelic diversity and polymorphism were also shown in lake trout from Lake Superior. In a study by Dorschner et al. (2000) 43 MH class II beta alleles were obtained from 74 fish in six different populations. This was followed up a few years later by Noakes et al. (2003) who identified 65 genotypes in 80 lake trout from the Apostle Islands within Lake Superior. Only 20% of alleles were shared between the siscowet and lean morphotypes.

1.13 Objectives of Proposed Research

The aim of this project was to assess the diversity of the major histocompatibility (MH) class II beta gene within Ontario lake trout populations, as well as differentiation among populations based on ancestry. Native, mixed ancestry and hatchery lake trout populations across Ontario were characterized for MH diversity and allelic states to assess their diversity and divergence. From this, a geographic map of MH structure and diversity was generated and compared to previous studies that relied on neutral markers. The introgression of MH class II beta alleles within the mixed-ancestry populations and the significance of adaptive genetic diversity within and among populations from different histories will be important for management and research of stocks.

If the MH class II beta gene works well as a population marker then significant differentiation between inland (wild) and hatchery based lake trout populations can be expected. The hypothesis is that although the majority of lake trout populations within Ontario were largely founded by a single Mississippian source, these populations will have diverged from one another and become genetically distinct due to habitat isolation and local adaptation. Historical stocking of hatchery-reared lake trout into the indigenous native populations may have caused a homogenizing effect on the native population's allelic diversity. Finally, higher amounts of allelic diversity compared to a single locus system will be present due to the existence of multiple MH class II beta loci. Multiple loci have been shown in other salmonids, for example Atlantic salmon (Harstad et al., 2008) and it may be the case for all salmonids.

Chapter 2 – Materials and Methods

2.1 Fish Samples for Population Study

Approximately 20 individual fish from each of 11 lake trout populations (Figure 2.1) from across Ontario were sampled for this study.

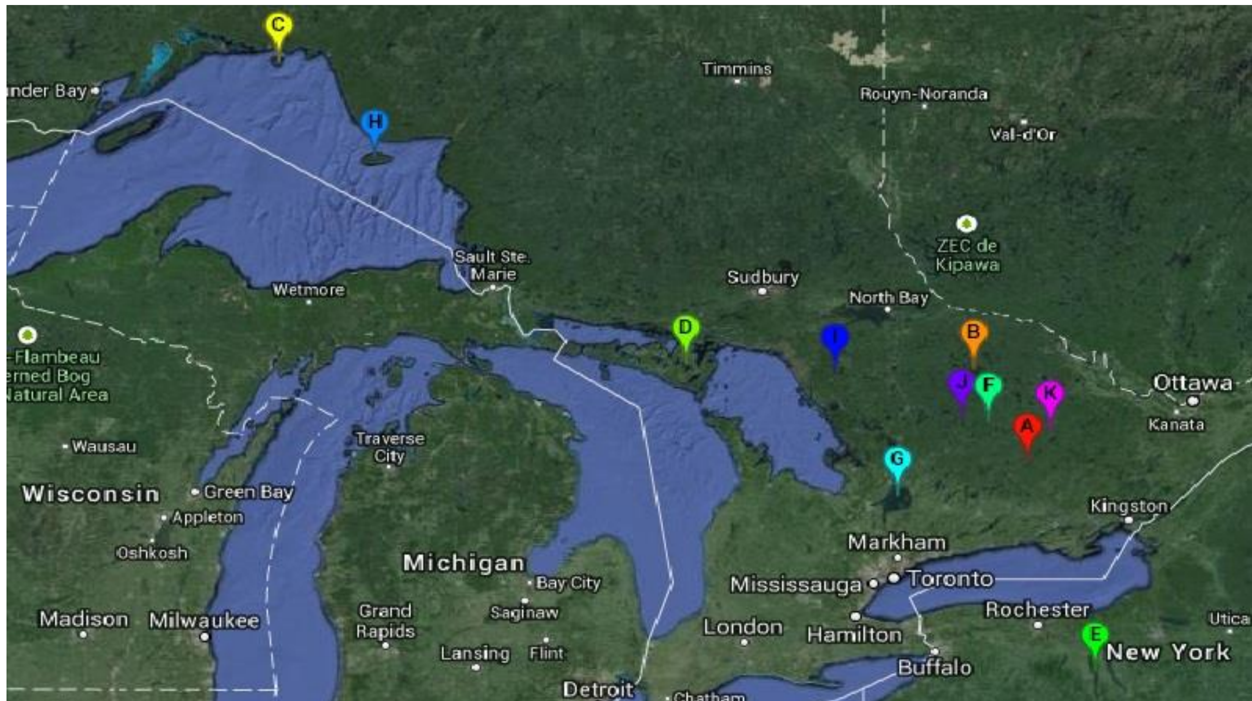


Figure 2.1. Eleven lake trout populations involved in this study, ten of which are located within Ontario and one in New York. Study lakes are indicated with different coloured markers. A – Dickey Lake, B – Opeongo Lake, C – Slate Islands, D – Lake Manitou, E – Seneca Lake, F – Kingscote Lake, G – Lake Simcoe, H – Michipicoten Lake, I – Hogan Lake, J – Macdonald Lake, K – Crystal Lake.

These populations were put into three main categories: native, hatchery stocking sources and native/historically stocked populations (Table 2.1).

Table 2.1: Lake trout samples across Ontario distributed into native, hatchery stocking sources and native/historically stocked populations.

Native	Hatchery Stocking Sources	Native/Historically Stocked
Macdonald Lake	Slate Islands	Kingscote Lake
Crystal lake	Manitou Lake	Opeongo Lake
Hogan Lake	Seneca Lake	Simcoe Lake
Dickey Lake	Michipicoten Lake	

Refugia group samples were collected from various areas in Canada to use for comparison to the Ontario populations. The specific lake trout populations for each refugia group can be seen in Table 2.2 and the locations are shown in Figure 2.2.

Table 2.2: Lake trout refugia samples across Canada. Numbers in brackets indicate the number of samples from each lake within the respective refugial groups.

Refugial Group	Lakes
Atlantic	Lac Archambault (1)
Mississippian	Hawley Lake (4)
Atlantic/Mississippian	Seneca Lake (5), Conferderation Lake (1)
Beringian	Island Lake (6), Toolik Lake (6)
Nahanni	Ugashik Lake (4), Kusawa Lake (4)
Montana	Waterton Lake (4)

Figure 2.2. Refugia lake trout sampling locations across Canada. Lakes are indicated with different coloured markers. A – Lac Archambault (Quebec), B – Hawley Lake (Ontario), C – Seneca Lake (New York), D – Confederation Lake (Ontario), E – Island Lake (Alaska), F – Toolik Lake (Alaska), G – Ugashik Lake (Alaska), H – Kusawa Lake (Yukon), I – Waterton Lake (Alberta).

2.2 DNA Extraction for Refugia Samples

Samples received in ethanol or as a dried fin clip were soaked in TE buffer (100mM Tris, 1mM EDTA) for 1-2 hours, followed by an overnight incubation in 197µl of fish extraction buffer (100mM Tris, 10mM EDTA, 240mM NaCl, 1% SDS) and 300ug/ml of Proteinase K at 55°C. The next morning, samples were removed from incubation and extracted using phenol-chloroform. One volume of phenol was added, centrifuged for 5min at 15 000 rpm and the top layer removed to a new tube and repeated. This was followed with adding one volume of chloroform, centrifuged for 5min at 15,000 rpm and the top layer removed to a new tube. Following the phenol-chloroform, RNase treatment of the DNA was performed using RNase A (1mg/mL) and incubated in a water bath at 37°C for 30 minutes. DNA was precipitated with

200µl of 2-propanol and incubated overnight at -20°C. Following centrifugation at 15,000 rpm for 15 minutes the DNA was resuspended in 400µl of TE. The DNA was then precipitated again using 1/10 volume of 3M Acetic acid (CH_2COONa pH 5.2) and 2 volumes of ethanol. The tubes were inverted a few times to mix and incubated overnight at -20°C for maximum yield. After incubation, DNA was pelleted by centrifugation for 15 min at 15,000 rpm and washed twice with 70% ethanol. The final DNA pellet was resuspended in 50 µl of water and stored at 4°C for short-term storage but -20°C for long-term storage.

2.3 PCR and Cloning

2.3.1 Polymerase Chain Reaction

In order to assist with differentiation of the two potential loci that may be present, a region of the intron sequence was included in the amplicon. PCR forward (LT FWD2: 5' –ACT CCT CAA AGG ACC TGC ATG GT– 3') and PCR reverse (LT REV2: 5' –CTC AGT GTT TAA ATA TAG TTG CAG CT– 3') primers were designed to match a conserved region at the beginning of the second exon, and within the second intron. The primers were based on sequences of full-length lake trout MH class II genes previously amplified using PCR forward (SAALDABF: 5' – GAT ACT CCT CAA AGG ACC TG– 3') and PCR reverse (SAALDABR: 5'- CTC AGC CAG GTC ACT CTG – 3') previously used in the laboratory (Conejeros et al., 2012).

Using LTFWD2 and LTREV2, lake trout samples were amplified using polymerase chain reaction (PCR) to give an amplicon of 500 bp in length. For each individual fish two independent amplification reactions were performed with an initial denaturation at 95°C for 2 min, plus 25

cycles of 95°C denaturation for 30 sec, 40 sec of 53°C of primer annealing and 72°C extension for 2 min. A final extension at 72°C for 10 min was performed to ensure the addition of an A overhang at the end of the complementary strand.

2.3.2 Gel Extraction, Ligation and Transformation

The PCR products were separated by gel electrophoresis on a 1% agarose gel and visualized with gel red (Biotium, Hayward, CA, USA) to confirm that the fragments were of the predicted length. The DNA amplicons were then extracted from the agarose gel using the QIAquick Gel Extraction Kit (Qiagen, Mississauga, ON) and ligated into pGEM- T easy vector following the manufacturer's instructions (Promega Corporation, Madison, WI). Following ligation, the products of each individual PCR reaction were transformed into XL1-Blue MRF *Escherichia coli* bacteria and plated on LB agar plates containing ampicillin (100ug/ml), 100ug/ml X-gal and 1uM IPTG (isopropyl-β-D-1-thiogalactopyranoside).

2.3.3 Miniprepation of plasmid DNA

Six to eight single white colonies from each PCR reaction were selected from the plate to be inoculated in an LB broth with ampicillin at a concentration of 100ug/mL. They were grown overnight in a shaker-incubator at 37°C with 200-225 rpm shaking. The next day, inoculated cultures were pelleted by centrifugation at 10,000 rpm for 5 minutes and the supernatant was removed as the pellets were placed on ice. The pellet was resuspended in Solution I (50mM glucose, 25 mM Tris-HCl pH 8.0, 10mM EDTA pH 8.0) by vortexing vigorously followed by the addition of Solution II (0.2M NaOH and 1% SDS). The tubes were mixed by inversion and

placed on ice for 5 minutes. After the addition of Solution III (3M potassium acetate and 11.5% (v/v) acetic acid) tubes were inverted multiple times, incubated on ice for 5 minutes and centrifuged at 15 000 rpm for 5 minutes. The supernatants were transferred to a new tube with the addition of RNase A (1mg/mL) and incubated in a water bath at 37°C for 30 minutes. Following incubation, 1 volume of pre-chilled isopropanol was added and tubes were inverted and incubated overnight in the -20°C to allow for maximum DNA precipitation. The next day, the DNA was pelleted by centrifugation at 15 000 rpm for 15 minutes, washed twice with pre-chilled 70% ethanol and dissolved in 20µl of milliQ water. The DNA solution was heated to 65-70°C for 10 minutes to ensure the DNA had dissolved into solution and concentration measured with a NanoDrop.

2.3.4 Restriction Enzyme Digestion

Digests were performed using the FastDigest® system (Fermentas, Burlington, ON) using the FastDigest® enzyme EcoRI following the manufacturer's instructions. Following digestion, each digest was run on a 1% agarose gel to ensure plasmids contained inserts of proper length.

2.4 Sequencing

Following determination of DNA concentrations, clones that were determined to have proper sized inserts were sent for sequencing. DNA was added to plates and sent to the Trent University DNA Sequencing Lab (Trent University, Peterbourgh, ON) or the Centre for Applied Genomics (SickKids, Toronto, ON) for sequencing. Plasmids were sequenced in both directions using the T7 and Sp6 primers.

2.5 Data Analysis

The sequences obtained were aligned using the software Muscle (Edgar, 2004). An alignment with previously sequenced full length class IIβ sequences from lake trout (Dorschner et al. 2000; Noakes et al. 2003) was used as a guide for locating the end of the second exon and beginning of the intron. Bioedit (Hall, 1999) was used to find and correct sequences containing singleton errors, define the open reading frames and trim sequences to correct length. DnaSP v5 (Librado and Rozas, 2009) was used to calculate the haplotype and nucleotide diversities of the sequences among and within the populations using Nei model (Nei, 1987). Mega 5 (Tamura et al., 2011) was used to calculate the rates of non-synonymous/synonymous substitutions using the Nei-Gojobori model (Nei and Gojobori, 1986) using the Jukes-Cantor correction.

2.6 Phylogeographic Analysis

Sequences of all unique alleles from refugial and study populations were analyzed using jModelTest (Posada, 2008) to carry out statistical selection of the model of best-fit for nucleotide substitution. The best model of evolution was used to compute neighbor-joining, maximum parsimony and maximum likelihood trees using PAUP*4.0b10 (Swofford, 2003). The neighbor-joining and maximum parsimony trees were bootstrapped 1000 times and values were added to branches that were consistent with the maximum likelihood tree.

2.7 Network Analysis

Genealogical relationships among nucleotide sequences were analyzed by constructing a haplotype network using the parsimony method of Templeton et al. (1992). It estimates the maximum number of substitutions to connect parsimoniously two alleles with a confidence connection limit or a maximum number of connection limit steps beginning with linking sequences with the smallest number of differences. It also estimates the most likely allele outgroup allowing the identification of the most ancestral alleles in the study. The analysis was performed with the software TCS vs. 1.21 (Clement *et al*. 2000) using a 95% confidence interval and gaps as a 5th state.

Chapter 3 – Results

3.1 Allelic Sequences of Ontario/New York populations

A total of fifty-seven unique MH class II β1 alleles were identified in 215 individual lake trout from 10 lakes within Ontario and one in New York with a 0.0648 intrapopulation diversity. The 234 nt exon sequence (Figure 3.1) encoded 78 amino acids from the β1 domain (Figure 3.2). Of the 57 alleles, 32 were unique and within one lake while 25 alleles were common throughout the multiple lakes in the study (Table 3.1). Slate Islands had the highest allelic richness (2.53) and Macdonald has the lowest of 1.74 (Table 3.2). Manitou Lake had the highest haplotype diversity (0.912) and Seneca lake had the highest nucleotide diversity (0.0719) while Kingscote had the lowest nucleotide and haplotype diversities (Table 3.2). Only two common alleles were identified within all the populations studied, Sana-DAB 4 which had an allele frequency of 0.243 within the populations as a whole and occurred at a frequency as high as 0.375 within an individual population, and Sana-DAB 5 which had an allele frequency of 0.126 within the populations as a whole and as high as 0.327 within a single population (Table 3.3).

Five alleles in this study (Sana-DAB 3, 11, 20, 21 and 48) were in common with the Dorschner et al., (2000) study on lake trout within Lake Superior. The common MH class II β allele nucleotide sequences were in GenBank with the accession numbers: AF13008, AF129994, AF129978, AF130030 and AF130010 (Figure 3.1).

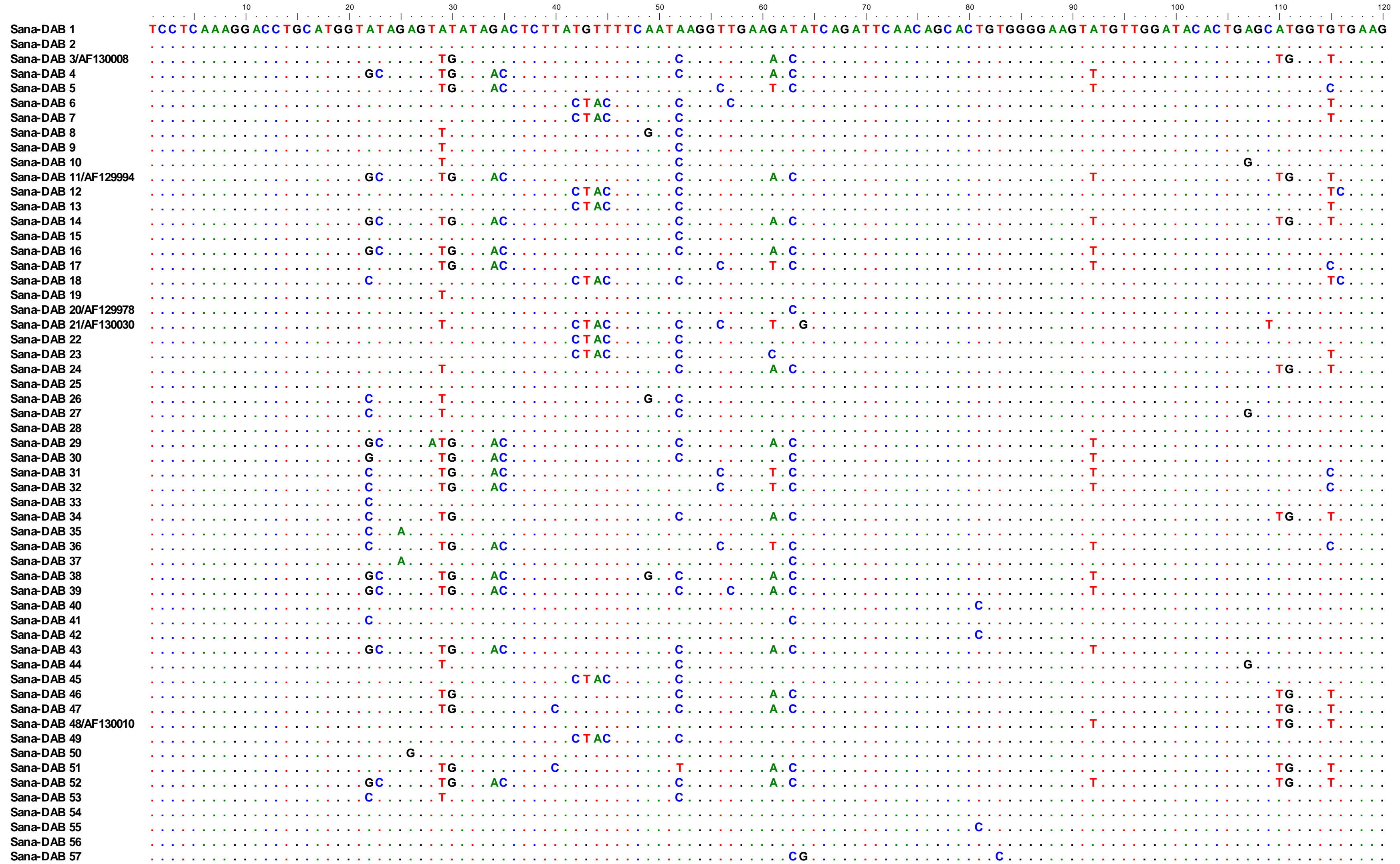


Figure 3.1. Nucleotide sequences of MH class II β1 unique alleles in lake trout from 11 populations. GenBank accession numbers are listed for the five common alleles in the Dorschner et al., (2000) study.

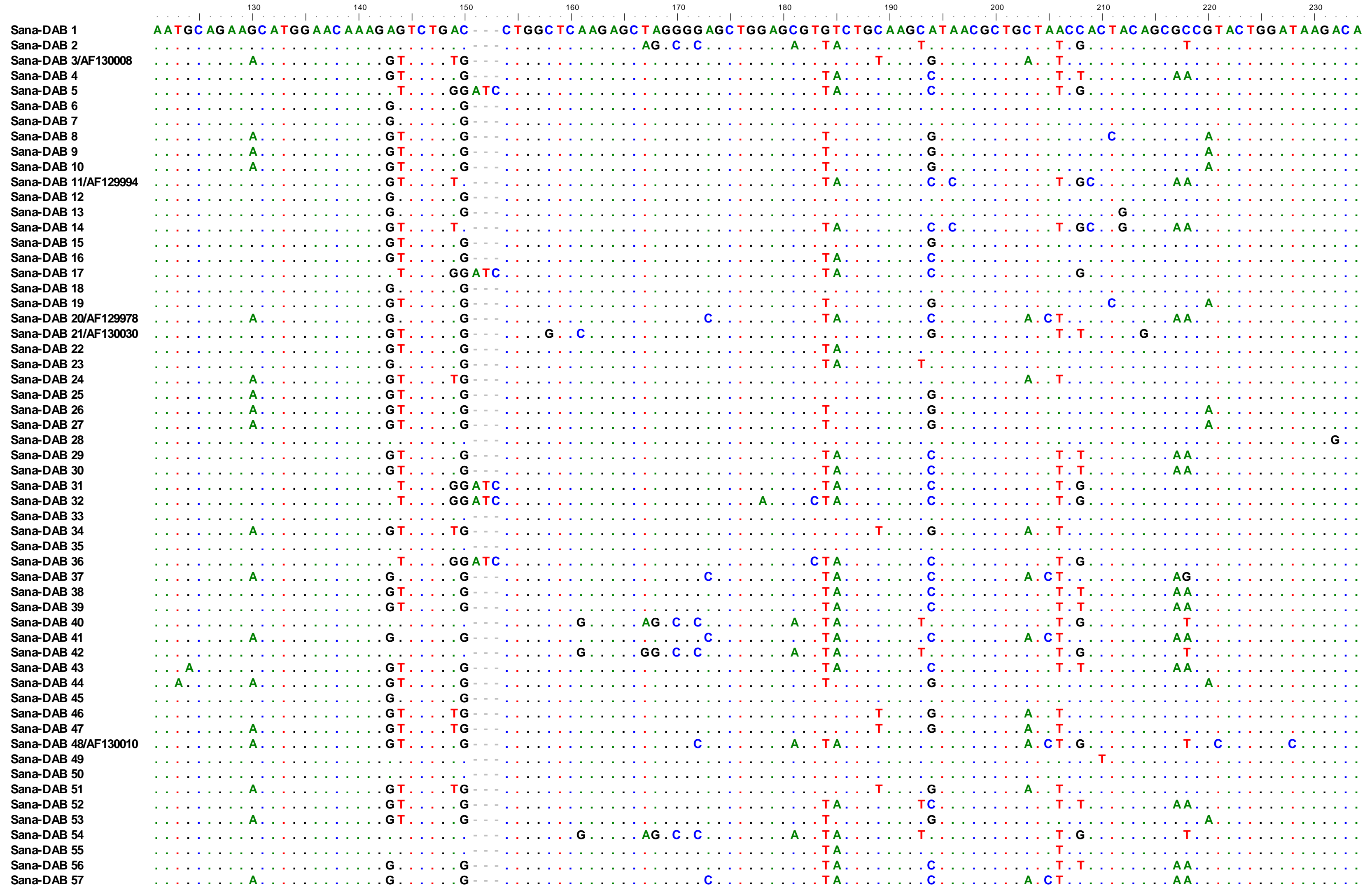


Figure 3.1 continued. Nucleotide sequences of MH class II β1 unique alleles in lake trout from 11 populations. GenBank accession numbers are listed for the five common alleles in the Dorschner et al., (2000) study.

Figure 3.2. Protein sequences of MH class II β1 unique alleles in lake trout from 11 populations. GenBank accession numbers are listed for the five common alleles in the Dorschner et al., (2000) study.

Table 3.1. Allelic diversity at Sana-DAB1 of lake trout populations studied. Ancestry was determined from previous studies on lake trout populations and hatchery strains (Grewe and Hebert 1988; Grewe et al. 1993; Halbisen and Wilson 2009; Stott 1998; Wilson and Hebert 1996, 1998).

Population	# of Fish	Fish with >2 alleles	Alleles Present	Total Alleles	Unique Alleles	Stocked (Yes, No or Source)	Ancestry (based on mtDNA/µsat)
Hogan Lake	21	4	1,2,3,4,5,6,7,8,9,10,11,43,44	13	2	No	Mississippian and Atlantic
Simcoe Lake	20	2	1,3,4,5,7,9,10,11,33,45	10	0	Yes	Diverse ancestry
Crystal Lake	20	10	1,3,4,5,7,8,9,10,12,13,14,15,16,17,31,46,47	17	5	No	Mississippian
Michipicoten Lake	20	5	1,3,4,5,7,8,9,11,14,18,46,47	12	1	Stocking Source	Diverse ancestry
Opeongo Lake	20	9	1,2,4,5,9,11,14,19,20,21,22,23,24,48,49	15	7	Yes	Mississippian and Atlantic
Slate Islands	19	10	1,3,4,5,7,14,25,26,27,33,47,50,51,52,53	15	5	Stocking Source	Diverse ancestry
Dickey Lake	20	6	1,4,5,7,10,20,27,28,29,30,31,32,45,53,54,55,56	17	5	No	Mississippian
Manitou Lake	20	6	2,3,4,5,7,8,10,20,27,31,32,33,34,35,43,54,57	17	2	Stocking Source	Diverse ancestry
Kingscote Lake	18	4	1,4,5,8,20,36,37,38,57	9	3	Yes	Mississippian
Seneca Lake	18	5	1,4,5,11,14,20,32,39,40,41,54,57	12	1	Stocking Source	Mississippian and Atlantic
Macdonald Lake	19	2	4,5,20,39,40,42,57	7	1	No	Mississippian
All Populations	215	63		144	32		

Table 3.2. Allelic richness (A_R), haplotype diversity (h), nucleotide diversity (π), non-synonymous (d_N) and synonymous (d_S) substitutions per population in lake trout MH class II β1 sequences.

Population	Allelic Richness (A_R)	Haplotype Diversity (h)	Nucleotide Diversity (π)	Non-Synonymous Substitutions (d_N)	Synonymous Substitutions (d_S)	D_N/d_S Ratio (K_a/K_s)
Hogan Lake	1.86	0.842	0.0579	0.0828	0.0421	1.97
Simcoe Lake	1.8	0.849	0.0593	0.0781	0.0424	1.84
Crystal Lake	2.45	0.887	0.0604	0.0758	0.0424	1.79
Michipicoten Lake	2.1	0.820	0.0600	0.0831	0.0476	1.75
Opeongo Lake	2.45	0.844	0.0658	0.0936	0.0479	1.95
Slate Islands	2.53	0.833	0.0568	0.0761	0.0333	2.35
Dickey Lake	2.2	0.893	0.0574	0.0769	0.0324	2.37
Manitou Lake	2.25	0.912	0.0620	0.0893	0.0358	2.49
Kingscote Lake	1.89	0.788	0.0548	0.0830	0.0192	4.32
Seneca Lake	1.89	0.879	0.0719	0.0937	0.0366	2.56
Macdonald Lake	1.74	0.809	0.0641	0.0905	0.0403	2.25
Averages	2.11	0.851	0.0609	0.0839	0.0382	2.33

Table 3.3. Allelic frequencies of all 57 unique alleles per population and study.

Allele Name	Hogan Lake	Simcoe Lake	Crystal Lake	Michipicoten	Opeongo	Slate Islands	Dickey	Manitou	Kingscote	Seneca	Macdonald	Total
Sana-DAB 1	0.128	0.056	0.061	0.061	0.041	0.063	0.091		0.029	0.029		0.053
Sana-DAB 2	0.026				0.102			0.022				0.015
Sana-DAB 3	0.128	0.028	0.082	0.102		0.042		0.022				0.040
Sana-DAB 4	0.359	0.306	0.306	0.327	0.204	0.375	0.205	0.200	0.029	0.147	0.061	0.243
Sana-DAB 5	0.077	0.111	0.082	0.082	0.327	0.063	0.182	0.089	0.265	0.029	0.030	0.126
Sana-DAB 6	0.026											0.002
Sana-DAB 7	0.026	0.194	0.061	0.122		0.042	0.023	0.044				0.049
Sana-DAB 8	0.026		0.061	0.041				0.022	0.324			0.040
Sana-DAB 9	0.051	0.056	0.041	0.020	0.020							0.018
Sana-DAB 10	0.051	0.139	0.041				0.023	0.022				0.024
Sana-DAB 11	0.026	0.056		0.020	0.020					0.029		0.013
Sana-DAB 12			0.061									0.007
Sana-DAB 13			0.020									0.002
Sana-DAB 14			0.041	0.020	0.020	0.042				0.206		0.029
Sana-DAB 15			0.020									0.002
Sana-DAB 16			0.020									0.002
Sana-DAB 17			0.041									0.004
Sana-DAB 18				0.020								0.002
Sana-DAB 19					0.020							0.002
Sana-DAB 20					0.041		0.182	0.089	0.235	0.029	0.121	0.060
Sana-DAB 21					0.061							0.007
Sana-DAB 22					0.020							0.002
Sana-DAB 23					0.020							0.002
Sana-DAB 24					0.020							0.002
Sana-DAB 25						0.021						0.002
Sana-DAB 26						0.021						0.002
Sana-DAB 27						0.146	0.068	0.067				0.029
Sana-DAB 28							0.023					0.002
Sana-DAB 29							0.023					0.002
Sana-DAB 30							0.023					0.002

Allele Name	Hogan Lake	Simcoe Lake	Crystal Lake	Michipicoten	Opeongo	Slate Islands	Dickey	Manitou	Kingscote	Seneca	Macdonald	Total
Sana-DAB 31			0.020				0.023	0.089				0.013
Sana-DAB 32							0.023	0.178		0.176		0.033
Sana-DAB 33		0.028				0.021		0.022				0.007
Sana-DAB 34								0.022				0.002
Sana-DAB 35								0.022				0.002
Sana-DAB 36									0.029			0.002
Sana-DAB 37									0.029			0.002
Sana-DAB 38									0.029			0.002
Sana-DAB 39										0.059	0.242	0.022
Sana-DAB 40										0.206	0.303	0.038
Sana-DAB 41										0.029		0.002
Sana-DAB 42											0.030	0.002
Sana-DAB 43	0.051							0.022				0.007
Sana-DAB 44	0.026											0.002
Sana-DAB 45		0.028					0.023					0.004
Sana-DAB 46			0.020	0.020								0.004
Sana-DAB 47			0.020	0.020		0.021						0.007
Sana-DAB 48					0.061							0.007
Sana-DAB 49					0.020							0.002
Sana-DAB 50						0.042						0.004
Sana-DAB 51						0.021						0.002
Sana-DAB 52						0.021						0.002
Sana-DAB 53						0.042	0.023					0.007
Sana-DAB 54							0.023	0.022		0.029		0.007
Sana-DAB 55							0.023					0.002
Sana-DAB 56							0.023					0.002
Sana-DAB 57								0.044	0.029	0.029	0.212	0.024

3.2 Refugia Sequences

A total of 35 fish from five different refugia groups were analyzed (Table 3.4) finding 2 Mississippian alleles, 8 Atlantic alleles, 10 Beringian alleles, 9 Nahanni alleles and 2 Montana alleles. These 41 alleles were derived from the same 234 nt exon sequence amplified in the Ontario/New York populations above (Figure 3.3 and 3.4) and common alleles were seen in the Atlantic/Beringian and Beringian/Nahanni refugial groups.

Table 3.4. Refugia allelic diversity at Sana-DAB1.

Population	# of Fish	Refugia Present based on mtDNA	Total Alleles	Unique Alleles
Lac Archambault	1	Atlantic	1	1
Seneca Lake	5	Atlantic, Mississippian	10	6
Confederation Lake	1	Atlantic, Mississippian	2	1
Hawley Lake	4	Mississippian	4	0
Island Lake	6	Beringian	13	3
Toolik Lake	6	Beringian	8	2
Ugashik Lake	4	Nahanni	8	1
Kusawa Lake	4	Nahanni	7	1
Waterton Lake	4	Montana	7	2
All Populations	35		60	17

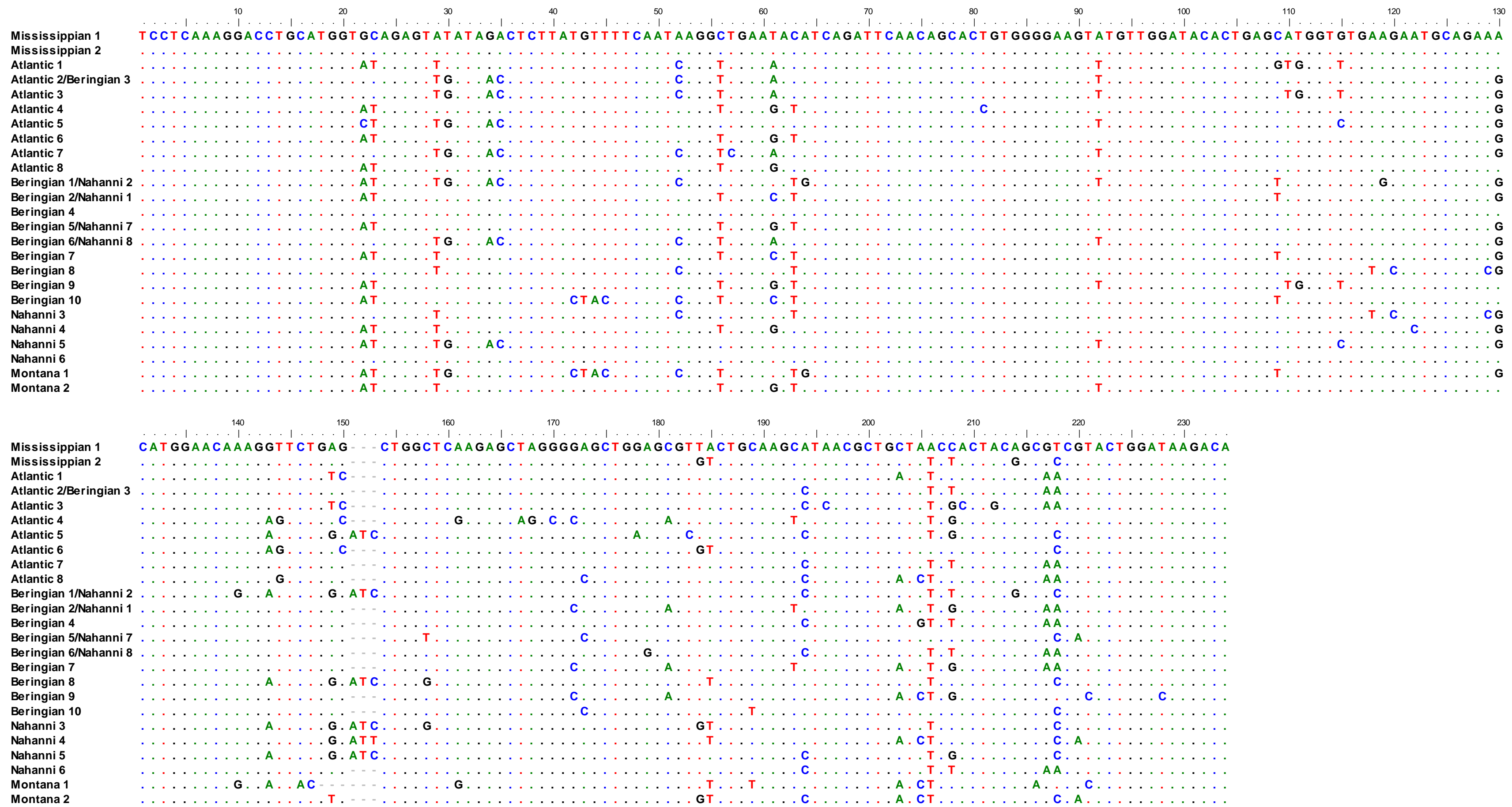


Figure 3.3. MH Class II β1 alleles in lake trout from Wisconsinan glacial refugia mtDNA lineages.

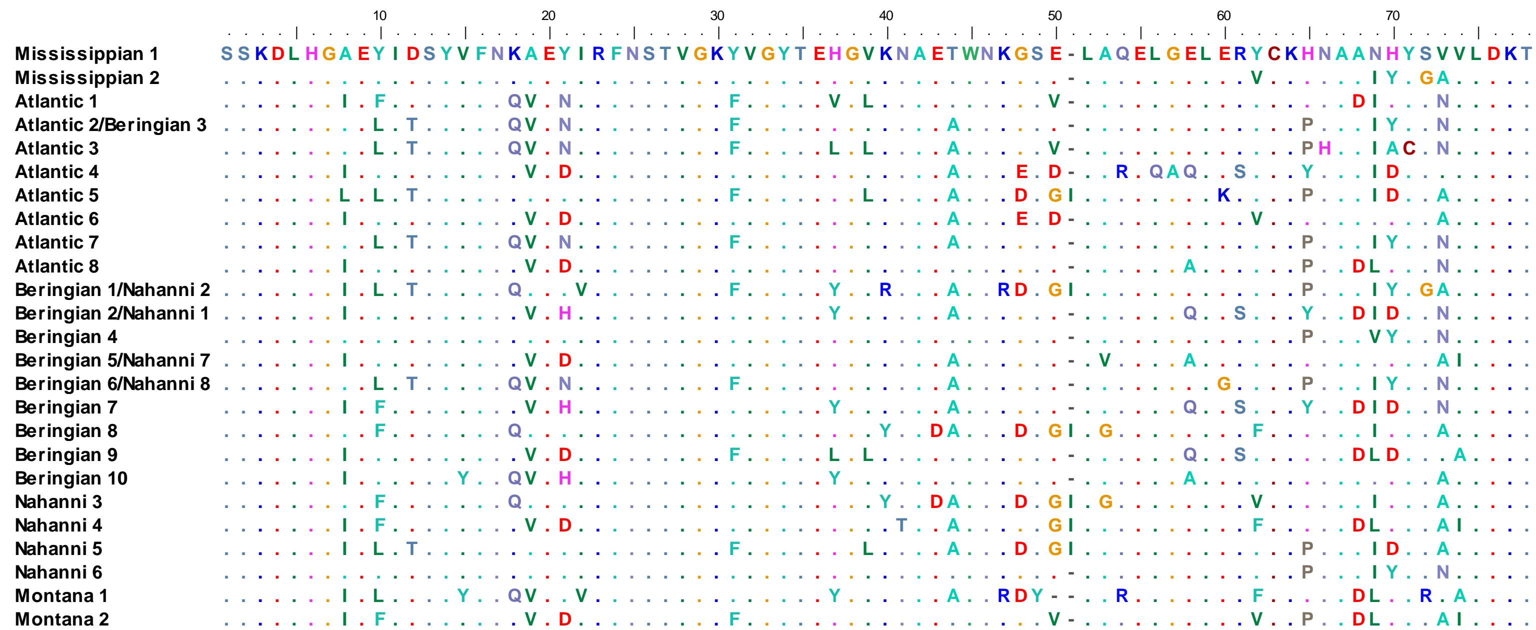


Figure 3.4. Protein sequences of the MH class II β1 alleles in lake trout from putative refugial groups.

3.3 Selection

A total D_N/D_S (non-synonymous/synonymous substitutions) ratio study average of 2.33 (Table 3.2) suggests that positive selection may be acting on the polymorphic residues (Hurst, 2002). Most populations exhibited ratios greater than 1 with the highest of 4.32 within Kingscote Lake indicating stronger positive selection. No populations exhibited a ratio below one, suggesting stabilizing selection.

3.4 Phylogenetic and Network Inferences

A phylogenetic tree and a parsimony network were constructed to look at all the alleles in this study and within each individual population. A maximum likelihood tree with all alleles in the study along with the refugia alleles (Figure 3.5) did not group alleles clearly into refugia groups with any bootstrap support or confidence. Similarly, the parsimony network of all study and refugia alleles (Figure 3.6) did not show any type of refugia pattern although common alleles from Atlantic, Beringian and Nahanni refugia groups were seen within the study populations.

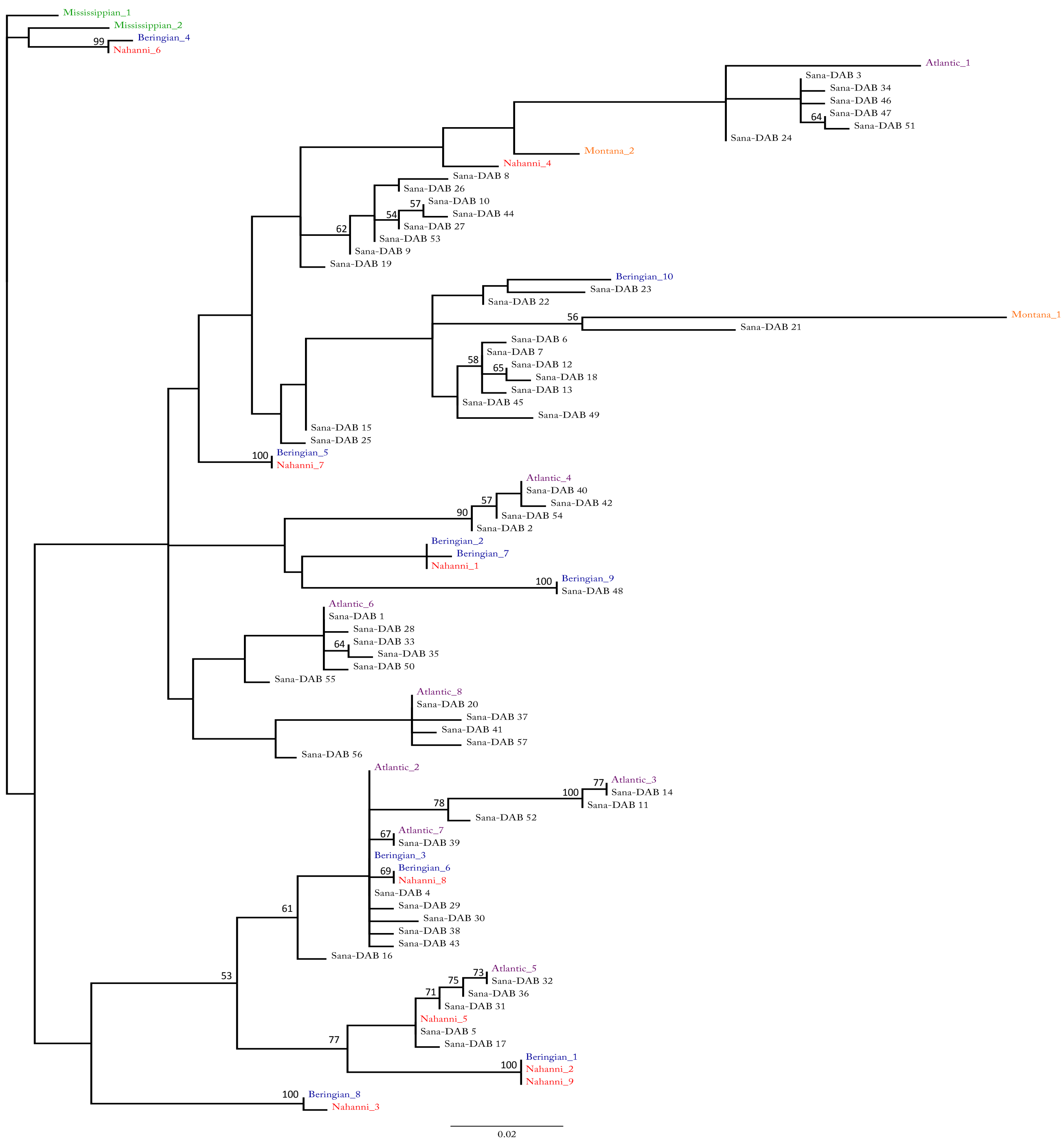


Figure 3.5. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 unique and refugia nucleotide sequences of lake trout within this study. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining analyses to branches that are consistent with the likelihood tree. Scale bar indicates nucleotide divergence of 0.02 (2%).

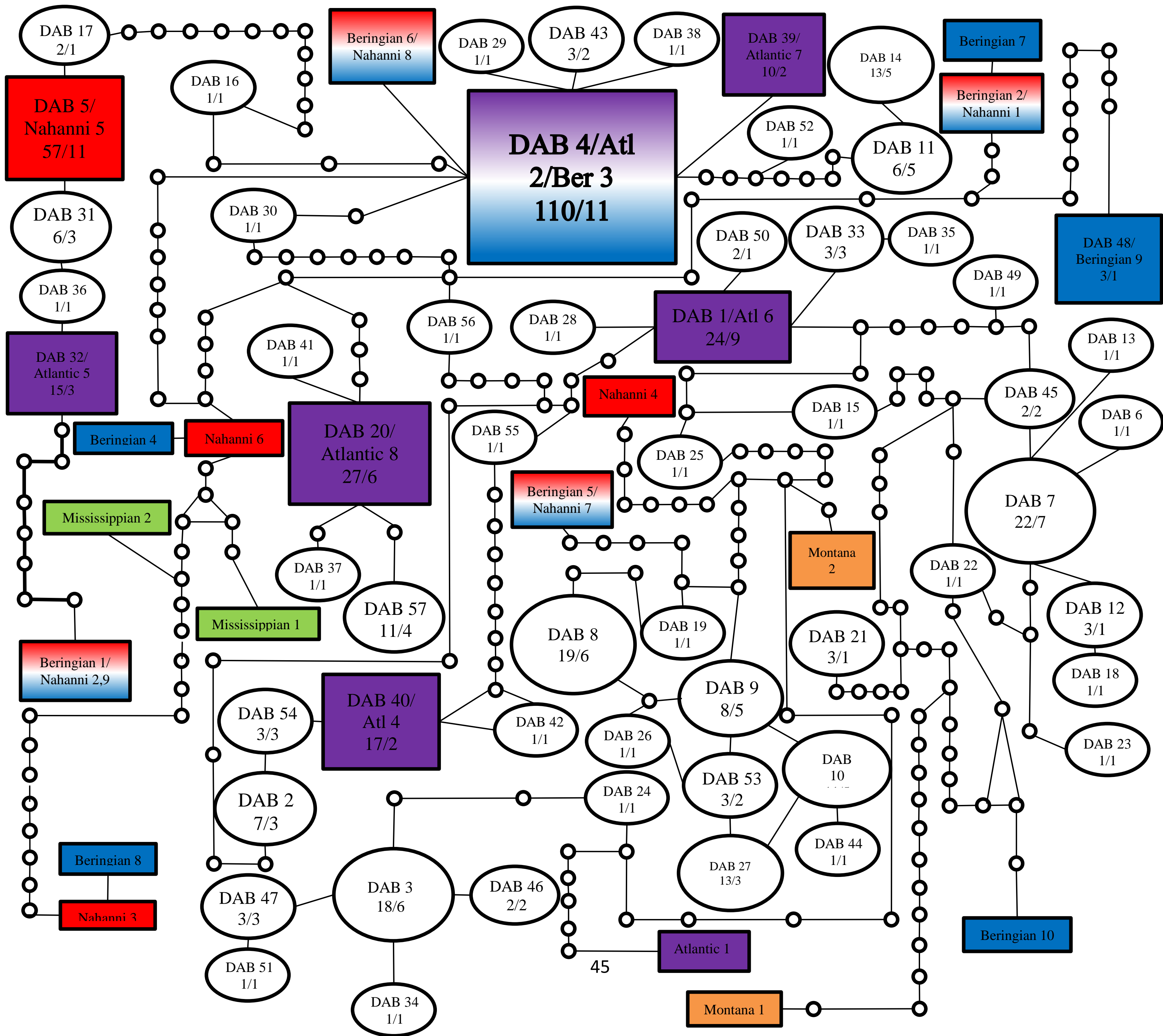
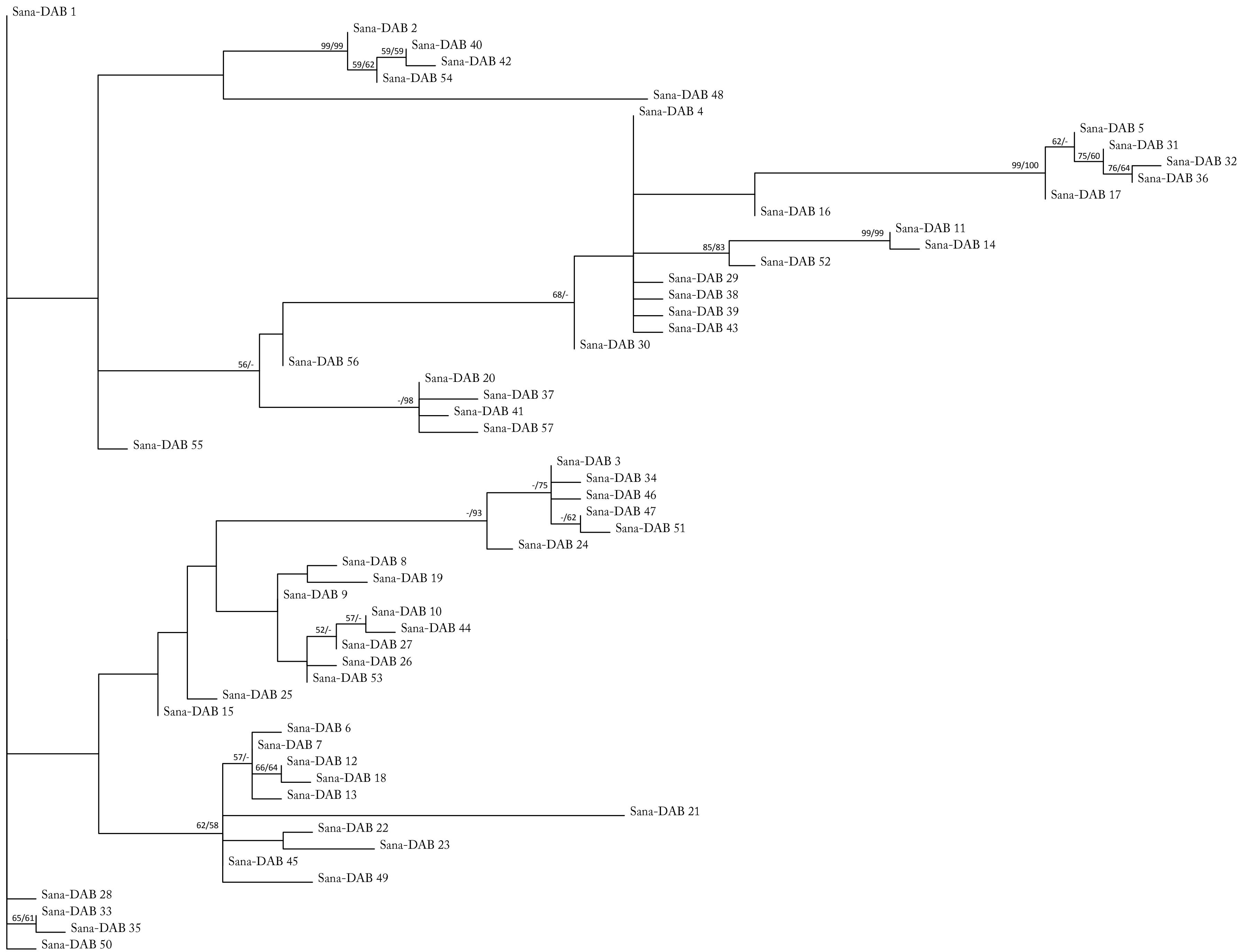

DAB 17
2/1
DAB 5/
Nahanni 5
57/11
DAB 16
1/1
Beringian 6/
Nahanni 8
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39/
Atlantic 7
10/2
DAB 14
13/5
Beringian 7
Beringian 2/
Nahanni 1
DAB 4/Atl
2/Ber 3
110/11
DAB 52
1/1
DAB 11
6/5
DAB 31
6/3
DAB 30
1/1
DAB 50
2/1
DAB 33
3/3
DAB 35
1/1
DAB 48/
Beringian 9
3/1
DAB 36
1/1
DAB 49
1/1
DAB 56
1/1
DAB 28
1/1
DAB 1/Atl 6
24/9
DAB 32/
Atlantic 5
15/3
DAB 41
1/1
DAB 13
1/1
Nahanni 4
DAB 15
1/1
DAB 45
2/2
DAB 6
1/1
Beringian 4
Nahanni 6
DAB 20/
Atlantic 8
27/6
DAB 55
1/1
DAB 25
1/1
Beringian 5/
Nahanni 7
DAB 7
22/7
Mississippian 2
Montana
2
DAB 37
1/1
DAB 22
1/1
DAB 57
11/4
DAB 8
19/6
DAB 19
1/1
DAB 21
3/1
DAB 12
3/1
Mississippian 1
Beringian 1/
Nahanni 2,9
DAB 18
1/1
DAB 9
8/5
DAB 54
3/3
DAB 40/
Atl 4
17/2
DAB 42
1/1
DAB 26
1/1
DAB 23
1/1
DAB
10
DAB 53
3/2
DAB 2
7/3
DAB 24
1/1
DAB 44
1/1
DAB 27
13/3
Beringian 8
DAB 47
3/3
DAB 3
18/6
DAB 46
2/2
Nahanni 3
Atlantic 1
Beringian 10
DAB 51
1/1
DAB 34
1/1
Montana 1

Figure 3.6. Network of MH class II β1 study and refugia nucleotide sequences of lake trout. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

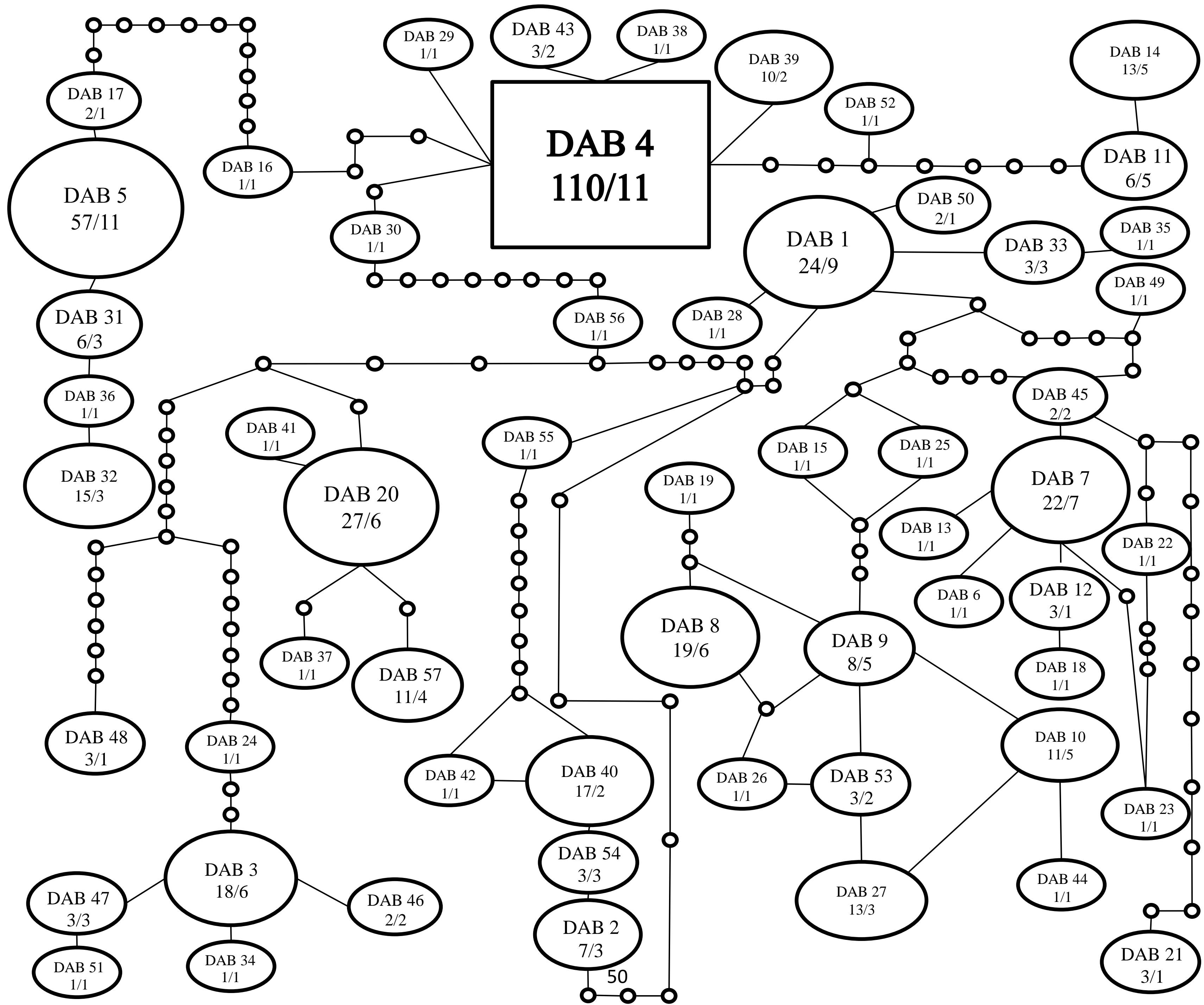
A maximum likelihood tree of the fifty-seven alleles in the eleven lakes (Figure 3.7) separated the alleles into two main clades when examined separately from the refugia alleles. In all 30% of fish from the populations contained three or more alleles (Table 3.1) and the formation of two major branches within Figure 3.7 suggests the possibility of multiple loci within lake trout. A parsimony network of all alleles (Figure 3.8a) was designed to show the differences and similarities among all alleles throughout the populations in the study. From this network, approximately nine different groupings of the alleles can be seen as demonstrated in Figure 3.8b grouped together based upon the mutational differences between alleles.

Using only the refugial alleles, a phylogenetic tree was computed to show whether MH class II β alleles across Canada exhibit a phylogeographic signal (Figure 3.9a). Since it was difficult to interpret the maximum likelihood tree with all the refugia alleles, a maximum likelihood tree based on the most common alleles by highest frequency (Figure 3.9b) was created and demonstrated a cleaner separation of alleles by geographical region. The Mississippian alleles grouped separately from the other refugia which separated into two clades. One included Atlantic, Beringian and Nahanni alleles while the other clade included Beringian, Nahanni and Montana alleles. Little bootstrap confidence or certainty supported these clades and groupings within the clades. A parsimony network was not designed for the refugia alleles due to the large number of mutational steps between many of the alleles.



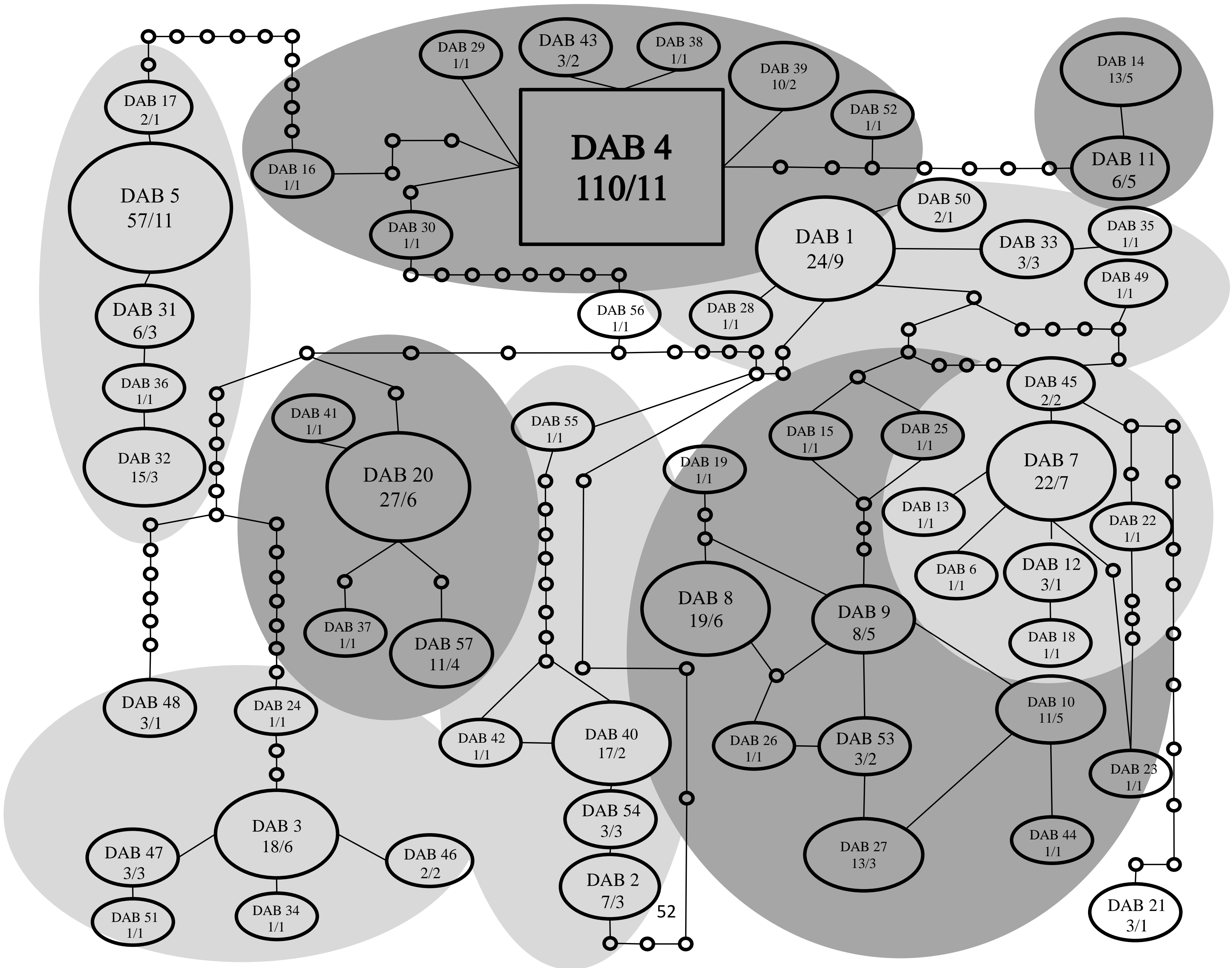
Sana-DAB 1
99/99
Sana-DAB 2
59/59
Sana-DAB 40
59/62
Sana-DAB 42
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
99/100
Sana-DAB 32
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
68/-
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
-/62
Sana-DAB 47
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.7. Unrooted maximum likelihood phylogenetic tree of the unique MH class II β1 nucleotide sequences of lake trout within this study. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).



DAB 4
110/11
DAB 5
57/11
DAB 1
24/9
DAB 20
27/6
DAB 7
22/7
DAB 8
19/6
DAB 3
18/6
DAB 40
17/2
DAB 32
15/3
DAB 14
13/5
DAB 27
13/3
DAB 57
11/4
DAB 10
11/5
DAB 39
10/2
DAB 9
8/5
DAB 2
7/3
DAB 11
6/5
DAB 31
6/3
DAB 43
3/2
DAB 33
3/3
DAB 53
3/2
DAB 54
3/3
DAB 47
3/3
DAB 48
3/1
DAB 12
3/1
DAB 21
3/1
DAB 17
2/1
DAB 50
2/1
DAB 45
2/2
DAB 46
2/2
DAB 29
1/1
DAB 38
1/1
DAB 52
1/1
DAB 16
1/1
DAB 30
1/1
DAB 35
1/1
DAB 49
1/1
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 19
1/1
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 18
1/1
DAB 37
1/1
DAB 24
1/1
DAB 42
1/1
DAB 26
1/1
DAB 23
1/1
DAB 44
1/1
DAB 51
1/1
DAB 34
1/1
50

Figure 3.8a. Network of MH class II β1 study nucleotide sequences of lake trout. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).



DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 17
2/1
DAB 4
110/11
DAB 11
6/5
DAB 16
1/1
DAB 5
57/11
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 32
15/3
DAB 20
27/6
DAB 19
1/1
DAB 7
22/7
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
52
DAB 51
1/1
DAB 34
1/1
DAB 21
3/1

Figure 3.8b. Network of MH class II β1 study nucleotide sequences of lake trout. Nine allele groups based upon mutational differences are separated by shadings above. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

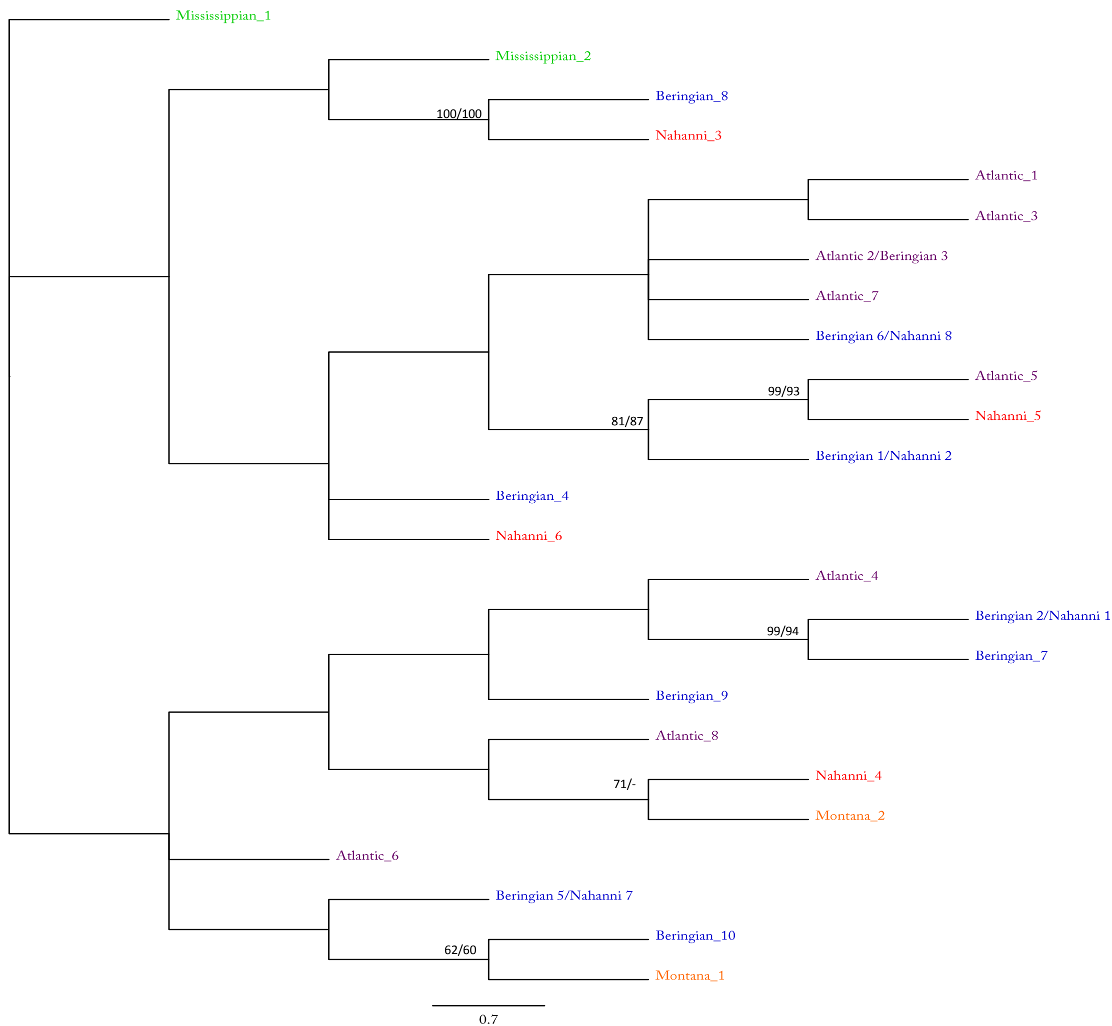


Figure 3.9a. Unrooted maximum likelihood phylogenetic tree of refugial MH class II β1 nucleotide sequences of lake trout within this study. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.7 (70%).

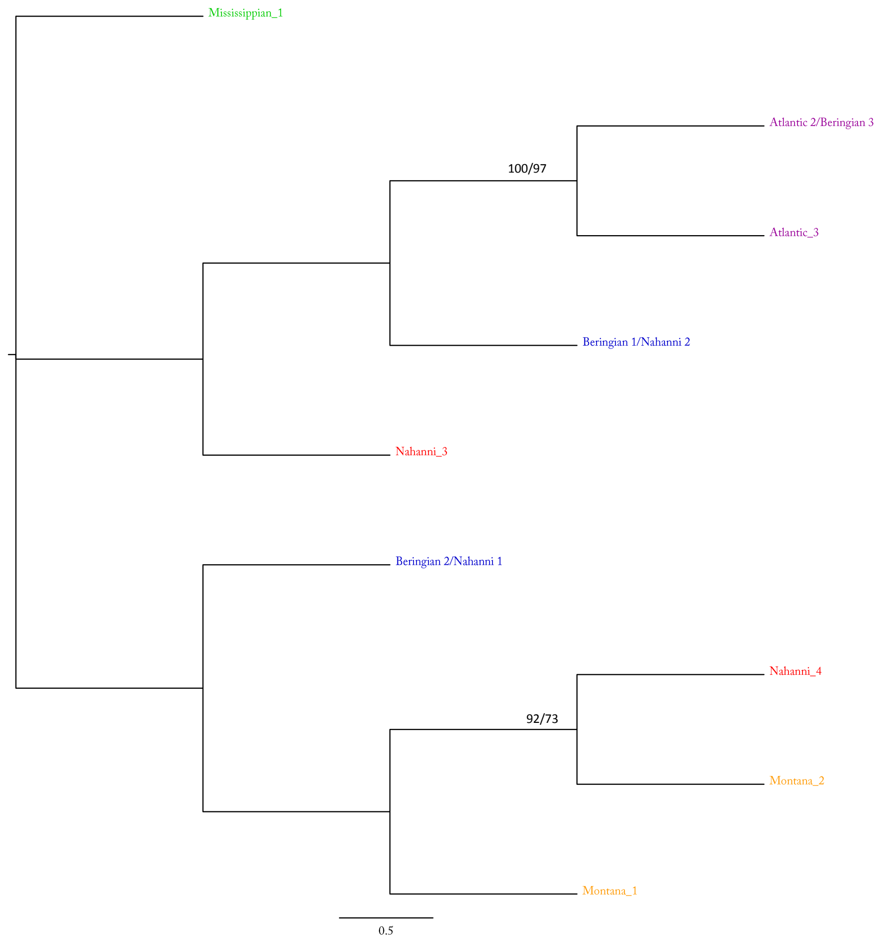


Figure 3.9b. Unrooted maximum likelihood phylogenetic tree of refugial most common MH class II β1 nucleotide sequences of lake trout within this study This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.5 (50%).

3.5 Native Populations – Hogan, Crystal, Dickey and Macdonald

Each of the four native populations (Hogan, Crystal, Dickey and Macdonald) contained some introgressed alleles in other populations but each lake had unique features of MH diversity. None of these native populations have any history of stocking. Crystal and Dickey Lake had the most unique alleles among native populations with five while Crystal had the highest allelic richness. Macdonald Lake had the highest nucleotide diversity with the fewest alleles but only contained alleles from 4 allele groupings from the parsimony network, the lowest amount within the entire study. Dickey Lake had fewer mutational steps between distant alleles when compared to Hogan, Crystal and Macdonald and contained the highest haplotype diversity among native populations.

3.5.1 Hogan Lake

Hogan Lake is a native population in Algonquin Park, Ontario colonized by Mississippian and Atlantic sources (Halbisen and Wilson, 2009). A total of 13 MH class II β1 alleles were discovered in the population with an allelic richness of 1.86 but only two were unique. Three common alleles (Sana-DAB 1, 3 and 4) make up over 50% of allele frequencies within the population with Sana-DAB 4 having the highest frequency at 0.359 (Table 3.3). The nucleotide diversity (0.0579) and haplotype diversity (0.842) were both below the mean diversity values of all the populations studied (Table 3.2). Non-synonymous substitutions of 0.0828 and the D_N/D_S ratio of 1.97 were below the mean values while the synonymous substitutions of 0.0421 was above the study mean (Table 3.2). The alleles of Hogan Lake occurred in groups or pairs with high bootstrap support within the phylogenetic tree (Figure 3.10). The most common

alleles in the population were located in separate clusters isolated from the others. The parsimony network of Hogan Lake had alleles from ~ 8 different groupings which were separated by as few as 8 mutational steps to as many as 28 mutational steps (Figure 3.11). The three common alleles Sana-DAB 1, 3 and 4 differed from each other between 20 – 32 mutational steps but within the phylogenetic tree these three most common alleles were divided with Sana-DAB 1 in a different clade from Sana-DAB 3 and 4.

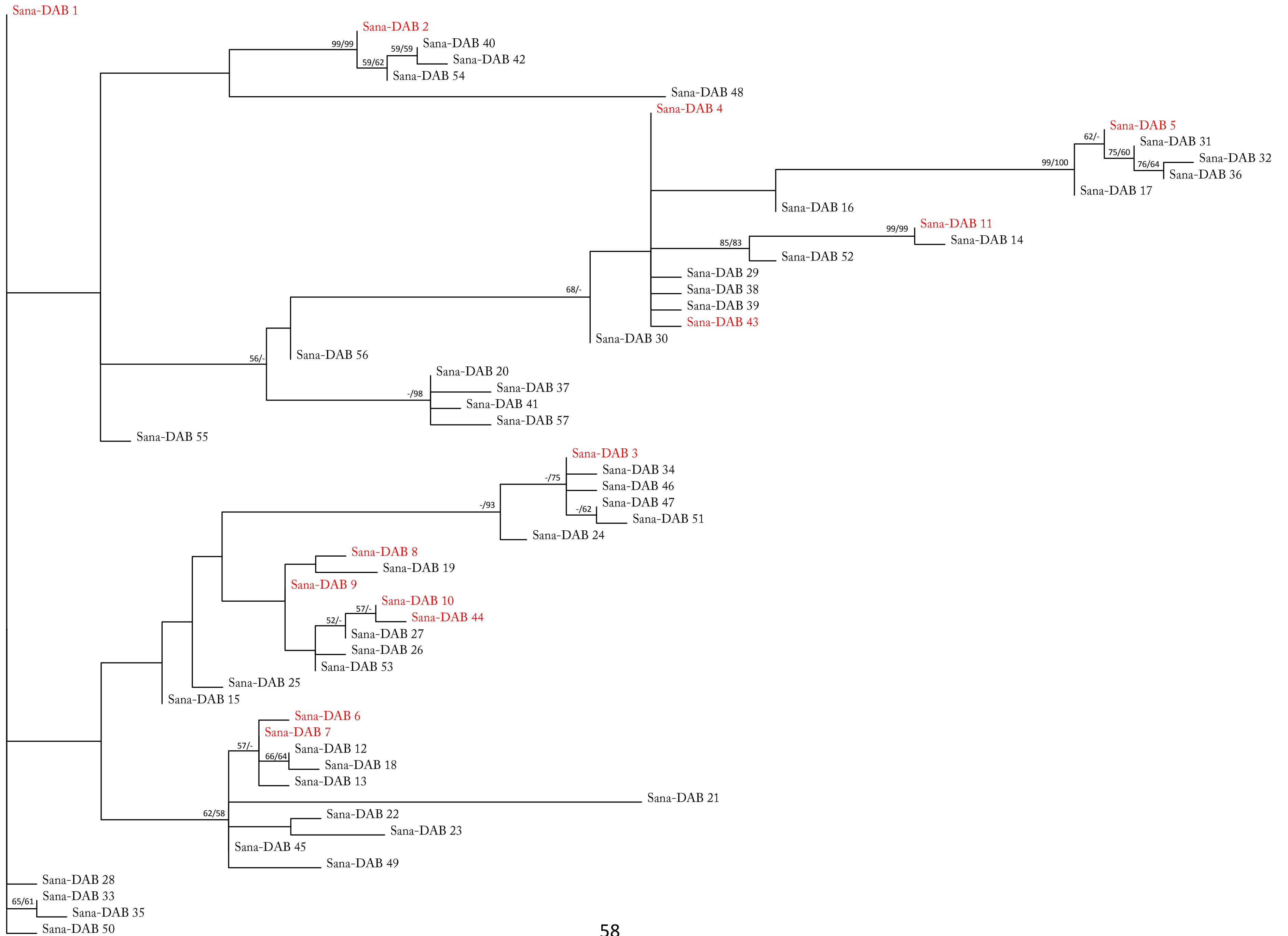
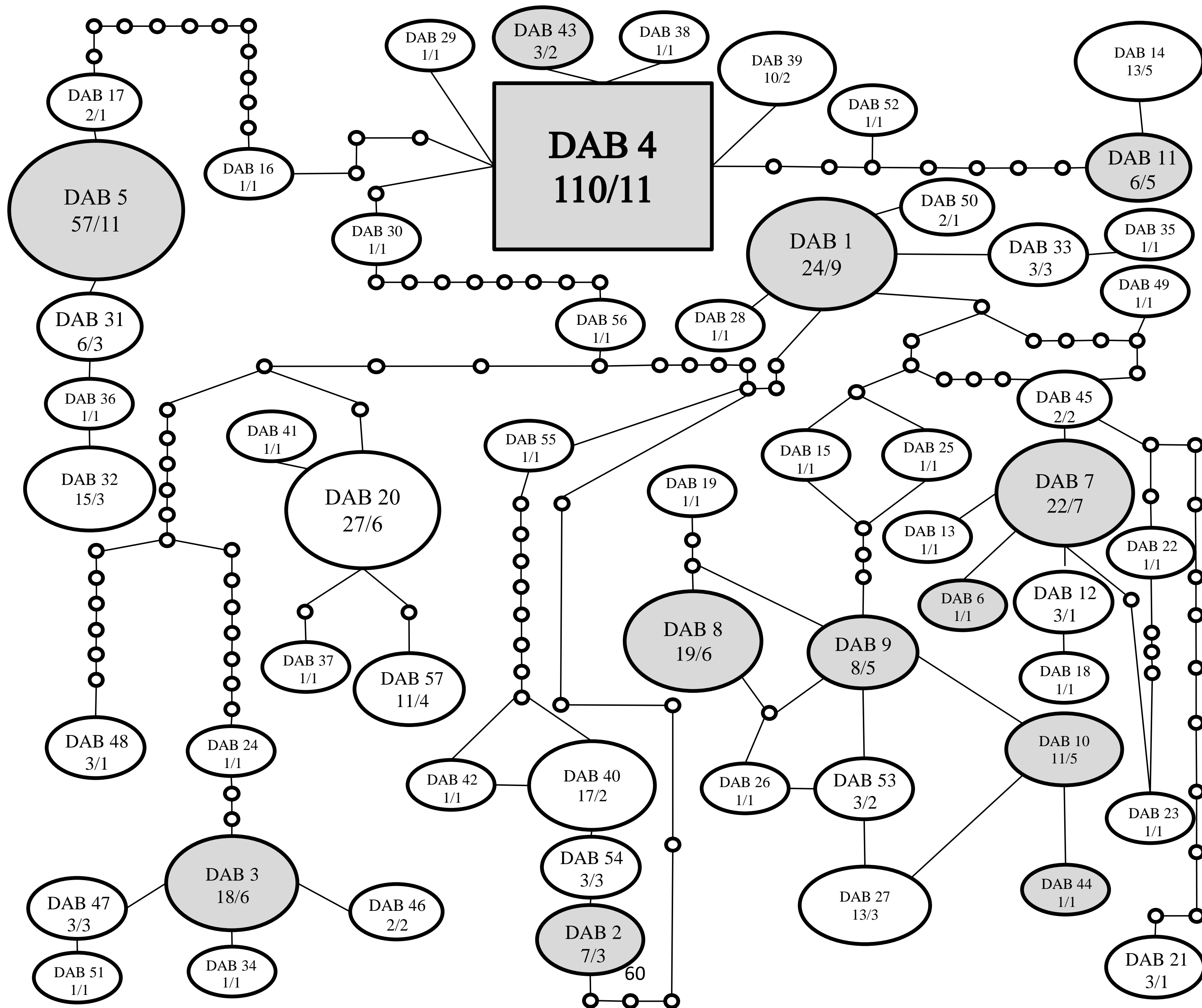

Sana-DAB 1
Sana-DAB 2
Sana-DAB 40
Sana-DAB 42
Sana-DAB 54
99/99
59/59
59/62
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
76/64
Sana-DAB 36
99/100
Sana-DAB 17
Sana-DAB 16
Sana-DAB 11
99/99
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
Sana-DAB 56
56/-
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
Sana-DAB 34
-/75
Sana-DAB 46
Sana-DAB 47
-/93
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.10. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Hogan Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values >50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

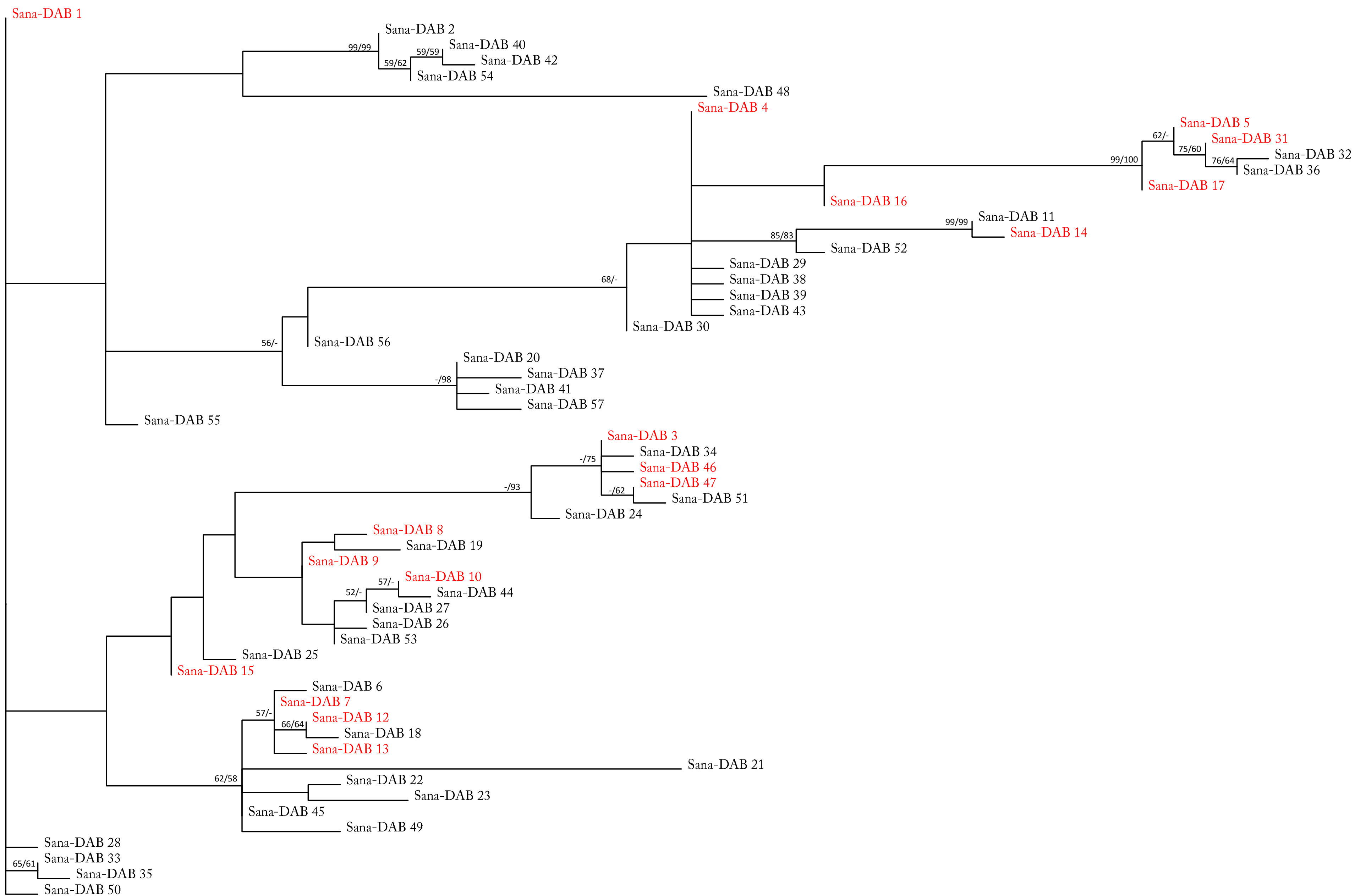


DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 11
6/5
DAB 17
2/1
DAB 16
1/1
DAB 5
57/11
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 32
15/3
DAB 20
27/6
DAB 19
1/1
DAB 7
22/7
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 37
1/1
DAB 57
11/4
DAB 8
19/6
DAB 9
8/5
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 10
11/5
DAB 23
1/1
DAB 3
18/6
DAB 54
3/3
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
DAB 21
3/1

Figure 3.11. Network of MH class II β1 nucleotide sequences of lake trout from Hogan Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

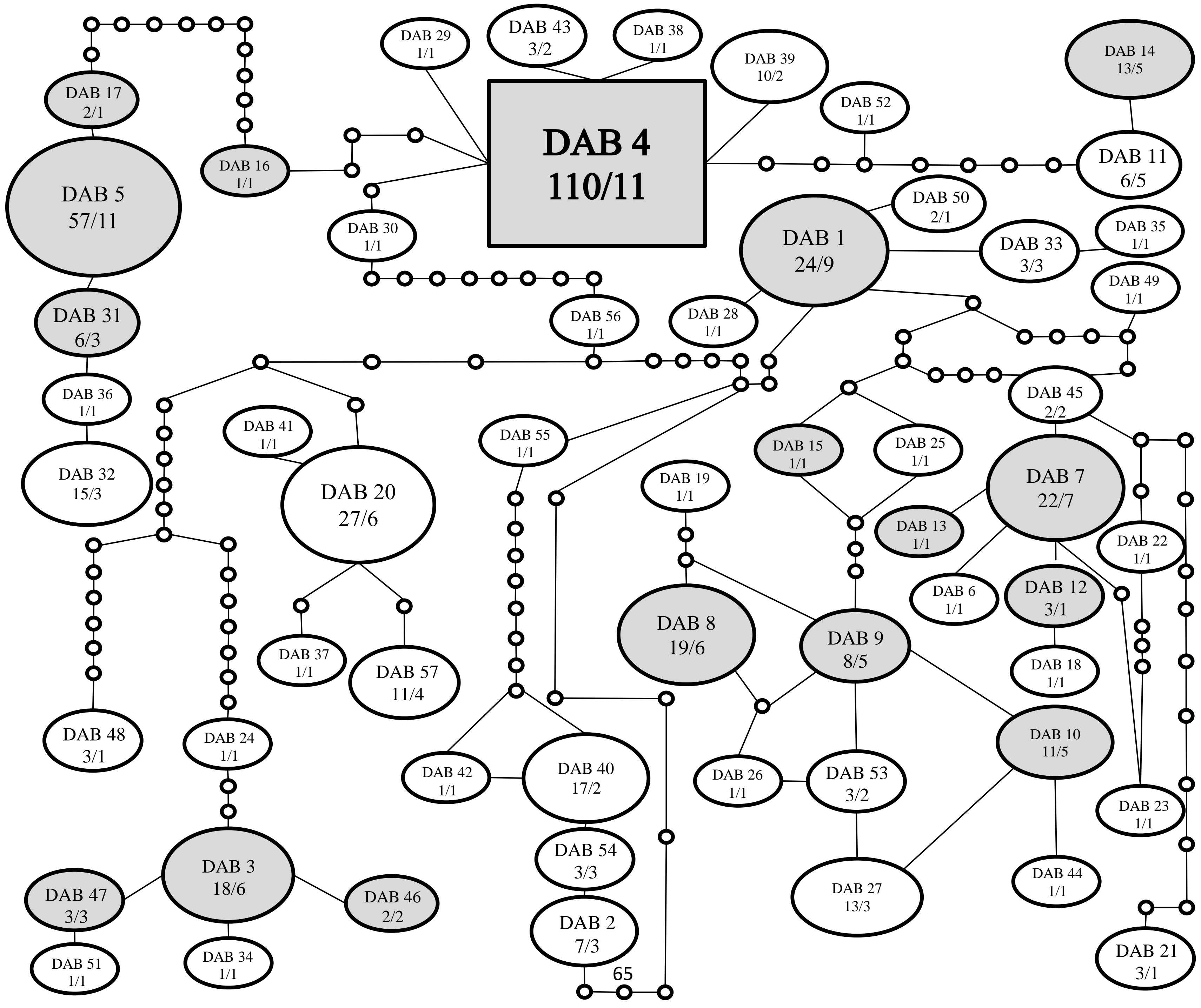
3.5.2 Crystal Lake

Crystal Lake is a native population in the Bancroft region of Ontario colonized by fish of Mississippian origin (Halbisen and Wilson, 2009). Tied for the highest number of alleles in a single population, a total of 17 MH class II β1 alleles were detected in the population with five of them being unique to Crystal Lake. Sana-DAB 4 was the most frequent allele within Crystal Lake, with a frequency of 0.306. No other allele had a frequency above 0.1 (Table 3.3). The nucleotide diversity (0.0604), non-synonymous substitutions (0.0758) and the D_N/D_S ratio (1.79) for Crystal Lake were all below the study averages (Table 3.2) while synonymous substitutions (0.0424), haplotype diversity (0.887) and allelic richness (2.45) were above study averages. The alleles of Crystal Lake separated into the two main clusters (Figure 3.12). Disagreements in the bootstrap support for the different clusters of alleles within Crystal Lake between the neighbour-joining and parsimony methods can be seen by the missing values. The parsimony network of Crystal Lake had alleles from ~ 7 different groupings, with some alleles separated by as many as 28 or 32 mutational steps (Figure 3.13). Unique alleles to Crystal Lake were located to be between or beside common, high frequency alleles in this network, suggesting recent divergence or regular "pruning" because of selection or a limited effective population size.



Sana-DAB 1
Sana-DAB 2
99/99
Sana-DAB 40
59/59
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
Sana-DAB 11
99/99
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
68/-
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
Sana-DAB 47
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.12. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Crystal Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).



DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 11
6/5
DAB 17
2/1
DAB 16
1/1
DAB 5
57/11
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 32
15/3
DAB 41
1/1
DAB 20
27/6
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 7
22/7
DAB 19
1/1
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
65
DAB 21
3/1

Figure 3.13. Network of MH class II β1 nucleotide sequences of lake trout from Crystal Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.5.3 Dickey Lake

Dickey Lake is a native population located south of Bancroft, Ontario colonized by Mississippian ancestry (Halbisen and Wilson, 2009). Similar to Crystal and Manitou, it contains a total of 17 MH class II β1 alleles with five of them being unique to the population. Three common alleles (Sana-DAB 4, 5 and 20) make up 50% of the alleles in the population with Sana-DAB 4 having the highest frequency at 0.205 (Table 3.3). The nucleotide diversity, non-synonymous and synonymous substitutions were below the study averages but the allelic richness, haplotype diversity and D_N/D_S ratio were greater (Table 3.2). Dickey Lake contained alleles from all major groups and most subgroups within the phylogenetic tree (Figure 3.14). The most common alleles (Sana-DAB 4 and 5) all are located within the same main cluster containing the majority of Dickey alleles. The parsimony network of Dickey Lake (Figure 3.15) contained alleles from ~ 7 groupings. The fact that the greatest separation in this lake between any alleles was only 17 mutational steps contributes to the low nucleotide diversity. Unique alleles such as Sana-DAB 55 and 56 were located between mutational steps from common alleles rather than in novel mutational directions like Sana-DAB 28 and 29. This unusual direction of evolution is not standard and suggests gene dispersal into the population.

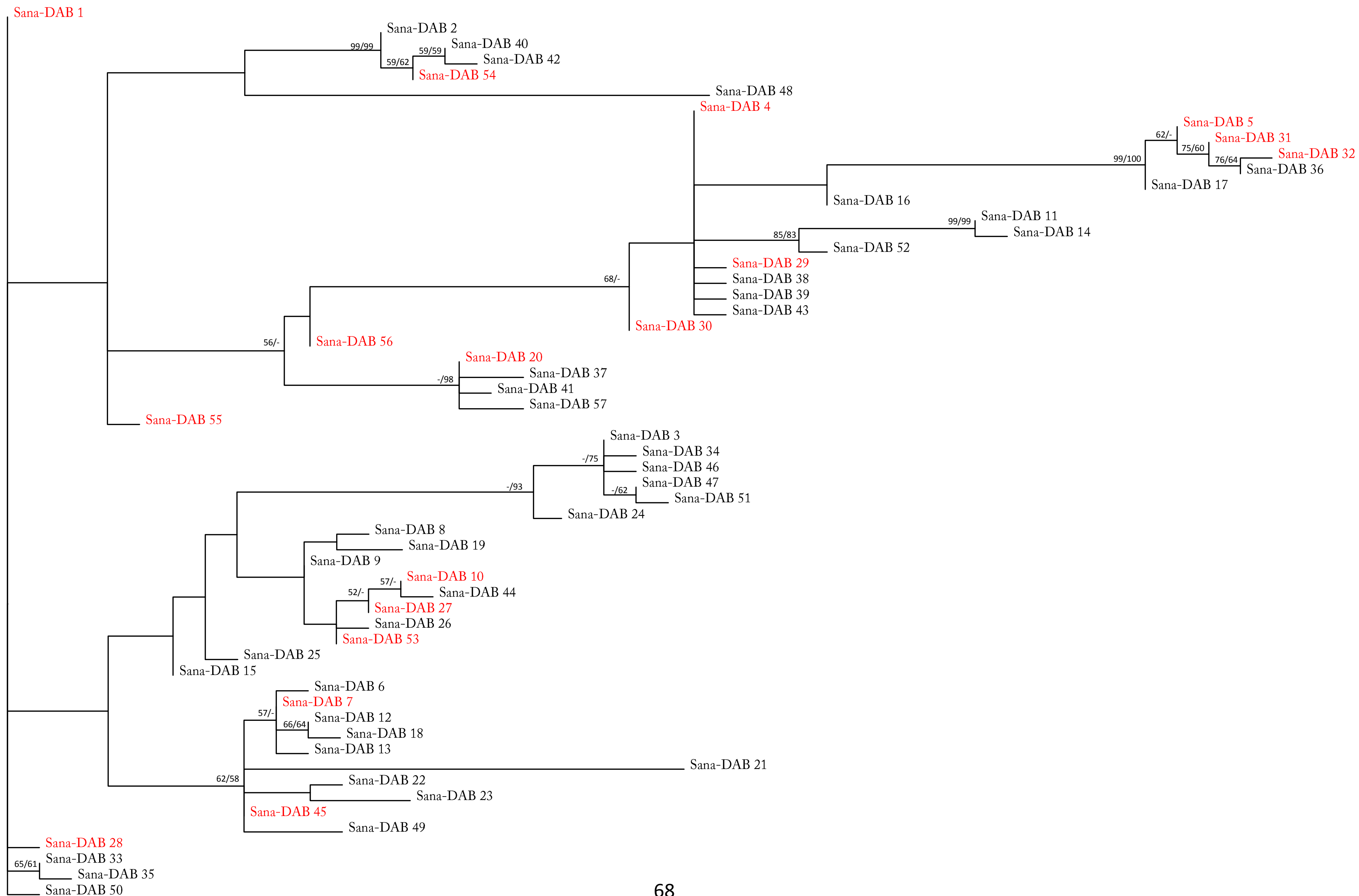

Sana-DAB 1
Sana-DAB 2
99/99
Sana-DAB 40
59/59
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
76/64
99/100
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
Sana-DAB 11
99/99
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
Sana-DAB 56
56/-
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
Sana-DAB 34
-/75
Sana-DAB 46
Sana-DAB 47
-/93
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

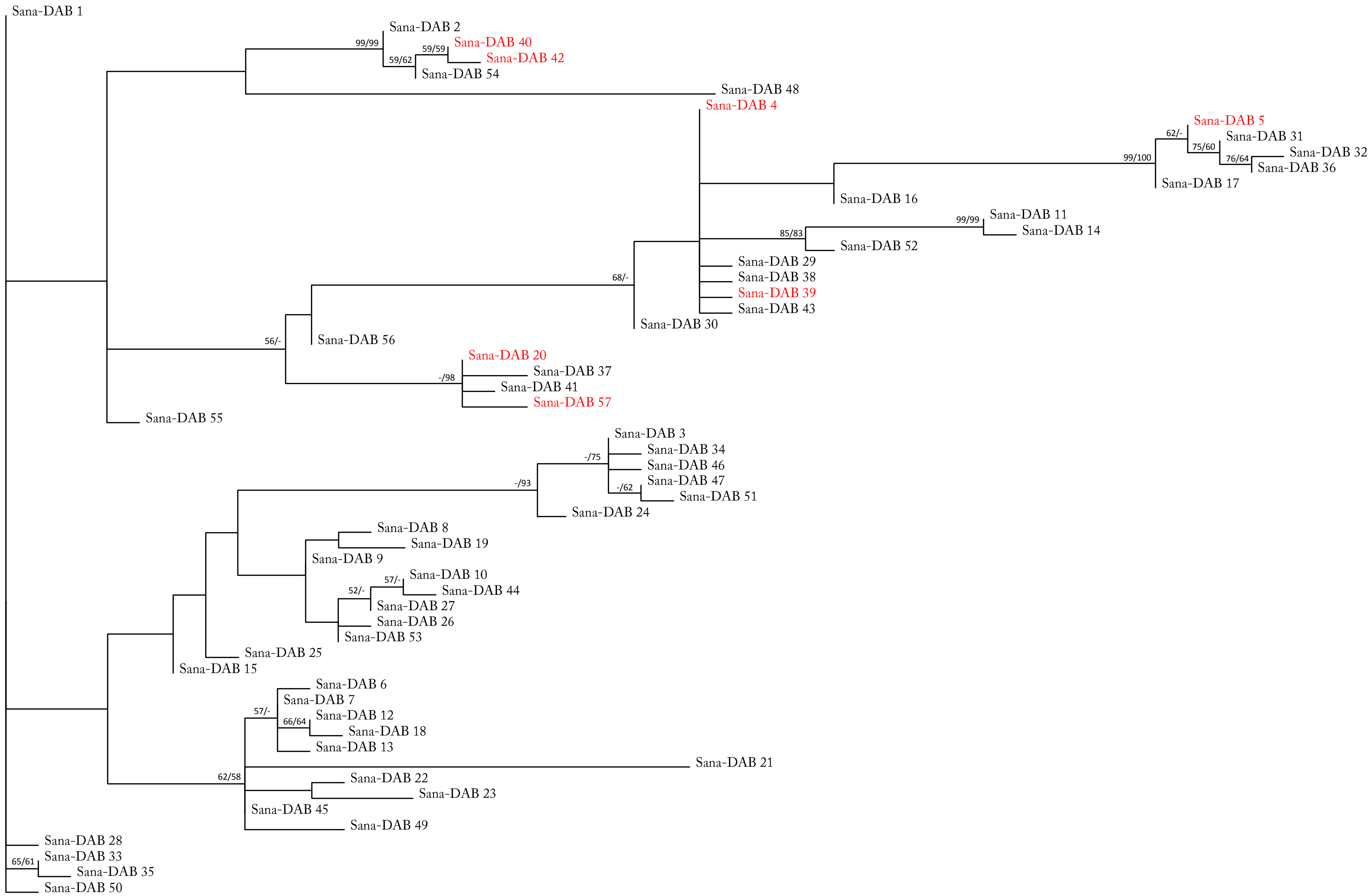
Figure 3.14. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Dickey Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

DAB 29 1/1
DAB 43 3/2
DAB 38 1/1
DAB 39 10/2
DAB 14 13/5
DAB 17 2/1
DAB 16 1/1
DAB 4 110/11
DAB 52 1/1
DAB 11 6/5
DAB 5 57/11
DAB 30 1/1
DAB 50 2/1
DAB 1 24/9
DAB 33 3/3
DAB 35 1/1
DAB 49 1/1
DAB 31 6/3
DAB 56 1/1
DAB 28 1/1
DAB 45 2/2
DAB 36 1/1
DAB 41 1/1
DAB 55 1/1
DAB 15 1/1
DAB 25 1/1
DAB 7 22/7
DAB 32 15/3
DAB 20 27/6
DAB 19 1/1
DAB 13 1/1
DAB 22 1/1
DAB 6 1/1
DAB 12 3/1
DAB 8 19/6
DAB 9 8/5
DAB 37 1/1
DAB 57 11/4
DAB 18 1/1
DAB 48 3/1
DAB 24 1/1
DAB 10 11/5
DAB 42 1/1
DAB 40 17/2
DAB 26 1/1
DAB 53 3/2
DAB 23 1/1
DAB 54 3/3
DAB 3 18/6
DAB 47 3/3
DAB 46 2/2
DAB 27 13/3
DAB 44 1/1
DAB 2 7/3
DAB 21 3/1
DAB 51 1/1
DAB 34 1/1

Figure 3.15. Network of MH class II β1 nucleotide sequences of lake trout from Dickey Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

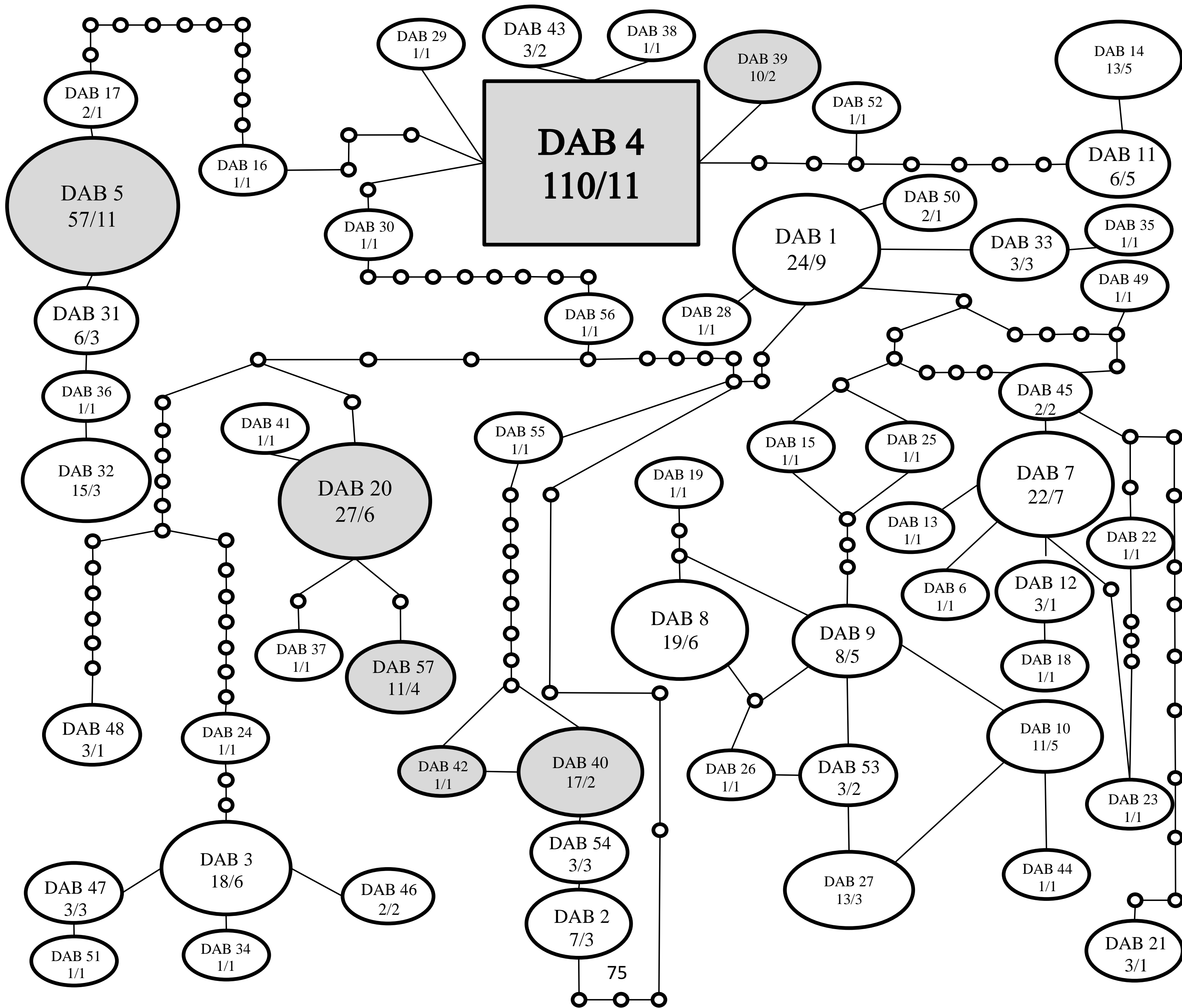
3.5.4 Macdonald Lake

Macdonald Lake is a native population located in Algonquin Park, Ontario reported to be colonized by only fish of a Mississippian refugia ancestry (Wilson and Hebert, 1996; Halbisen and Wilson, 2009). A total of 7 MH class II β1 alleles were discovered, which was the least among all populations in the study and only one of them was unique. Macdonald Lake had the lowest allelic richness in the study. Four common alleles (Sana-DAB 20, 39, 40 and 57) made up 88% of the alleles in the population with Sana-DAB 40 having the highest frequency at 0.303 (Table 3.3). Macdonald Lake was one of very few lakes that did not have Sana-DAB 4 as one of its most common high frequency alleles. The nucleotide diversity of 0.0641, non-synonymous substitutions of 0.0905 and synonymous substitutions of 0.0403 were all above the average for the study but the haplotype diversity and D_N/D_S ratio were below (Table 3.2). Macdonald Lake alleles were only detected in one of the two main clusters (Figure 3.16); no alleles within the second cluster were discovered in this population. This occurred in only two populations in this study, Macdonald and Seneca. The most common alleles within Macdonald being different from other populations in the study were detected in subgroups separate from the study's highest frequency alleles with high bootstrap confidence of 98% and 99%. The parsimony network of Macdonald Lake (Figure 3.17) contained alleles from the least number of groupings with only 4 having separation of 17 – 20 mutational steps between them. The main common alleles are separated by as few as 2 to as many as 28 mutational steps.



Sana-DAB 1
99/99
Sana-DAB 2
59/59
Sana-DAB 40
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
62/-
Sana-DAB 5
Sana-DAB 31
75/60
Sana-DAB 32
76/64
Sana-DAB 36
99/100
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
Sana-DAB 47
-/62
Sana-DAB 51
-/93
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
57/-
Sana-DAB 10
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.16. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Macdonald Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).



DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 14
13/5
DAB 52
1/1
DAB 11
6/5
DAB 17
2/1
DAB 5
57/11
DAB 16
1/1
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 32
15/3
DAB 20
27/6
DAB 19
1/1
DAB 7
22/7
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 21
3/1
DAB 51
1/1
DAB 34
1/1
75

Figure 3.17. Network of MH class II β1 nucleotide sequences of lake trout from Macdonald Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.6 Hatchery Stocking Sources – Michipicoten, Slate Islands, Manitou and Seneca

The genetic compositions of the stocking sources varied substantially, with each stock having distinguishable differences from one another. All alleles within the stocking sources can be seen within Figure 3.18 demonstrating the significant number of stocked alleles within the populations in this study. A total of 33 of the 57 unique alleles were within the four stocking sources. The Michipicoten stock contains three (Sana-DAB 9, 18 and 46), Slate Islands six (Sana-DAB 25, 26, 50, 51, 52 and 53), Manitou six (Sana-DAB 2, 10, 31, 34, 35 and 43) and Seneca three (Sana-DAB 39, 40 and 41) unique alleles not seen in the other stocking strains. The Michipicoten and Slate Islands stocks (both from Lake Superior) contain very similar genetic profiles when compared to one another with the Michipicoten population having a higher nucleotide diversity and the Slate Islands population having a higher haplotype diversity containing more total and unique alleles, thus a higher allelic richness. The main common allele groups that Michipicoten and Slate Islands are missing occur in Manitou and Seneca but these two stocks are very different from each other as they contain four differences in main groups (Sana-DAB 3, 7, 8-10, and 11). Seneca Lake has the highest nucleotide diversity among the stocking sources and the study but contains the least unique alleles.

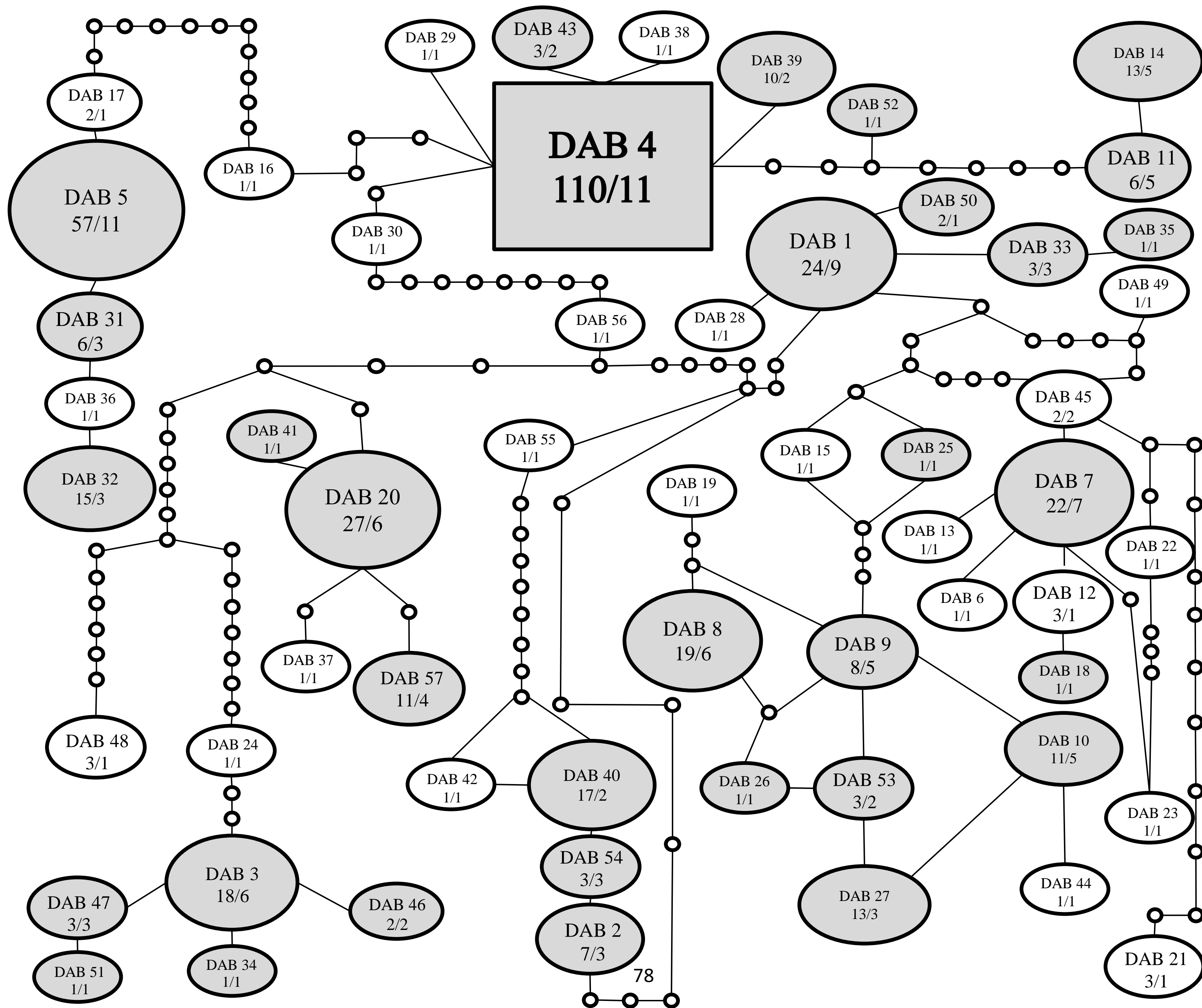

DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 11
6/5
DAB 17
2/1
DAB 16
1/1
DAB 5
57/11
DAB 30
1/1
DAB 1
24/9
DAB 50
2/1
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 32
15/3
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 7
22/7
DAB 20
27/6
DAB 19
1/1
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 27
13/3
DAB 44
1/1
DAB 47
3/3
DAB 46
2/2
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
DAB 21
3/1

Figure 3.18. Network of MH class II β1 nucleotide sequences of lake trout, coloured alleles show all stocked alleles in the lakes studied. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.6.1 Michipicoten

Michipicoten is a stocking source population from eastern Lake Superior (Ontario) that is composed of fish with a diverse ancestry (Halbisen and Wilson, 2009). It contained a total of 12 MH class II β1 alleles with only one being unique to the population. The Michipicoten stocking strain has three alleles (Sana-DAB 9, 18 and 46) which were not detected in the other stocking strains in this study. Three common alleles (Sana-DAB 3, 4 and 7) make up over 50% of allele frequencies within the population with Sana-DAB 4 having the highest frequency at 0.327 (Table 3.3). Allelic richness, haplotype and nucleotide diversity, non-synonymous substitutions and the D_N/D_S ratio were all below the study averages with only synonymous substitutions above (Table 3.2). The Michipicoten population had the lowest D_N/D_S ratio of 1.75 in the entire study which still suggested positive selection. Alleles from Michipicoten were separated into two main groups in the phylogenetic tree (Figure 3.19). The highest frequency allele, Sana-DAB 4 occurred in a separate cluster from the two other highly common alleles. Sana-DAB 3 and 7 occurred in different subgroups suggesting the most common alleles in this population were genetically different from one another or contained different ancestry. The parsimony network of Michipicoten (Figure 3.20) had alleles from ~ 7 different groupings but only one or two alleles were detected from each. There were many individual alleles separated from others within the population by 8 – 32 mutational steps, which is different from the other populations that usually show larger numbers of alleles grouping closer together with those groups separated by more mutational steps. The most common alleles (Sana-DAB 3, 4 and 7) differed in the parsimony network between 28 – 36 mutational steps.

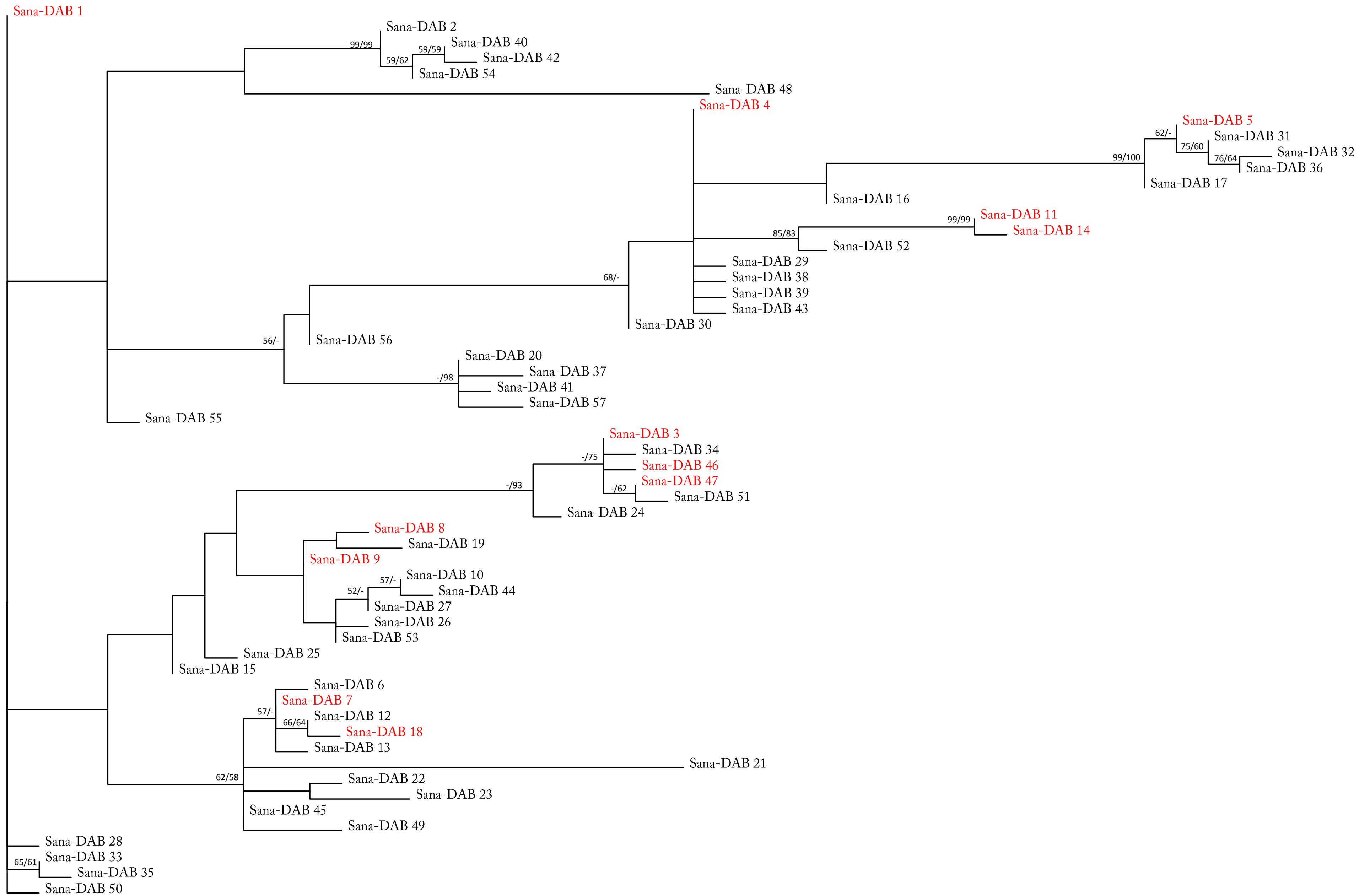
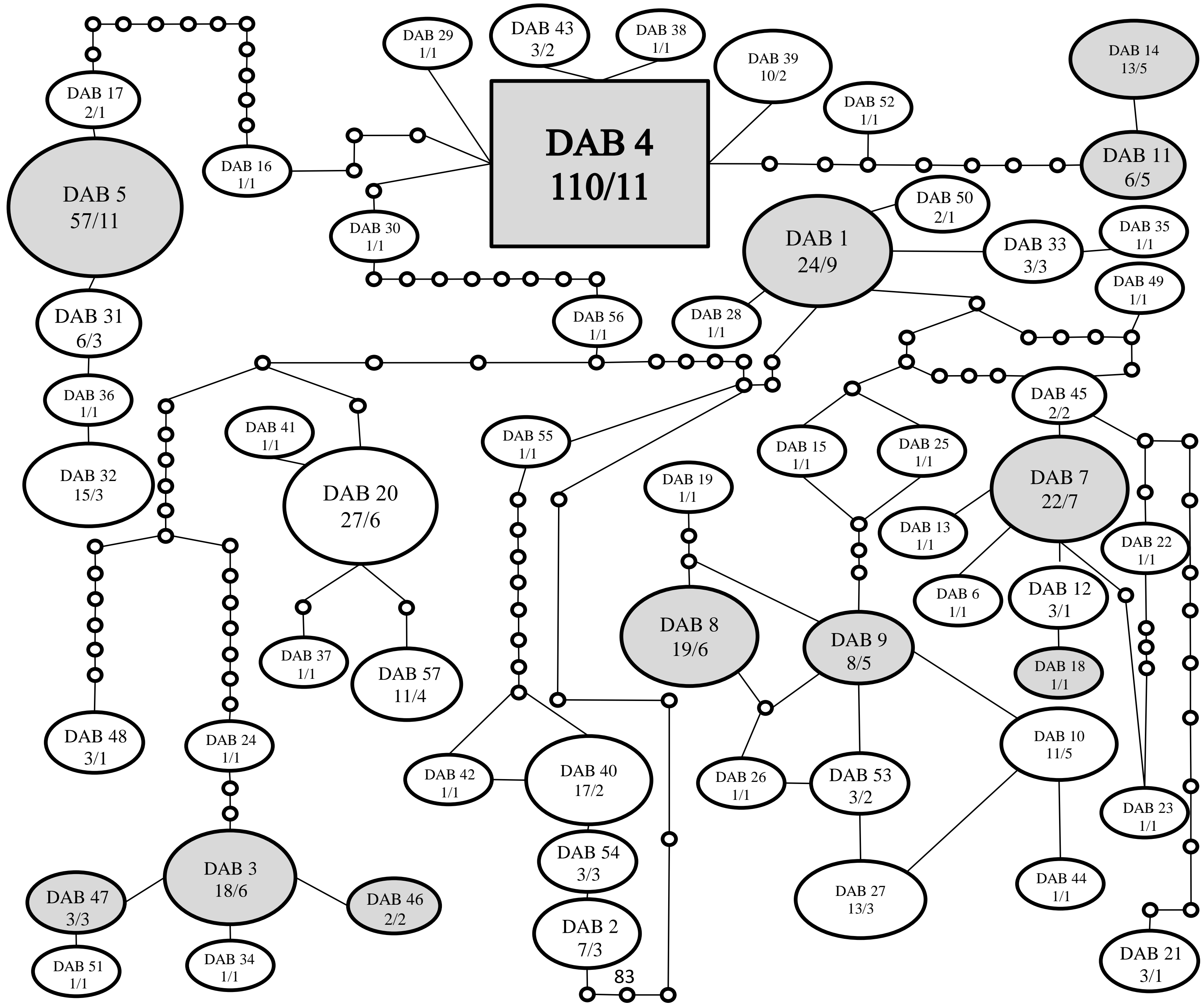

Sana-DAB 1
Sana-DAB 2
99/99
Sana-DAB 40
59/59
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
Sana-DAB 47
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
57/-
Sana-DAB 10
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.19. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Michipicoten. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).



DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 14
13/5
DAB 52
1/1
DAB 11
6/5
DAB 17
2/1
DAB 5
57/11
DAB 16
1/1
DAB 30
1/1
DAB 1
24/9
DAB 50
2/1
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 32
15/3
DAB 41
1/1
DAB 20
27/6
DAB 55
1/1
DAB 45
2/2
DAB 15
1/1
DAB 25
1/1
DAB 7
22/7
DAB 19
1/1
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 18
1/1
DAB 37
1/1
DAB 57
11/4
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
DAB 21
3/1
83

Figure 3.20. Network of MH class II β1 nucleotide sequences of lake trout from the OMNR Michipicoten broodstock. Only coloured alleles occur within the population. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.6.2 Slate Islands

Slate Islands is a stocking source population that originated from the Great Lakes region in northern Lake Superior and contains a diverse ancestry (Halbisen and Wilson, 2009). There were 15 total MH class II β1 alleles within this population with five being unique. Just two common alleles, Sana-DAB 4 and 27, made up 50% of allele sequences in the population with Sana-DAB 4 having the highest allelic frequency in the population, and indeed of all populations in this study, at 0.375 (Table 3.3). Slate Islands had the highest allelic richness of the study (2.53) but the haplotype and nucleotide diversity, non-synonymous and synonymous substitutions were all below the study averages while the D_N/D_S ratio was exactly average (Table 3.2). Slate Island alleles were detected in both main groups in the phylogenetic tree (Figure 3.21) with the majority occurring within the cluster separate from the highest frequency allele (Sana-DAB 4) in the study. This grouping contained the other highly common allele within this population, Sana-DAB 27. The parsimony network of Slate Islands (Figure 3.22) contained alleles from ~ 7 groupings with some separated by as many as 28 or 32 mutational steps. This stocking population was missing two commonly stocked allele groups around Sana-DAB 20 and 40 which were commonly recognized in Manitou and Seneca Lake stocks. The two most common alleles from the Slate Islands population differed from one another by 31 mutational steps within the parsimony network.

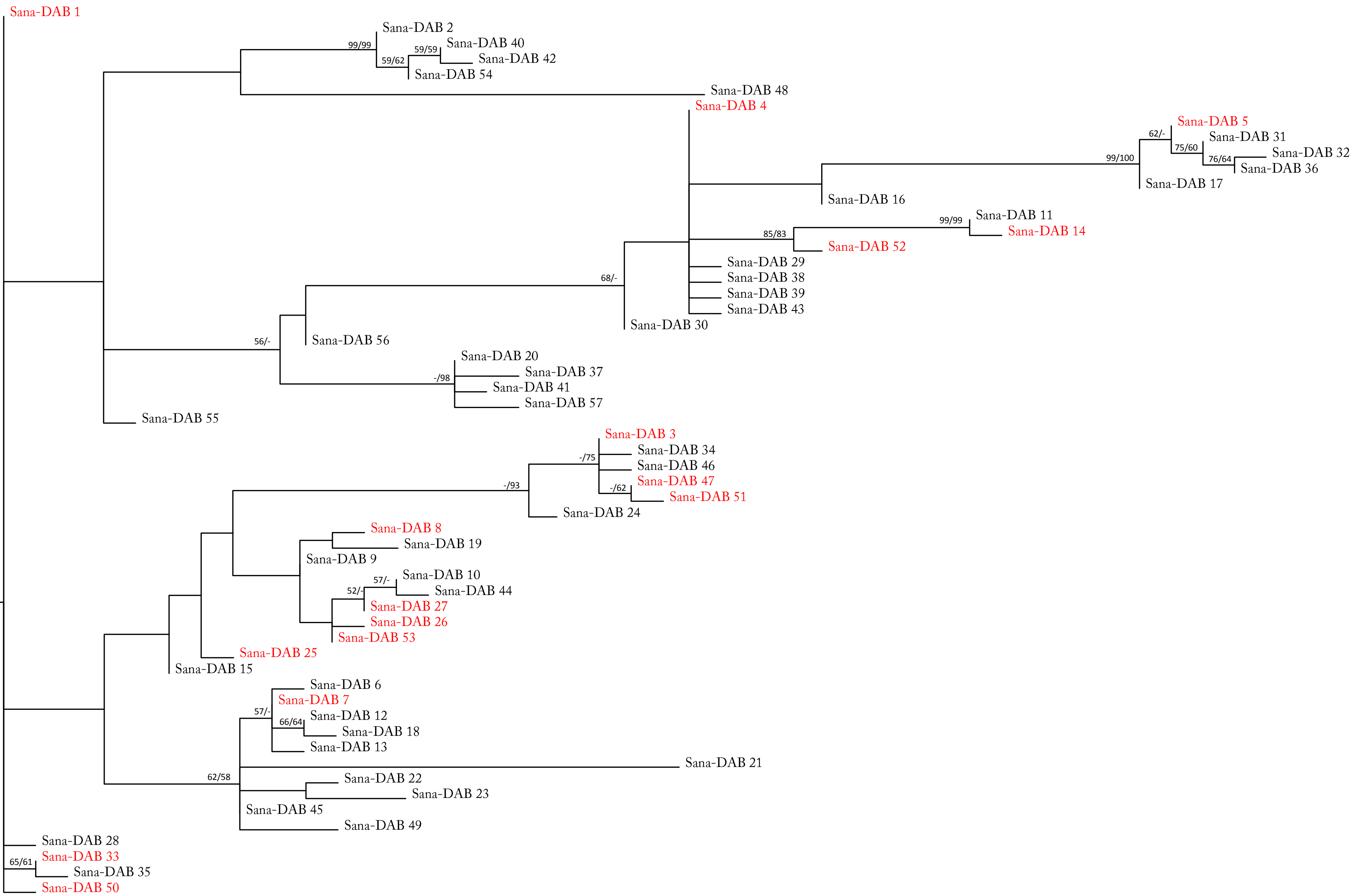

Sana-DAB 1
Sana-DAB 2
99/99
Sana-DAB 40
59/59
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
68/-
Sana-DAB 39
Sana-DAB 43
Sana-DAB 30
Sana-DAB 56
56/-
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
Sana-DAB 47
-/62
Sana-DAB 51
-/93
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.21. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Slate Islands. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

DAB 4
110/11

DAB 29
1/1

DAB 43
3/2

DAB 38
1/1

DAB 39
10/2

DAB 14
13/5

DAB 52
1/1

DAB 11
6/5

DAB 17
2/1

DAB 16
1/1

DAB 5
57/11

DAB 30
1/1

DAB 50
2/1

DAB 1
24/9

DAB 33
3/3

DAB 35
1/1

DAB 49
1/1

DAB 31
6/3

DAB 56
1/1

DAB 28
1/1

DAB 36
1/1

DAB 41
1/1

DAB 55
1/1

DAB 45
2/2

DAB 15
1/1

DAB 25
1/1

DAB 32
15/3

DAB 20
27/6

DAB 19
1/1

DAB 7
22/7

DAB 13
1/1

DAB 22
1/1

DAB 6
1/1

DAB 12
3/1

DAB 8
19/6

DAB 9
8/5

DAB 37
1/1

DAB 57
11/4

DAB 18
1/1

DAB 48
3/1

DAB 24
1/1

DAB 10
11/5

DAB 42
1/1

DAB 40
17/2

DAB 26
1/1

DAB 53
3/2

DAB 23
1/1

DAB 54
3/3

DAB 3
18/6

DAB 47
3/3

DAB 46
2/2

DAB 27
13/3

DAB 44
1/1

DAB 2
7/3

88

DAB 51
1/1

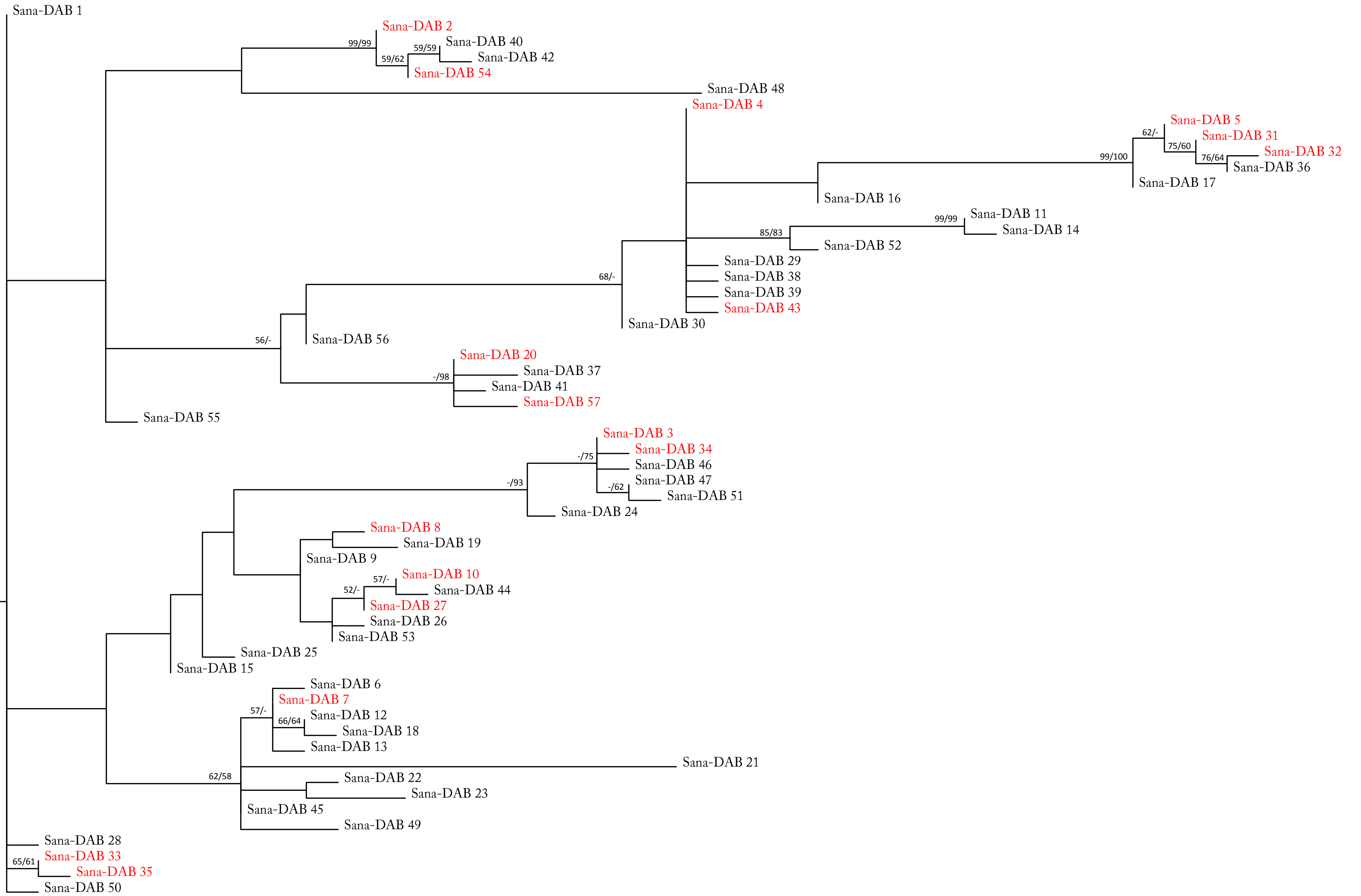
DAB 34
1/1

DAB 21
3/1

Figure 3.22. Network of MH class II β1 nucleotide sequences of lake trout from the OMNR Slate Islands broodstock. Only coloured alleles occur within the population. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

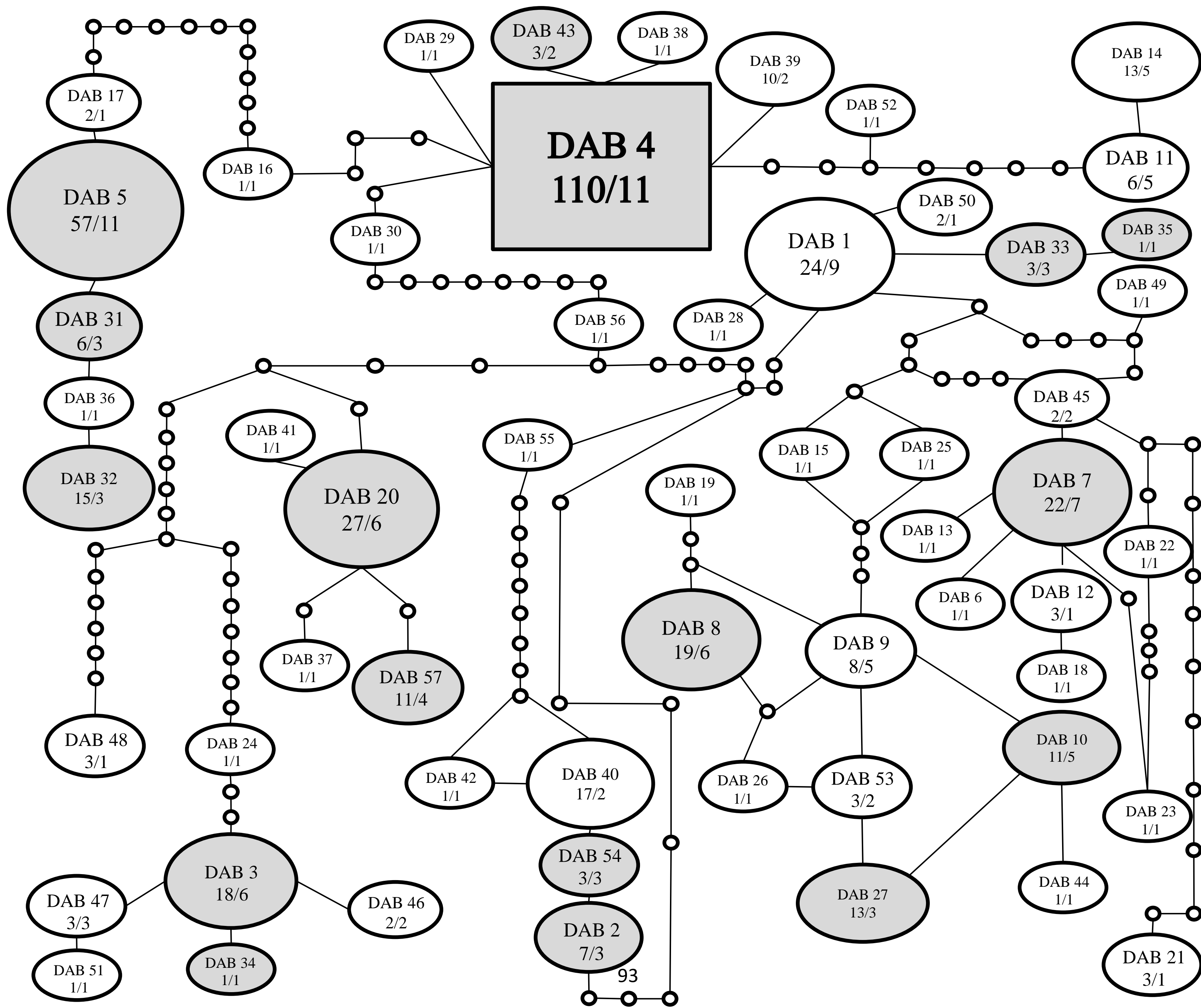
3.6.3 Manitou Lake

Manitou Lake is a stocking source population that originated from the northern Great Lakes, specifically Manitoulin Island, Ontario and contains a diverse ancestry (Halbisen and Wilson, 2009). The lake had a native lake trout population that was supplementally stocked with lake trout from two rivers in eastern Lake Superior (OMNR, 2005). In common with the native lakes of Crystal and Dickey, Manitou Lake population had the highest number of MH class II β1 alleles with 17 but only had two that were unique. Sana-DAB 4 and 32 at frequencies of 0.2 and 0.178 were the only two alleles to have frequencies above 0.1 within the Manitou Lake population (Table 3.3). Manitou had the highest haplotype diversity (0.912) in the study as well as had values above the study averages for allelic richness, nucleotide diversity, non-synonymous and synonymous substitutions and D_N/D_S ratio (Table 3.2). The phylogenetic tree of Manitou Lake contained a widespread allelic representation of all groups and almost all subgroups (Figure 3.23). The most common alleles Sana-DAB 4 and 32 were located within the same major grouping but within different subgroups with high certainty. The parsimony network of Manitou Lake (Figure 3.24) showed alleles within 8 of the major allele groups, except the grouping of Sana-DAB 11 and 14 which were identified in the stocking sources of Seneca and Michipicoten. There were a total of 20 mutational steps between the two common alleles of Sana-DAB 4 and 32.



Sana-DAB 1
Sana-DAB 2
99/99
59/59
Sana-DAB 40
59/62
Sana-DAB 42
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
85/83
Sana-DAB 14
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
68/-
Sana-DAB 39
Sana-DAB 43
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
-/98
Sana-DAB 37
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
Sana-DAB 47
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
57/-
Sana-DAB 10
52/-
Sana-DAB 44
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.23. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Manitou Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).



DAB 29 1/1
DAB 43 3/2
DAB 38 1/1
DAB 39 10/2
DAB 14 13/5
DAB 17 2/1
DAB 16 1/1
DAB 4 110/11
DAB 52 1/1
DAB 11 6/5
DAB 5 57/11
DAB 30 1/1
DAB 50 2/1
DAB 1 24/9
DAB 33 3/3
DAB 35 1/1
DAB 49 1/1
DAB 31 6/3
DAB 56 1/1
DAB 28 1/1
DAB 36 1/1
DAB 45 2/2
DAB 41 1/1
DAB 55 1/1
DAB 15 1/1
DAB 25 1/1
DAB 32 15/3
DAB 20 27/6
DAB 19 1/1
DAB 7 22/7
DAB 13 1/1
DAB 22 1/1
DAB 6 1/1
DAB 12 3/1
DAB 8 19/6
DAB 9 8/5
DAB 37 1/1
DAB 57 11/4
DAB 18 1/1
DAB 48 3/1
DAB 24 1/1
DAB 10 11/5
DAB 42 1/1
DAB 40 17/2
DAB 26 1/1
DAB 53 3/2
DAB 23 1/1
DAB 3 18/6
DAB 54 3/3
DAB 47 3/3
DAB 46 2/2
DAB 27 13/3
DAB 44 1/1
DAB 2 7/3
DAB 51 1/1
DAB 34 1/1
DAB 21 3/1

Figure 3.24. Network of MH class II β1 nucleotide sequences of lake trout from Manitou Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.6.4 Seneca Lake

Seneca Lake is a stocking source population in the Finger Lakes region of New York State, USA thought to be colonized by both Mississippian and Atlantic origins (Wilson and Herbert 1996, 1998). It contained a total of 12 MH class II β1 alleles within the population with only one being unique. Four common alleles; Sana-DAB 4, 14, 32 and 40, make up almost 75% of the alleles in the population with Sana-DAB 14 and 40 having the highest frequencies at 0.206 (Table 3.3). Seneca Lake had the highest nucleotide diversity (0.0719) within the study while having haplotype diversity, non-synonymous substitutions and a D_N/D_S ratio above the average. Allelic richness and synonymous substitutions for the Seneca population were both below the study averages (Table 3.2). As seen in Macdonald Lake, Seneca Lake alleles were only detected in one of the two main clusters in the phylogenetic tree (Figure 3.25); no alleles within the second cluster were discovered in this population. The parsimony network of Seneca Lake (Figure 3.26) contained alleles from ~ 6 groupings with 8 – 19 mutational steps between them. This stocking strain looks very different from other strains (Manitou, Michipicoten and Slate Islands) as it is missing common alleles from the groupings Sana-DAB 3, 7, and 8/9/10/27. The main common alleles in Seneca Lake have as many as 27 mutational steps and as little as 9 between them.

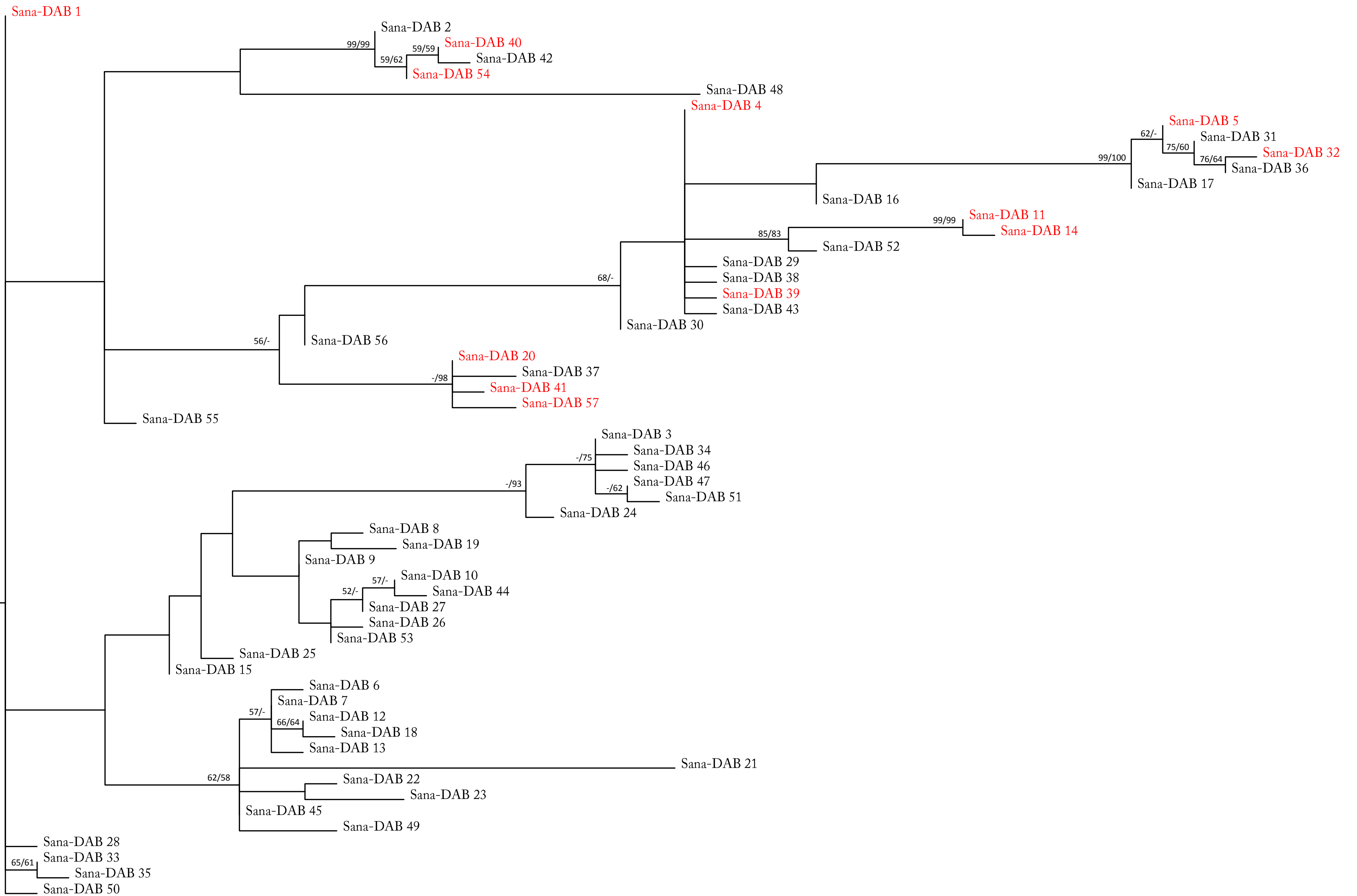

Sana-DAB 1
Sana-DAB 2
99/99
59/59
Sana-DAB 40
59/62
Sana-DAB 42
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
68/-
Sana-DAB 39
Sana-DAB 43
Sana-DAB 30
Sana-DAB 56
56/-
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
Sana-DAB 47
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
52/-
Sana-DAB 44
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
66/64
Sana-DAB 12
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
65/61
Sana-DAB 33
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.25. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Seneca Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

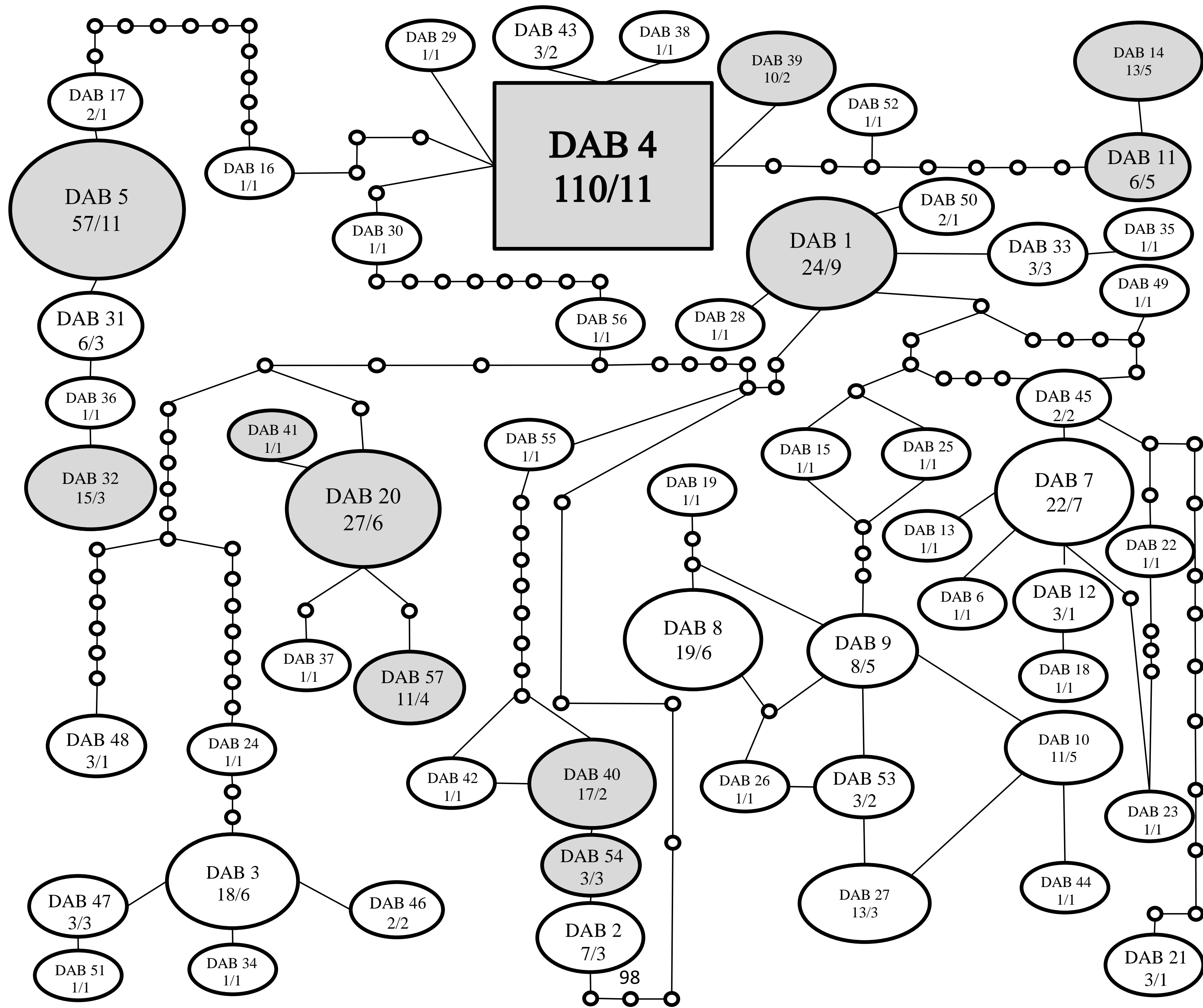

DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 14
13/5
DAB 52
1/1
DAB 11
6/5
DAB 17
2/1
DAB 5
57/11
DAB 16
1/1
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 32
15/3
DAB 41
1/1
DAB 20
27/6
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 7
22/7
DAB 19
1/1
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
DAB 21
3/1
98

Figure 3.26. Network of MH class II β1 nucleotide sequences of lake trout from Seneca Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.7 Native but Stocked – Simcoe, Opeongo and Kingscote

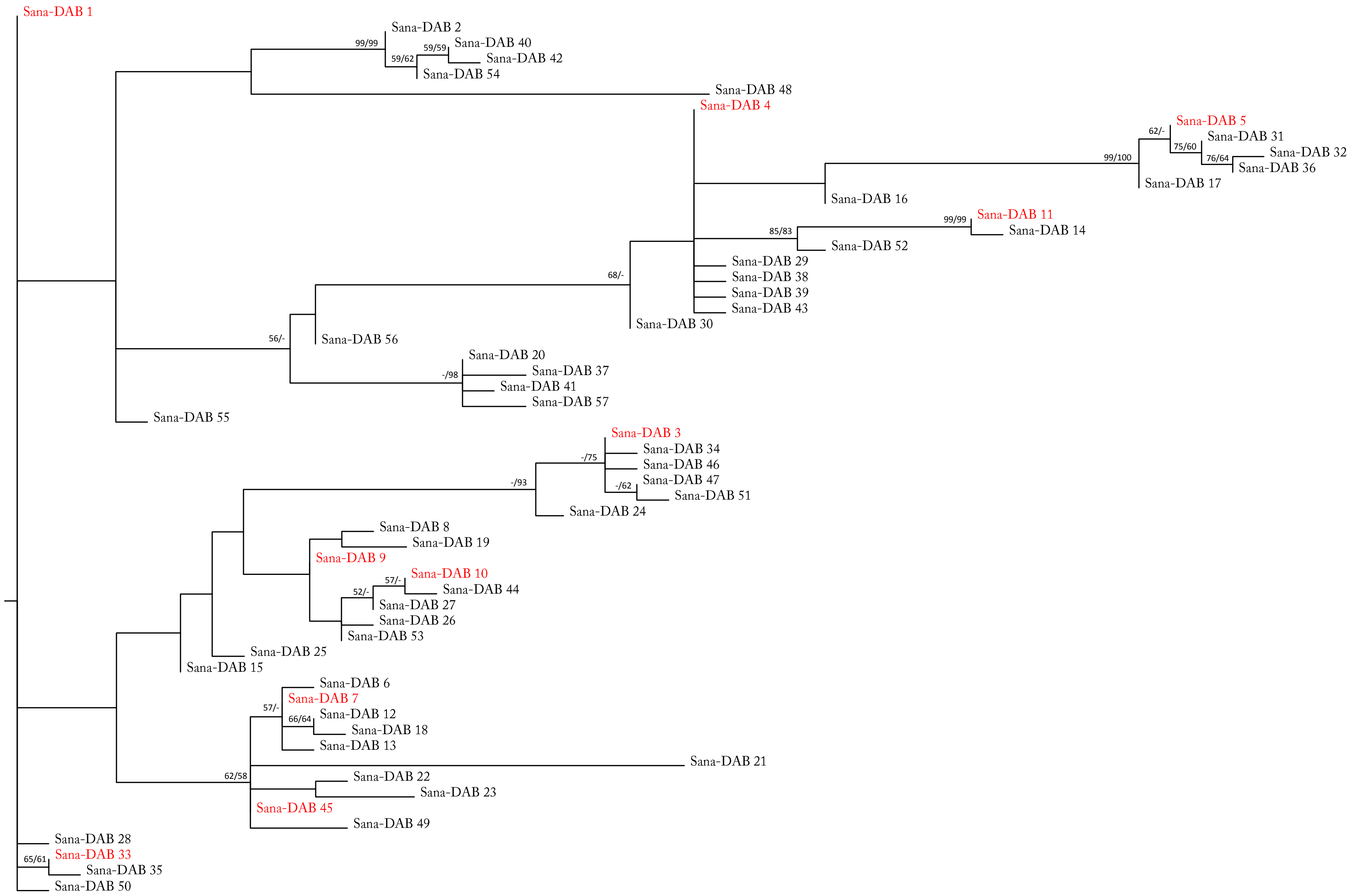
Stocking within native lakes can either help rejuvenate low diversity, diminish the diversity or remove native alleles within populations or have no effect. The potential effects of the hatchery stocking sources from this study have had on the genetic profiles of Simcoe, Opeongo and Kingscote Lakes can be seen with the number of alleles seen in the stocking sources (Table 3.5). The historical stocking of Simcoe Lake many times may have contributed to the removal of any native allele profile the population once had since no unique alleles and a decreased total amount of alleles were discovered (Table 3.1). In addition, this lake exhibited an lower haplotype and nucleotide diversity than the study average (Table 3.2). Stocking within Opeongo Lake seems to have enhanced the gene pool of the population by increasing its haplotype and nucleotide diversity along with unique alleles in the population (which is the highest in the study among all populations) and probably increasing the total number of alleles within the population. Stocking within Kingscote Lake does not seem to be helping as this population has the lowest haplotype and nucleotide diversities of all of the populations in the study. Even with the well-documented stocking attempts into Kingscote, it does not look like many main stocking alleles have been able to gain a foothold within the Kingscote population's genetic profile.

Table 3.5. Alleles from hatchery stocking sources in common with stocked populations in this study. Circled alleles represent alleles unique to only that stocking source of the four studied but were seen naturally in native unstocked populations in this study.

	Alleles Present in Population	**Michipicoten Lake**	**Slate Islands Lake**	**Manitou Lake**	**Seneca Lake**	**Alleles not in stocking sources**
Simcoe Lake	1, 3, 4, 5, 7, 9, 10, 11, 33, 45	1, 3, 4, 5, 7, 9, 11	1, 3, 4, 5, 7, 33	3, 4, 5, 7, 10, 33	1, 4, 5, 11	45
Opeongo Lake	1, 2, 4, 5, 9, 11, 14, 19, 20, 21, 22, 23, 24, 48, 49	1, 4, 5, 9, 11, 14	1, 4, 5, 14	2, 4, 5, 20	1, 4, 5, 11, 14, 20	19, 21, 22, 23, 24, 48, 49
Kingscote Lake	1, 4, 5, 8, 20, 36, 37, 38, 57	1, 4, 5, 8	1, 4, 5, 8	4, 5, 8, 20, 57	1, 4, 5, 20, 57	36, 37, 38

3.7.1 Lake Simcoe

Lake Simcoe is a stocked population in southern Ontario colonized with fish from diverse ancestries (Grewe and Hebert 1988; Grewe et al. 1993). A total of 10 MH class II β1 alleles were identified in this population with none of them being unique. Four common alleles (Sana-DAB 4, 5, 7 and 10) make up over 75% of the allele frequencies within the population with Sana-DAB 4 having the highest frequency at 0.306 (Table 3.3). The allelic richness, haplotype and nucleotide diversities, non-synonymous substitutions and the D_N/D_S ratio were all below the study averages while the synonymous substitutions were above (Table 3.2). The alleles of Lake Simcoe were located in both main clusters within the phylogenetic tree (Figure 3.27). The most common alleles making up the majority of frequencies in the population were separated with two in each main grouping. Sana-DAB 4 and 5 were located in the upper grouping while Sana-DAB 7 and 10 in the lower (Figure 3.27) where more Simcoe alleles were grouped. The parsimony network of Lake Simcoe (Figure 3.28) had alleles from ~ 7 different groupings, most separated by 7 – 9 mutational steps and two groups by 17 and 28 steps. The four most common alleles differed from each other between 17 – 47 mutational steps (Figure 3.28). The genetic profile of Lake Simcoe contained many common alleles in the stocking sources (Figure 3.18 and Table 3.5) but Lake Simcoe is missing the common stocking population allele groups of Sana-DAB 20 and 40 (Figure 3.27) which occur in both Manitou and Seneca stocks. Lake Simcoe contained the unique alleles Sana-DAB 9 and 10 which are only discovered within Michipicoten and Manitou stocks.



Sana-DAB 1
Sana-DAB 2
99/99
59/59
Sana-DAB 40
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
99/100
76/64
Sana-DAB 36
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
-/98
Sana-DAB 37
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
-/93
Sana-DAB 47
-/62
Sana-DAB 51
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
57/-
Sana-DAB 10
52/-
Sana-DAB 44
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
57/-
Sana-DAB 7
66/64
Sana-DAB 12
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
65/61
Sana-DAB 33
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.27. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Lake Simcoe. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

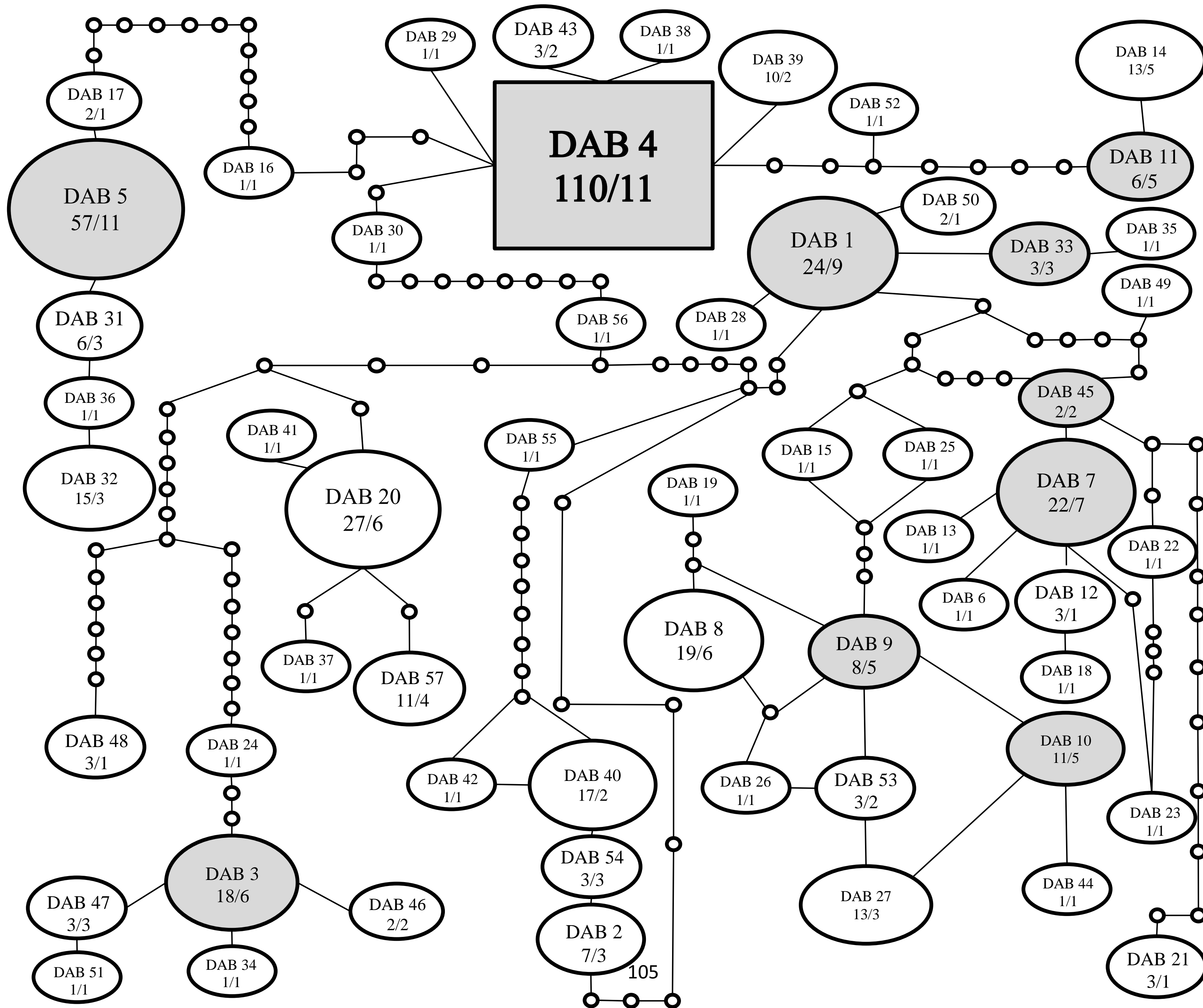

DAB 4
110/11
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 11
6/5
DAB 17
2/1
DAB 16
1/1
DAB 5
57/11
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 32
15/3
DAB 20
27/6
DAB 19
1/1
DAB 7
22/7
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 21
3/1
DAB 51
1/1
DAB 34
1/1

Figure 3.28. Network of MH class II β1 nucleotide sequences of lake trout from Lake Simcoe. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.7.2 Opeongo Lake

Opeongo Lake is a native lake located in Algonquin Park, Ontario colonized by Mississippian and Atlantic ancestry (Halbisen and Wilson, 2009) that has been historically stocked. Almost half of the 15 total MH class II β1 alleles were unique making Opeongo Lake the population with the most unique alleles (seven) in the study. Two common alleles (Sana-DAB 4 and 5) made up over 50% of alleles within the population. Sana-DAB 5 had the highest frequency at 0.327 making Opeongo Lake one of only two lakes which did not have Sana-DAB 4 as the highest allelic frequency (Table 3.3). Allelic richness, nucleotide diversity, non-synonymous and synonymous substitutions were all above study averages. The haplotype diversity and the D_N/D_S ratio were below (Table 3.2). The Opeongo Lake alleles were split between the two main clusters in the phylogenetic tree (Figure 3.29). Most alleles were individually separated from subgroups while only two pairs were grouped together. Both Sana-DAB 4 and 5, the two most common alleles in the study, occurred in the same main cluster but different subgroups with bootstrap support. The parsimony network of Opeongo (Figure 3.30) showed the presence of alleles from all 9 major groupings but contained a genetic profile of generally isolated alleles in which only one pair of alleles had a single a mutational step between them. All other alleles had a minimum of 4 to a maximum of 15 mutational steps to the next closest allele. The two most common alleles in Opeongo Lake differed from one another by 17 mutational steps. Opeongo Lake contained alleles from stocking sources, most notably the unique alleles Sana-DAB 2 and 9 which are only seen within Michipicoten and Manitou stocks (Table 3.5).

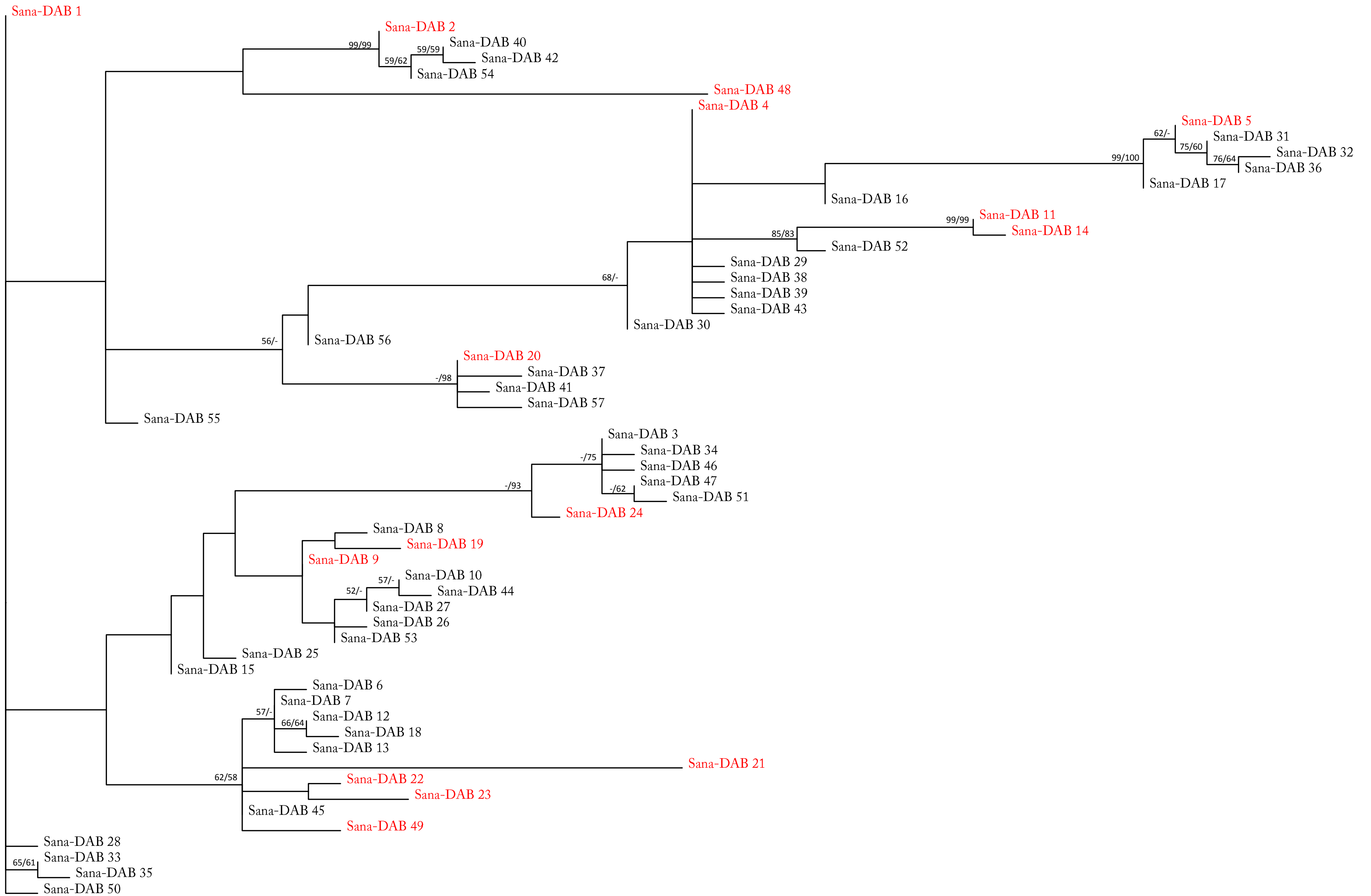
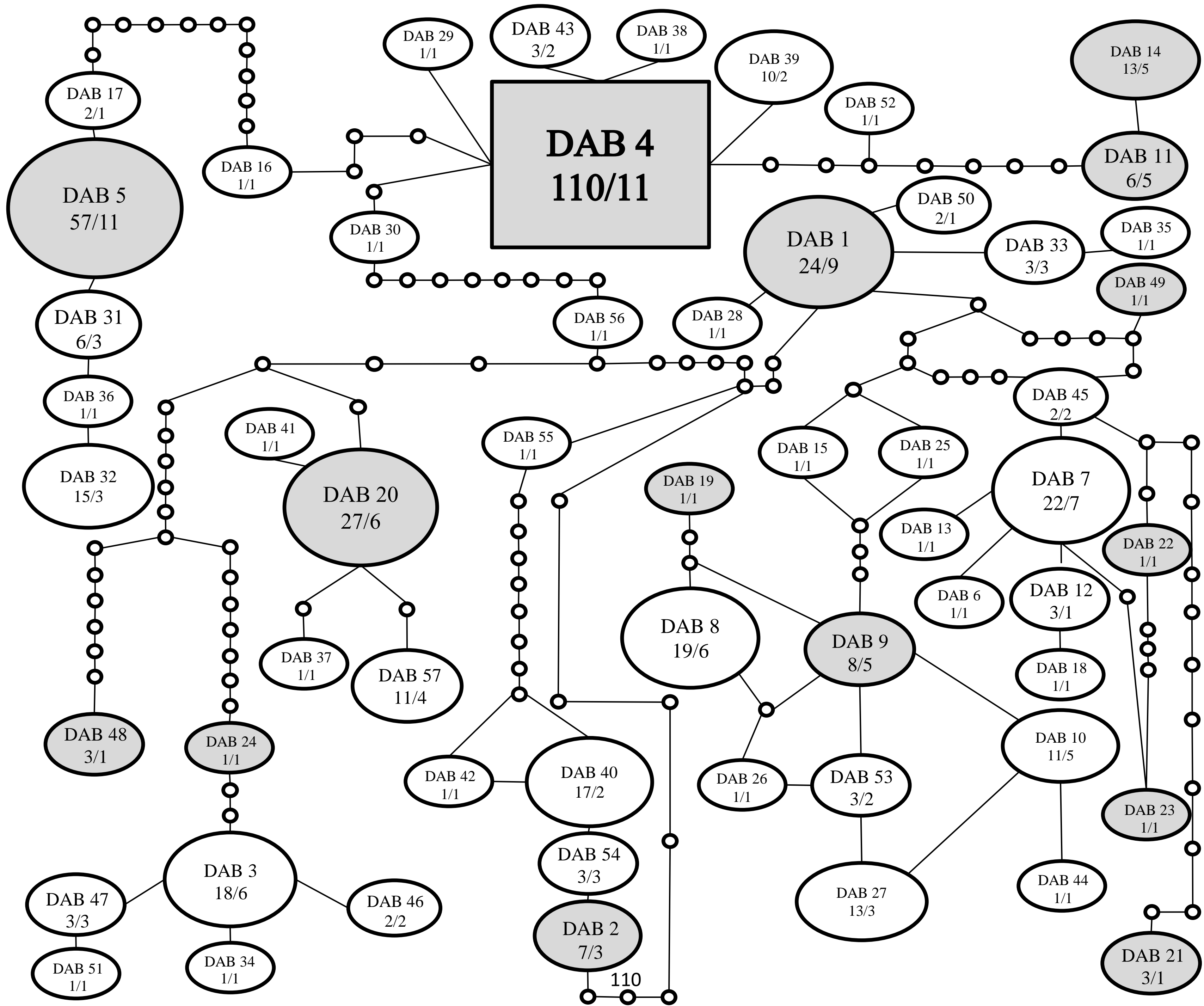

Sana-DAB 1
Sana-DAB 2
99/99
59/59
Sana-DAB 40
59/62
Sana-DAB 42
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
62/-
Sana-DAB 5
Sana-DAB 31
75/60
Sana-DAB 32
76/64
Sana-DAB 36
99/100
Sana-DAB 17
Sana-DAB 16
99/99
Sana-DAB 11
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
56/-
Sana-DAB 56
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
-/75
Sana-DAB 34
Sana-DAB 46
Sana-DAB 47
-/62
Sana-DAB 51
-/93
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
57/-
Sana-DAB 10
52/-
Sana-DAB 44
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
66/64
Sana-DAB 12
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.29. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Opeongo Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

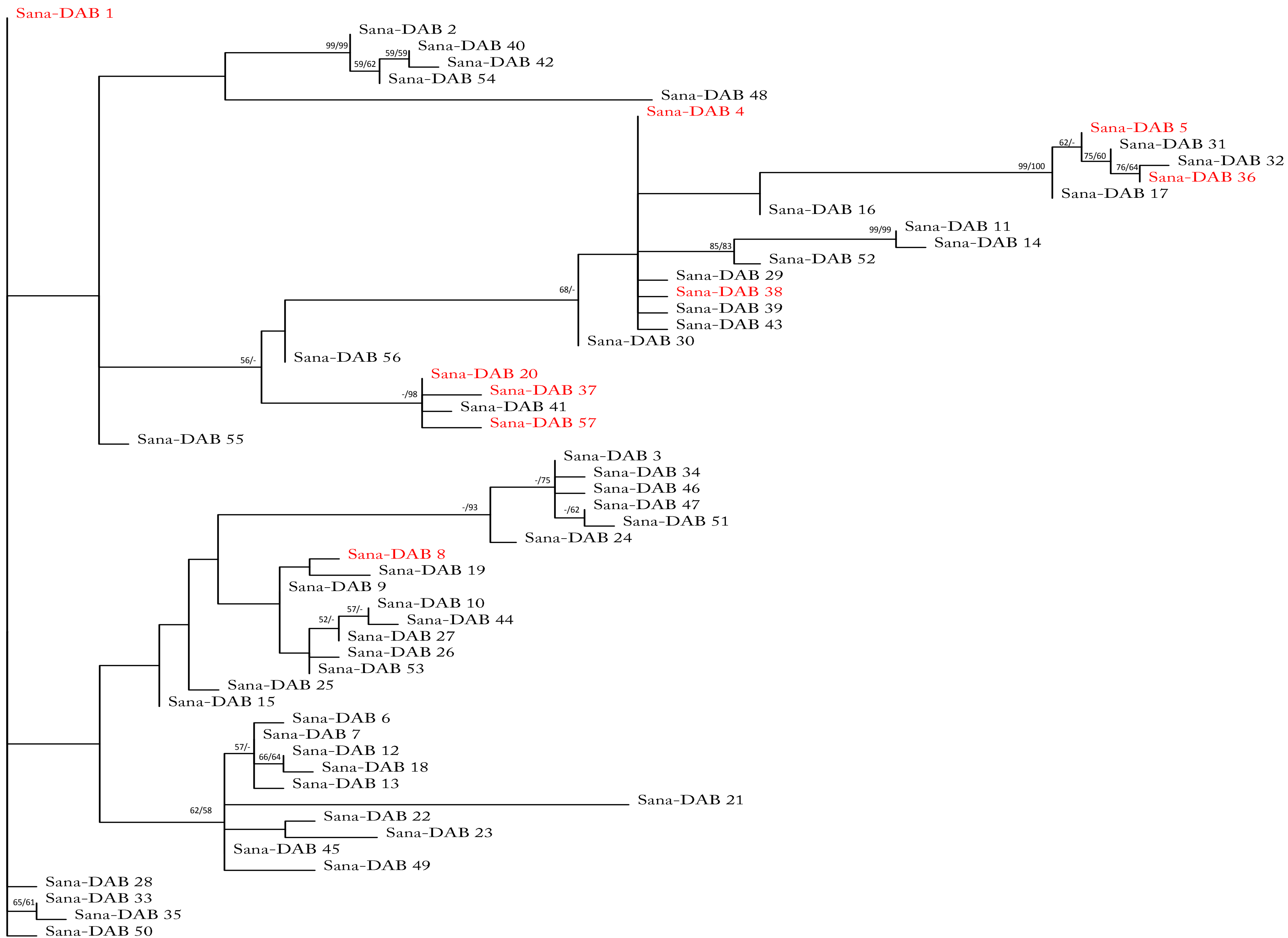


DAB 4
110/11
DAB 5
57/11
DAB 17
2/1
DAB 16
1/1
DAB 29
1/1
DAB 43
3/2
DAB 38
1/1
DAB 39
10/2
DAB 52
1/1
DAB 14
13/5
DAB 11
6/5
DAB 30
1/1
DAB 50
2/1
DAB 1
24/9
DAB 33
3/3
DAB 35
1/1
DAB 49
1/1
DAB 31
6/3
DAB 56
1/1
DAB 28
1/1
DAB 36
1/1
DAB 45
2/2
DAB 41
1/1
DAB 55
1/1
DAB 15
1/1
DAB 25
1/1
DAB 32
15/3
DAB 20
27/6
DAB 19
1/1
DAB 7
22/7
DAB 13
1/1
DAB 22
1/1
DAB 6
1/1
DAB 12
3/1
DAB 8
19/6
DAB 9
8/5
DAB 37
1/1
DAB 57
11/4
DAB 18
1/1
DAB 48
3/1
DAB 24
1/1
DAB 10
11/5
DAB 42
1/1
DAB 40
17/2
DAB 26
1/1
DAB 53
3/2
DAB 23
1/1
DAB 54
3/3
DAB 3
18/6
DAB 47
3/3
DAB 46
2/2
DAB 27
13/3
DAB 44
1/1
DAB 2
7/3
DAB 51
1/1
DAB 34
1/1
110
DAB 21
3/1

Figure 3.30. Network of MH class II β1 nucleotide sequences of lake trout from Opeongo Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.7.3 Kingscote Lake

Kingscote Lake is a native lake that has been stocked in the past located in Algonquin Park, Ontario which has been colonized by the Mississippian lineage (Halbisen and Wilson, 2009). The Kingscote Lake population contained a total of 9 MH class II β1 alleles, which was the 2nd least among populations in the study, with a third of them being unique. Three common alleles, Sana-DAB 5, 8 and 20 made up almost 82% of the alleles in the population with Sana-DAB 8 having the highest allele frequency at 0.324 (Table 3.3). This was one of two lakes to not have Sana-DAB 4 as one of highest allele frequencies within the lake. Kingscote Lake had the lowest haplotype (0.788) and nucleotide (0.0548) diversities contributing to a low amount of genetic variation within the population (Table 3.2). The allelic richness and non-synonymous substitutions were below the average along with the lowest synonymous substitutions in the study. This contributed to the highest D_N/D_S ratio of all populations at 4.32, suggesting a high amount of positive balancing selection (Table 3.2). The majority of Kingscote alleles were detected in one main cluster with only a single allele (Sana-DAB 8) albeit one of the most common in this population located in the other main grouping (Figure 3.31). The parsimony network of Kingscote Lake (Figure 3.32) contained alleles from only 5 groupings separated by as many as 17 mutational steps. For a population with a rich stocking history, Kingscote Lake was missing many common alleles (Figure 3.18) as well as four common groupings, Sana-DAB 3, 7, 11 and 40 seen in stocking populations. Most alleles in Kingscote Lake were common in at least two but more likely 3 or all 4 stocking strains in this study (Table 3.5) with no unique alleles from stocking sources being discovered within the population.



Sana-DAB 1
Sana-DAB 2
99/99
Sana-DAB 40
59/59
Sana-DAB 42
59/62
Sana-DAB 54
Sana-DAB 48
Sana-DAB 4
Sana-DAB 5
62/-
Sana-DAB 31
75/60
Sana-DAB 32
76/64
Sana-DAB 36
99/100
Sana-DAB 17
Sana-DAB 16
Sana-DAB 11
99/99
Sana-DAB 14
85/83
Sana-DAB 52
Sana-DAB 29
Sana-DAB 38
Sana-DAB 39
Sana-DAB 43
68/-
Sana-DAB 30
Sana-DAB 56
56/-
Sana-DAB 20
Sana-DAB 37
-/98
Sana-DAB 41
Sana-DAB 57
Sana-DAB 55
Sana-DAB 3
Sana-DAB 34
-/75
Sana-DAB 46
Sana-DAB 47
-/62
Sana-DAB 51
-/93
Sana-DAB 24
Sana-DAB 8
Sana-DAB 19
Sana-DAB 9
Sana-DAB 10
57/-
Sana-DAB 44
52/-
Sana-DAB 27
Sana-DAB 26
Sana-DAB 53
Sana-DAB 25
Sana-DAB 15
Sana-DAB 6
Sana-DAB 7
57/-
Sana-DAB 12
66/64
Sana-DAB 18
Sana-DAB 13
Sana-DAB 21
62/58
Sana-DAB 22
Sana-DAB 23
Sana-DAB 45
Sana-DAB 49
Sana-DAB 28
Sana-DAB 33
65/61
Sana-DAB 35
Sana-DAB 50
0.02

Figure 3.31. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout within Kingscote Lake. Only the alleles coloured in red were detected in this population. This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 0.02 (2%).

DAB 29
1/1

DAB 43
3/2

DAB 38
1/1

DAB 39
10/2

DAB 14
13/5

DAB 17
2/1

DAB 16
1/1

DAB 4
110/11

DAB 52
1/1

DAB 11
6/5

DAB 5
57/11

DAB 30
1/1

DAB 50
2/1

DAB 1
24/9

DAB 33
3/3

DAB 35
1/1

DAB 49
1/1

DAB 31
6/3

DAB 56
1/1

DAB 28
1/1

DAB 36
1/1

DAB 45
2/2

DAB 41
1/1

DAB 55
1/1

DAB 15
1/1

DAB 25
1/1

DAB 32
15/3

DAB 20
27/6

DAB 19
1/1

DAB 7
22/7

DAB 13
1/1

DAB 22
1/1

DAB 6
1/1

DAB 12
3/1

DAB 8
19/6

DAB 9
8/5

DAB 37
1/1

DAB 57
11/4

DAB 18
1/1

DAB 48
3/1

DAB 24
1/1

DAB 10
11/5

DAB 42
1/1

DAB 40
17/2

DAB 26
1/1

DAB 53
3/2

DAB 23
1/1

DAB 54
3/3

DAB 3
18/6

DAB 47
3/3

DAB 46
2/2

DAB 44
1/1

DAB 27
13/3

DAB 2
7/3

DAB 51
1/1

DAB 34
1/1

DAB 21
3/1

Figure 3.32. Network of MH class II β1 nucleotide sequences of lake trout from Kingscote Lake. Only coloured alleles occur within the lake. Genealogical relationships were estimated by the parsimony method of Templeton et al. (1992). All refugia alleles are marked as rectangles. Size of ovals is proportional to the number of individuals sharing a particular allele. The numbers below the allele names inside the ovals are the number of fish with that allele followed by the number of populations/locations that allele was identified in. The single line indicates one mutation between alleles (small circles dividing single lines represent missing alleles). Sana-DAB 4 (marked as the largest rectangle) was chosen to have the highest outgroup probability by the program (Castelloe & Templeton, 1994; Donnelly & Tavaré, 1986).

3.8 Comparison of MH class II β1 alleles in Salmonid fishes

A phylogenetic tree of lake trout, arctic charr, rainbow trout and Atlantic salmon detected multiple significant clusters with bootstrap support (Figure 3.33). Atlantic salmon and rainbow trout alleles clustered into individual groupings with high certainty ranging from 80 – 100% although none of these alleles clustered with any significance to alleles of the *Salvelinus* genus. Three pairings of lake trout and arctic charr alleles clustered together with significant bootstrap support in the 70% certainty range and as high as 100% certainty.

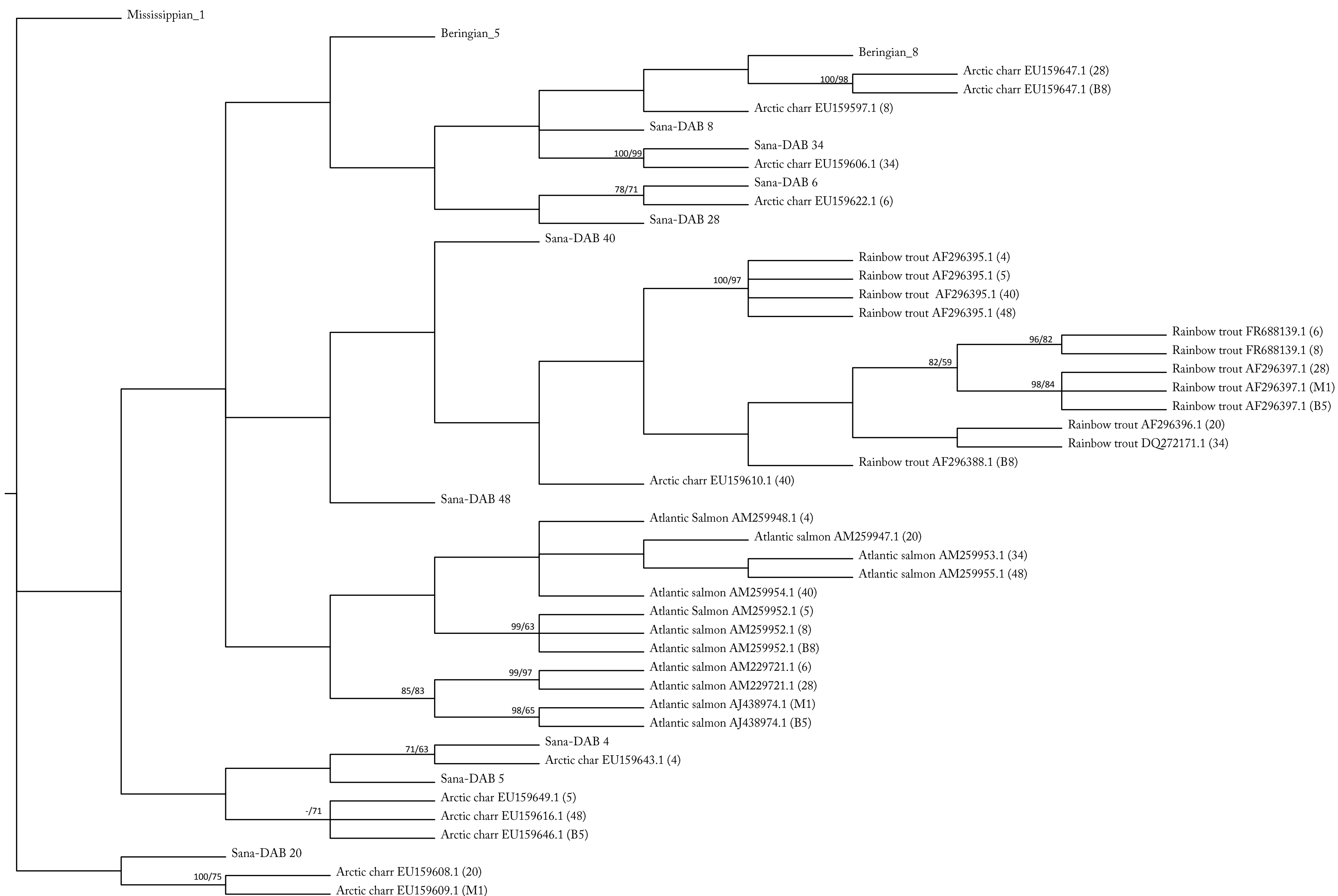

Mississippian_1
Beringian_5
Beringian_8
100/98
Arctic charr EU159647.1 (28)
Arctic charr EU159647.1 (B8)
Arctic charr EU159597.1 (8)
Sana-DAB 8
100/99
Sana-DAB 34
Arctic charr EU159606.1 (34)
78/71
Sana-DAB 6
Arctic charr EU159622.1 (6)
Sana-DAB 28
Sana-DAB 40
Rainbow trout AF296395.1 (4)
100/97
Rainbow trout AF296395.1 (5)
Rainbow trout AF296395.1 (40)
Rainbow trout AF296395.1 (48)
96/82
Rainbow trout FR688139.1 (6)
Rainbow trout FR688139.1 (8)
82/59
98/84
Rainbow trout AF296397.1 (28)
Rainbow trout AF296397.1 (M1)
Rainbow trout AF296397.1 (B5)
Rainbow trout AF296396.1 (20)
Rainbow trout DQ272171.1 (34)
Rainbow trout AF296388.1 (B8)
Arctic charr EU159610.1 (40)
Sana-DAB 48
Atlantic Salmon AM259948.1 (4)
Atlantic salmon AM259947.1 (20)
Atlantic salmon AM259953.1 (34)
Atlantic salmon AM259955.1 (48)
Atlantic salmon AM259954.1 (40)
Atlantic Salmon AM259952.1 (5)
99/63
Atlantic salmon AM259952.1 (8)
Atlantic salmon AM259952.1 (B8)
99/97
Atlantic salmon AM229721.1 (6)
85/83
Atlantic salmon AM229721.1 (28)
Atlantic salmon AJ438974.1 (M1)
98/65
Atlantic salmon AJ438974.1 (B5)
Sana-DAB 4
71/63
Arctic char EU159643.1 (4)
Sana-DAB 5
Arctic char EU159649.1 (5)
-/71
Arctic charr EU159616.1 (48)
Arctic charr EU159646.1 (B5)
Sana-DAB 20
Arctic charr EU159608.1 (20)
100/75
Arctic charr EU159609.1 (M1)
2.0

Figure 3.33. Unrooted maximum likelihood phylogenetic tree of the MH class II β1 nucleotide sequences of lake trout (*Salvelinus namaycush*), arctic charr (*Salvelinus alpinus*), rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*). This tree was generated using TIM1+I+G model of evolution using jModelTest (Posada, 2008). Bootstrap values > 50% are added from neighbour joining/ parsimony analyses to branches that are consistent with the likelihood tree (- represents no available bootstrap value). Scale bar indicates nucleotide divergence of 2.0 (200%).

Chapter 4 – Discussion

The MH class II β gene is one of the most polymorphic nuclear encoded genes known, with the peptide binding region it encodes changing rapidly within populations. This allows a greater resistance to pathogens and a better chance of survival (Grimholt et al., 2003). The MH genes with their variability and local adaptation within individual populations living in close geographical proximity thus provide an excellent opportunity to differentiate recently diverged populations. The phylogenetic trees constructed during this project subdivided into various small clusters of related alleles often representing population-specific sequences. Patterns show that lake trout populations do share lineages of common alleles, but over time sequences can diverge into several local alleles within populations. This study looked at the variability of the MH class II β gene in three different types of lake trout populations: native, hatchery stocking sources or mixed ancestral/stocked populations and the possible differences that could exist between them.

4.1 MH Polymorphism of the β1 domain

A high level of polymorphism was seen in this study with 57 alleles being detected in the 11 different populations that were studied. Even between populations within tight geographical ranges, a large amount of MH allele variation was seen. Only five out of the forty-three alleles in the Dorschner et al. (2000) study were identical to alleles discovered within this study (Figure 3.1). Geographical proximity of sampling locations between the two would have suggested a higher amount of common alleles to be detected. In this study, two sampling locations were within Lake Superior (Slate Islands and Michipicoten Lake) while Dorschner et al. (2000) looked at six different sampling locations within Lake Superior's southern and western ends. Slate

Islands and Michipicoten Lakes are two well-known hatchery stocking sources in the northern and eastern end of Lake Superior and shared 8 common alleles between them. The lack of similarity in alleles between the two studies although surprising was possible for a few reasons. This study and Dorschner et al. (2000) used different sets of primers thus providing a primer bias toward the alleles in each study. Distinct separate populations may have arose within Lake Superior due to multiple refugial groups colonizing Lake Superior following the glaciations (Wilson and Hebert, 1996; 1998), various stocking attempts by different stocks throughout Lake Superior (Kerr, 2001), as well as local adaptation and selection. One of the contributors to the variation in this study may have been a high degree of selection in each individual lake, as the frequency of non-synonymous mutations exceeded those of synonymous mutations in all populations. All populations had D_N/D_S ratios greater than one indicating positive selection for all lakes for an average in this study of 2.33, while Kingscote had the highest ratio at 4.32 (Table 3.2). These D_N/D_S values were lower than those reported for the MH class II β peptide binding regions of previous studies in lake trout from Dorschner et al. (2000) and arctic charr in Conejeros et al. (2013) possibly due to different sets of primers along with a species (arctic charr) that has a wider geographical distribution and abundance.

A main hypothesis in the literature is that balancing selection acts on MH genes to help maintain variation and multiple alleles within a population. As the fish move to new environments, exposure to novel pathogens and parasites increases the frequency of MH alleles that can recognise them and thus diversity is maintained. Additionally, heterozygotes are able to present to a broader array of antigens than homozygotes therefore they are favoured through heterozygote advantage (Bernatchez and Landry, 2003; Spurgin and Richardson, 2010). In this

study 79% of individuals (170 of 216) were heterozygotes. Parasites try to exploit the MH immune response by mutating away from recognition by the most common MH alleles, providing a selective advantage to new rare-alleles (Bernatchez and Landry, 2003; Spurgin and Richardson, 2010). Although selection may be one of the forces maintaining polymorphism in the β1 domain, it is not the only force affecting the gene.

Another driver of variability and polymorphism in MH genes can be behavioral mechanisms, like mate choice. In this case, reproductive behavior favours pairing of mates with different MH allele compositions (Penn and Potts, 1999). This sexual selection provides the offspring with increased pathogen resistance as they are heterozygotes with a greater protective ability against a broader range of pathogens than homozygotes (Landry et al., 2001; Forsberg et al., 2007). MH diversity is also used as a mechanism to prevent inbreeding from occurring as salmonids have the ability to differentiate between kin and non-kin when mating, creating more variability (Olsen et al, 1998, Olsen et al, 2002, Rajakaruna et al. 2006). Specifically the MH class II gene has been shown to influence kin discrimination and recognition in juvenile brook trout (*Salvelinus fontinalis*) and Atlantic salmon (*Salmo salar*) (Rajakurna et al., 2006).

Lake trout also potentially have historical factors diminishing their variability and polymorphism. Lake trout are a glacial relict species, thus their present day populations have been strongly influenced by Pleistocene glaciation events and demographics. Many populations may have been formed from small refugia gene pools following the most recent glaciations, limiting the amount of genetic diversity within these populations and the time available for mutations to build up. As lake trout populations form from small gene pools, particularly

landlocked populations, genetic drift can become an influential force affecting the make-up of the populations causing them to diverge and lose variability. The force of genetic drift is directly linked to the size of the population it is acting on with smaller populations being affected much more strongly. In this study, many of the populations are located in smaller lakes within Ontario and therefore genetic drift would have been a much more prevalent force overpowering selection shaping the gene pools of each population.

4.2 MH genes as a Phylogeographic Marker and Refugia Lineages

In this study, the MH class II β gene was examined for possible use in phylogeography, similar to the application of mtDNA and μsats as molecular markers in past studies. Previous studies (Wilson & Hebert, 1996; 1998) used mtDNA as a phylogeographic marker to differentiate lake trout populations as being derived from six possible refugia from the Pleistocene glaciations: two Mississippian, one Atlantic, one Nahanni, one Beringian and one Montana. One of the goals of this project was to see if MH genes would differentiate the populations from five main refugia in a similar manner, or if they would provide a different picture.

The refugia sample sizes that could be obtained and analysed for this study were small making the differentiation of the refugia groups difficult to see when analysed with the study samples. A maximum likelihood tree based on all the alleles within the study including refugial alleles (Figure 3.5) did not group alleles clearly into refugial groups. The parsimony network of the study and refugial alleles (Figure 3.6) did not show any clear patterns either. It was expected that since the majority of Ontario populations were largely founded by a Mississippian source

that the alleles would cluster well with the Mississippian refugial samples but this was not the case. The majority of alleles in this study grouped with the other refugia groups within the maximum likelihood tree and network (Figure 3.5 and 3.6). This is more likely a consequence of the limited refugial sample sizes that were used in this study, especially for the Mississippian refugia. Although, it is not surprising that many of the populations within Ontario contain alleles from various refugial groups as secondary contact among refugial groups would have occurred through the central portion of the species range especially within areas where the proglacial lakes once existed (Wilson and Hebert, 1996).

The best separation of the refugia groups was seen in the maximum likelihood tree based on the most common alleles in each refugia population (Figure 3.9b) providing the cleanest separation of the three main clades by geographical region. Ontario populations sampled were Hawley and Confederation Lake which were both founded by individuals from a single refugia group, Mississippian, based upon mtDNA (Wilson and Hebert 1996; 1998). These two populations had only one common allele (Mississippian 1) that grouped separately from all other refugia alleles, albeit not with high bootstrap confidence. Apart from this, two separate main clades formed.

In the first clade, a grouping of alleles from Seneca Lake in New York (Atlantic 2 and 3), Island and Toolik Lakes in Alaska (Beringian 1 and 3), Ugashik Lake in Alaska (Nahanni 2 and 3) and Kusawa Lake in the Yukon (Nahanni 2 and 3). This was very similar to Wilson & Herbert's (1996, 1998) six lineage theory with the separation of Mississippian (A), Atlantic (B) and Nahanni (D) into one major clade. Upon bootstrapping analysis of the neighbour joining and

maximum parsimony trees, Atlantic and Beringian alleles grouped with 100% and 97% certainty (Figure 3.9b). Beringian and Nahanni as well as Atlantic and Beringian shared common alleles which may be the result of the fact that common alleles were particularly successful in those areas and/or that they are derived from stocks that shared common ancestral lineages. The extent of genetic similarity and divergence among the Atlantic with Beringian and Nahanni refugia demonstrated no clear correlation with their geographic location similar to what Wilson and Hebert (1998) discovered with their mtDNA analysis. The alleles were closely related yet the three refugia occupy opposite ends of the country geographically. Secondary contact among these lineages during glacial cycles could have caused alleles to disperse or mix despite the geographical distance between the Atlantic and Beringian/Nahanni refuges being quite great. These glacial refugial groups probably originated from pre/mid-Pleistocene populations that were separated by glacial advances followed by divergence within different refugia (Wilson and Hebert, 1998).

In the second clade, a grouping of alleles from Waterton Lake in Alberta (Montana 1 and 2) and alleles from Island (Beringian 2), Toolik (Beringian 2) and Ugashik Lakes in Alaska (Nahanni 1 and 4) and Kusawa Lake in the Yukon (Nahanni 1 and 4) were present. Upon bootstrapping analysis, two groups of alleles emerged: one group of common Nahanni and Beringian alleles and another group of Montana and Nahanni alleles clustered with 92% and 73% confidence values (Figure 3.9b). The existence of common alleles between these populations demonstrates that they originated from pre/mid-Pleistocene populations and that common ancestral alleles probably persist in these populations, along with some secondary contact between the groups in interglacial periods. The branching pattern of the clusters within

this clade demonstrates a clearer correlation with the geographic location of these populations and ancestral refugia groups not seen in the clade discussed above.

The Seneca population lake trout in this study is a stocking source containing both Mississippian and Atlantic haplotypes from mtDNA studies (Wilson and Hebert, 1996; 1998). This population has been heavily supplemented through stocking dating all the way back to 1984 using unknown source population(s) (OMNR, 2005) which may have caused mixed evolutionary origins. This may explain the inconsistencies seen in Figures 3.9a from 3.9b where Atlantic alleles were seen in all the major clusters grouped with all five main refugia within the maximum likelihood tree albeit with no significance or may simply be the result of small sampling sizes.

4.3 Does the salmonid MH class II β gene exist as multiple loci?

A common observation in past MH studies (Conejeros et al., 2008; Conejeros et al., 2013; Dixon et al., 1996; Dorschner et al., 2000; Kruiswijk et al., 2004, 2005; Miller et al., 1996; Miller et al., 1997; Miller et al., 2001; Noakes et al., 2003) has been the number of alleles per individual detected in population studies. For many of these studies, this has been attributed to the high levels of polymorphism in MH genes without looking into other possible causes. Dorschner et al. (2000) did suggest the reason they amplified such a high number of alleles per individual could be due to the amplification of more than one locus. However, they did state that no more than two alleles were detected in any individual and concluded that if indeed the alleles identified were derived from duplicate loci two major branches should be evident on the phylogenetic tree, which was not the case in their study. A shortcoming of their study was they would probably not find more than two alleles in most individual fish because they did not pick

enough clones: only two clones per individual were picked for some populations and then two to six clones for other populations were picked and sequenced. Thus, there is a reasonable chance that many alleles may have been missed in this study when only two clones were picked (Lenz and Becker, 2008).

In this thesis, every individual fish had a minimum of six clones picked and sequenced. It was determined that approximately 30% (63 of 216 individuals) had three or more alleles (Table 3.1), despite the fact that precautions to minimize the chances of false alleles occurring were taken. There was variation within the different populations, as a few populations had as many as 10 fish of the 20 analyzed showing three or more alleles while other populations had as little as two fish of the 20 analyzed showing three or more alleles (Table 3.1). Artifacts can be produced during PCR amplification and give arise to heteroduplex and chimeric amplicons which do not represent *bona fide* alleles (Kanagawa, 2003). Heteroduplexes become mosaic sequences from two parent heterologous sequences as a result of the *Escherichia coli* mismatch repair system during cloning and chimeras are amplicons that contain sequence motifs from two different alleles (Kanagawa, 2003). The PCR protocol used here, unlike previous studies, was designed to minimize the formation of PCR artifacts when amplifying MH class IIβ. More support for the presence of two loci was revealed in the phylogenetic analysis of the study. The maximum likelihood tree of all the alleles in this study (Figure 3.7) showed two very distinct main clades present with a few sequence outliers. This pattern was even seen in a smaller scale in some of the individual population phylogenetic trees in the study. As stated by Dorschner et al. (2000), if alleles from two different loci were being amplified, two distinct major branches should be seen in the phylogenetic tree as demonstrated here (Figure 3.7).

The evidence for multiple loci above is intriguing, but why do some populations display it more than others? A simple answer would be that the second locus of certain populations may be more easily amplified as a result of primer bias. Depending on the similarity of the loci within a population the primer set may be unreliable in the amplification of alleles from both loci. Another explanation is possible, which is that lake trout class II β genes are linked and arranged in haplotypes that may vary in their number of loci. An example of this was seen in a study that did MH typing on cichlid fishes of the Great East African Lakes (Malaga-Trillo, 1998). The typing of the cichlid fishes demonstrated that the class II β gene is arranged in haplotypes that could possibly vary in their number of loci. The number of functional loci in any species may vary but will be low as it is believed that large numbers of functional loci reduce the size of a species' T-cell repertoire and its ability to respond to parasites (Malaga-Trillo, 1998).

4.4 Genetic Profiles of Stocking Sources

Natural selection along with genetic drift, mutation and gene flow within lakes can genetically mould populations. Stocks are known as populations with local adaptations resulting from exposure to the specific environments they live in (Cooke et al., 2001). Stock management works by managing the individual subgroup populations (lakes) of a species like lake trout rather than the whole species. In the proper circumstances fish stocking can be used as an effective management tool but can also be dangerous if the stocks are not well matched to the physical environment and native fish population (OMNR, 2005).

In this study, four hatchery stocking sources were investigated: Manitou, Michipicoten, Seneca and Slate Islands. There were only two common alleles that were detected within all four stocking sources with Manitou and Slate Islands having six alleles and Seneca and Michipicoten having three alleles different from the other stocking sources. These differences in allelic compositions are probably the result of local adaptations to the native environments of these stocking sources.

The Michipicoten and Slate Islands stocking sources' origins are native to Lake Superior and for this reason both stocks have been used in the upper Great Lakes. In Lake Superior, Michipicoten is used for the eastern end close to its native location while Slate Islands is used for the western end of the lake (OMNR, 2005). The MH genetic profiles in this study showed similarity between the two sources as they contained the same 7 allelic groups (Figure 3.20 and 3.22), both missing alleles from the main groupings of Sana-DAB 20 and 40. The differences in nucleotide diversity between the sources was less than 1%, while Slate Islands had a higher haplotype diversity and allelic richness with the Michipicoten strain having more mutational steps between the most common alleles. This MH analysis agrees with the allozymes data (Ihssen et al., 1988) that these two stocks genetic profiles are similar with only a few differences, such that both sources could be stocked into the same geographical location within the upper Great lakes.

The two major allele groups missing within Michipicoten and Slate Islands (Sana-DAB 20 and 40) were detected in both Manitou and Seneca stocks. Manitou was the stocking source containing the most major allele groups containing 8 of the 9 allele groups within this study

(Figure 3.24). This source also had the greatest number of different alleles, but a lot of variability with only two allele frequencies above 0.1. A previous study using allozymes by Ihssen et al. (1988) determined the Manitou stock to be genetically distinct from the other stocks in the study. For that reason, this strain has been used for rehabilitative purposes in hard water, inland lakes with communities and physical characteristics similar to Lake Manitou (OMNR, 2005). The MH genetic profile of Manitou reinforced these findings as this stock was missing the Sana-DAB 1 and 11/14 major groupings that were seen in both Michipicoten (Figure 3.20) and Slate Island (Figure 3.22) stocks and also contained many differences from the Seneca population. Thus the Manitou population possessed a unique genetic composition in comparison to the other stocking sources.

The most unique stocking source of the four studied was Seneca as it was missing three main allele groups (Sana-DAB 3, 7 and 8) that were seen in all other stocking sources. As a result, the Seneca stock contained the lowest allelic richness of stocking sources: only 6 major allele groupings with a total of 12 alleles. The Seneca stock did, however, have the highest nucleotide diversity of the study, a higher haplotype diversity than Michipicoten and Slate Island and the highest D_N/D_S ratio of the stocking sources suggesting natural selection as a main driving force maintaining genetic variation within this stock. The distinct MH genetic profile of Seneca may be the result of its mixed ancestry along with different geographical origins from the other stocks. Due to the uniqueness of Seneca Lake's genetic profile, caution needs to be taken when using this stock for stocking purposes. The OMNR have tried using this stock in the restoration efforts within Lake Ontario, since the extinction of native Lake Ontario lake trout, with some success. It is unknown whether the Seneca stock is genetically similar to the extinct native lake

trout within Lake Ontario but they do share similar geographical environments which may contribute to the success of the lake trout restoration efforts.

4.5 Genetic Profiles of Native Populations

All the native populations contained introgressed alleles commonly occurring within other populations throughout the study. The native populations of Hogan, Crystal, Dickey and Macdonald did, however, differ from one another in their genetic profiles. Each also contained genetic profiles that differed from the stocking sources or stocked populations in that they contained fewer highly common alleles but more unique or alleles detected in a lower number of populations. All the native populations live in smaller isolated landlocked lakes compared to the others populations in this study making them highly susceptible to genetic drift greatly affecting the make-up of the populations causing them to diverge and lose variability over time.

Crystal and Dickey Lake populations had high allelic richness with the highest totals within the study and alleles within 7 of the 9 groupings. Crystal Lake, the smallest of the native lakes, contained a very diverse gene profile with frequencies spread out among all 17 alleles in the population compared to Hogan and Dickey where three alleles made-up 50% of their populations. Crystal Lake had diverging sequences within its population, as Sana-DAB 15 and 16 with no significance and Sana-DAB 17 with high bootstrap confidence separated into their own subgroups away from highly common alleles (Figure 3.12). Thus these alleles show the start of population-specific adaptation to their local environment and the potential pathogens present. Dickey Lake contained unique alleles, Sana-DAB 55 and 56 which were between mutational steps from common alleles. This is an unusual direction of evolution is not standard and is either

the result of gene dispersal or remnants of ancestral allele groups that once existed but have been removed. Dickey Lake contained the lowest nucleotide diversities but highest haplotype diversity of native populations with alleles located within 7 of the 9 allelic groupings. The parsimony network showed the lowest number of mutational steps between the most distant alleles. Dickey Lake may have originated from a smaller gene pool than some of the other native populations making them more susceptible to effects like genetic drift causing the population to lose alleles from rare groupings producing a loss of genetic variability.

The native profile of Macdonald Lake, though distinct, seems to be a population that also shows limited diversity due to population biology or environmental factors. This population had the lowest allelic richness and number of MH class II β alleles in the study with only four major allele groupings represented. Only four alleles made up almost 90% of this population's allelic composition, but it had the highest nucleotide diversity behind Opeongo and Seneca Lakes. Sana-DAB 40 was the most frequent allele in the population, which was different from most populations in this study where Sana-DAB 4 was the most common. Macdonald Lake's small genetic profile contains a distinct and unique makeup different from other populations.

4.6 Genetic Profiles of Stocked Populations

Stocking is one of the techniques that can be used to enhance and rejuvenate naturally reproducing native fish for habitats and populations of concern. In some habitats where species become extinct, stocking strains are introduced to try to restore the population that once lived in that habitat. Hatchery fish can differ genetically from wild native populations, making stocking a complicated, often questioned practice as it is not always clear whether it will enhance wild

populations or destroy them. Three populations in this study, Simcoe, Opeongo and Kingscote Lakes, have been stocked at various times throughout their history.

Simcoe Lake is a population that has been stocked many times and as a result any unique genetic profile that population once may have had has been removed. The genetic profile of Simcoe contained a total of 10 MH class II β alleles, nine of which are commonly within stocking sources (Table 3.5) containing the 2nd lowest allelic richness of the study as well as haplotype and nucleotide diversities below average (Table 3.2). No unique alleles were discovered in this population meaning there is probably little or no local adaptation to this environment taking place. Based on its genetic composition, Simcoe Lake is missing alleles from two major groupings which commonly occurred within the Manitou and Seneca stocks but contains Sana-DAB 10 which is seen in the Manitou stock and not any others. Thus, extensive stocking within Simcoe probably has not been from the Seneca stock but more likely has been from the Michipicoten and Slate Islands stocking strains which also do not contain those two allele groupings within their genetic profiles. It is possible that Manitou may have been stocked into Simcoe Lake at one time and only remnants of introgression remain (Sana-DAB 10) or a different stock containing Sana-DAB 10 was used or this allele is native to the area.

Kingscote Lake seems to be a population that has adapted locally to its environment and continues to do so, making it an interesting population to study. It has been seen as a unique population that is both visually and genetically distinct from neighbouring populations (Wilson and Mandrak, 2004) as well as shown “to conform to a population-level native genetic profile regardless of continuous and substantial historical stocking” (Halbisen and Wilson, 2009). In this

MH study, Kingscote Lake had the least total number of alleles within the population (9), with three of them being unique. Each of the six common alleles in this population occurred throughout the four native populations as well as the stocking sources in this study (Table 3.5). Therefore, it is possible that even though Kingscote has been historically stocked no stocked alleles or at a minimum very few have been incorporated into its genetic profile through stocking. This may be the result of non-compatible stocks not matched correctly to this unique native population and environment. Kingscote had the lowest haplotype and nucleotide diversities but the highest D_N/D_s ratio within the study suggesting positive selection is occurring. The lake trout native to Kingscote may have adapted to specific environmental pathogens present while other stocks may not have the ability to do so. This may be a direct result of the allele, Sana-DAB 8 having the highest frequency in the population rather than the typical Sana-DAB 4 in most populations in this study. The differences within Kingscote may also be the result of its small, landlocked population size allowing genetic drift to overpower natural selection causing the population not only to diverge from others but reducing its genetic variability.

Opeongo Lake has been extensively stocked although it has been questioned whether this process has been successful or not. Opeongo contains remnants of introgression having unique alleles from both Michipicoten and Manitou stocks (Table 3.5). Opeongo had one of the largest gene pools in the study with alleles from all major groupings being present and contributing to high allelic richness, haplotype and nucleotide diversities and number of unique alleles. Thus the MH class II β genetic profile for Opeongo may not prove the success of stocking as the extensive gene pool and diversity within Opeongo could also be the result of its strong healthy native profile rather than an enhanced profile as a result of stocking. Common alleles within Opeongo

and stocking sources are commonly detected within all populations studied whether they are native or stocked thus it is difficult to determine whether these alleles occurred naturally in Opeongo Lake or not. Diverging alleles within Opeongo Lake were seen with a unique clade (Sana-DAB 21, 22, 23 and 49) within the maximum likelihood tree (Figure 3.29) but with no bootstrap support. These diverging alleles could be local variants of the DAB 7 grouping that once existed (Figure 3.30).

4.7 Trans-Species Polymorphism

Trans-species polymorphism which was first shown to be a characteristic of mammalian MHC alleles (Mayer et al., 1988) appears to be common in fishes like cichlids (Klein, Satta, & O'hUigin, 1993), Chinook and Pacific salmon (Garrigan and Hedrick, 2001) as well as Atlantic salmon and brown trout (Stet et al., 2002). In this study lake trout were compared with arctic charr, rainbow trout and Atlantic salmon for transspecific lineages. The Atlantic salmon and rainbow trout alleles clustered together with high confidence but none of them grouped with the *Salvelinus* (lake trout or arctic charr) significantly as they all diverged too long ago, 10 – 20 million years ago. Although the two *Salvelinus* genera, lake trout and arctic charr clustered in three pairings with high certainty (Figure 3.33) demonstrating trans-species evolution as these species diverged approximately three million years ago.

Chapter 5 – General Conclusions

5.1 Conclusions

This study used the MH class II β gene as a non-neutral population marker to differentiate populations of lake trout within Ontario and New York. This marker not only differentiates populations based on genetic variability of individuals but also on their interactions with the environment due to its role in detecting the presence of pathogens. High levels of genetic diversity were seen throughout this study with population gene pools impacted by the Pleistocene glaciation history followed by strong influential population forces like genetic drift and natural selection acting on them. Traces of local adaptation and selection were seen in the MH class II β sequences of the studied populations with a large amount of MH polymorphism demonstrated. The MH class II β gene showed promise providing a phylogeographic signal, but many unanswered questions still remain. The genetic profiles of the Ontario/New York populations showed differences in unique characteristics even though most came from common ancestry. Historical stocking within the populations has shown very limited success as stocking strains do not seem to have been able to incorporate their genes into populations like Opeongo and Kingscote with great success or have removed the native profile from populations like Simcoe, perhaps leaving a maladapted population in place.

5.2 Future Work

The use of a selective gene as a marker provides new and exciting insight into the population genetics of lake trout populations. In order to assess the significance and comparability of this marker to others a few questions need to be resolved. Firstly, it needs to be

determined whether or not the lake trout MH class II β gene exists as multiple loci to allow for a proper population-level genetic analysis. Solving this problem would allow the MH class II β gene to be used as a population marker with comparable statistics and analysis done as performed on other genetic markers. There are a few techniques that can be carried out to help confirm this and provide definitive proof. A southern blot can be done to determine the copy number of the MH class II β gene. This will provide proof of whether lake trout do indeed contain multiple loci at the MH class II β gene or not. If they do, tests will need to be done to determine the inheritance of the alleles or which alleles belong to which locus. The inheritance of the alleles present can be determined through MH gene mapping using family studies. This is done by looking at the segregation of alleles from the parents to determine which are being inherited independently. Another way to tackle the problem would be through next generation sequencing but this route is very costly. If no solution can be obtained, a new way to compute important population statistics such as F-statistics when using MH genes as population markers needs to be determined. Many population models are designed based on specific rules and parameters that the MH genes break because it is not known how many loci the genes contain and which alleles belong to which loci. For this reason, conventional population statistics such as Hardy-Weinberg cannot be performed and a new form or variation of population statistics/models needs to be designed.

Following the resolution of the technical issues above, a more extensive look at the MH class II β alleles discovered in the refugia groups and whether or not MH class II β genes can carry a phylogeographic signal will need to be carried out. This study only briefly looked at the potential of MH as a phylogeographic marker and with limited sample sites and individuals

examined in some populations, conclusions were difficult to obtain. Finally, future studies should be expanded to look at locations outside of Ontario and instead focus on lake trout populations across Canada. The use of the MH class II β gene as a population marker may provide insight into the adaptation populations are undergoing within their local environments. This knowledge will provide fish population managers the ability to identify and assess potential adaptive genetic diversity within and among populations which cannot be accomplished with conventional neutral marker systems.

References

Aguilar, A., & Garza, J. C. (2006). A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). *Molecular Ecology*, 15, 923–37.

Araki, H., Cooper, B., & Blouin, M. S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, 318, 100–103.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18, 489–522.

Bacon, L. (1987). Influence of the major histocompatibility complex on disease resistance and productivity. *Poultry Science*, 66, 802–811.

Bacon, L., Hunt, H., & Cheng, H. (2000). A review of the development of chicken lines to resolve genes determining resistance to diseases. *Poultry Science*, 79, 1082–1093.

Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Bology*, 16, 363–377.

Bingulac-Popovic, J., Figueroa, F., Sato, A., Talbot, W. S., Johnson, S. L., Gates, M., Klein, J. (1997). Mapping of Mhc class I and class II regions to different linkage groups in the zebrafish, *Danio rerio*. *Immunogenetics*, 46, 129–134.

Black, G. A. (1983a). Origin, distribution, and postglacial dispersal of a swimbladder nematode, *Cystidicola stigmatura*. *Canadian Journal of Fisheries and Aquatic Sciences*, 40, 1244–1253.

Black, G. A. (1983b). *Cystidicola farionis* (Nematoda) as an indicator of lake trout (*Salvelinus namaycush*) of Bering ancestry. *Canadian Journal of Fisheries and Aquatic Sciences*, 40, 2034–2040.

Brannon, E. L., Amend, D. F., Cronin, M. A., Lannan, J. E., LaPatra, S., McNeil, W. J., Westers, H. (2011). The controversy about salmon hatcheries. *Fisheries*, 29, 12–31.

Briggs, J. (1986). Introduction to the zoogeography of North American fishes. In C. Hocutt & E. O. Wiley (Eds.), *The Zoogeography of North American Freshwater Fishes*. John Wiley & Sons Inc. pp. 1–16.

Brown, W. M., George, M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *76*, 1967–1971.

Burgess, S. C., Basaran, B. H., & Davison, T. F. (2001). Resistance to Marek's disease Herpesvirus-induced Lymphoma is multiphasic and dependent on host genotype. *Veterinary Pathology*, 38, 129–142.

Castelloe J., Templeton AR (1994). Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution*, 3, 102–113.

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657-1660.

Conejeros, P., Phan, A., Power, M., Alekseyev, S., O'Connell, M., Dempson, B., & Dixon, B. (2008). MH class II alpha polymorphism in local and global adaptation of arctic charr (*Salvelinus alpinus* L.). *Immunogenetics*, 60, 325–37.

Conejeros, P., Power, M., Alekseyev, S., Caron, F., Gantner, N., & Dixon, B. (2012). Global MH Class II β polymorphism in arctic charr (*Salvelinus alpinus* L.) and the adaptation to local environments. *Immunogenetics*.

Consortium, MHC Sequncing. (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature*, 401, 921–923.

Cooke, S. J., Kassler, T. W., & Philipp, D. P. (2001). Physiological performance of largemouth bass related to local adaptation and interstock hybridization: implications for conservation and management. *Journal of Fish Biology*, 59, 248-268.

Crossman, E. J., & McAllister, D. E. (1986). Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungava Bay and the Arctic Archipelago. In C. H. Hocutt & E. O. Wiley (Eds.), *The Zoogeography of North American Freshwater Fishes*. John Wiley & Sons Inc. pp. 53–104.

Danchin, E., Vitiello, V., Vienne, A., Richard, O., Gouret, P., McDermott, M. F., & Pontarotti, P. (2004). The major histocompatibility complex origin. *Immunological Reviews*, 198, 216–232.

De Eyto, E., McGinnity, P., Consuegra, S., Coughlan, J., Tufto, J., Farrell, K., Stet, R. J. M. (2007). Natural selection acts on Atlantic salmon major histocompatibility (MH) variability in the wild. *Proceedings. Biological sciences / The Royal Society*, 274(1611), 861–9.

Dixon, B, & Stet, R. J. (2001). The relationship between major histocompatibility receptors and innate immunity in teleost fish. *Developmental and Comparative Immunology*, 25, 683–699.

Dixon, B, van Erp, S. H., Rodrigues, P. N., Egberts, E., & Stet, R. J. (1995). Fish major histocompatibility complex genes: an expansion. *Developmental and Comparative Immunology*, 19, 109-133.

Dixon, Brian. (2008). René Stet's impact on the study of teleost major histocompatibility genes: evolution from loci to populations. *Immunogenetics*, 60, 77–82.

Dixon, B., Nagelkerke, L. A. J., Sibbing, F. A., Egberts, E., & Stet, R. J. (1996). Evolution of MHC class II β chain-encoding genes in the Lake Tana barbel species flock (*Barbus intermedius complex*). *Immunogenetics*, 44, 419 – 431.

Donnelly P., Tavaré S. (1986). The ages of alleles and a coalescent. *Advances in Applied Probability*, 18, 1–19.

Dorschner, M. O., Duris, T., Bronte, C. R., Burnham Curtis, M. K., & Phillips, R. B. (2000). High levels of MHC class II allelic diversity in lake trout from Lake Superior. *The Journal of Heredity*, 91, 359–363.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.

Engelhard, V. H. (1994). Structure of peptides associated with MHC class I molecules. *Current Opinion in Immunology*, 6, 13–23.

Evans, D., & Willox, C. C. (1991). Loss of exploited, indigenous populations of lake trout, *Salvelinus namaycush*, by stocking of non-native stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 48, 134–147.

Flajnik, M. F., & Du Pasquier, L. (1990). The major histocompatibility complex of frogs. *Immunological Reviews*, 113, 47–63.

Flint, R. F. (1971). *Glacial and Quaternary Geology*. New York: John Wiley & Sons Inc.

Forsberg, L. A., Dannewitz, J., Petersson, E., & Grahn, M. (2007). Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout - females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology*, 20, 1859-1869.

Fosse, V. M., Hordvik, I., Bergersen, O., Endresen, C., & Lie, O. (1991). Cloning and characterization of the MHC class I and class II genes from Atlantic salmon (*Salmo salar*). *Developmental & Comparative Immunology*, *15*, S38.

Freeland, J. R. (2005). *Molecular Ecology*. England: John Wiley & Sons Inc.

Garrigan, D., & Hedrick, P. W. (2001). Class I MHC polymorphism and evolution in endangered California Chinook and other Pacific salmon. *Immunogenetics*, 53, 483–489.

Glamann, J., Juul-Madsen, H., & Simonsen, M. (1991). Isolation of a cDNA clone encoding MHC class II β Chain in Rainbow Trout. *Developmental & Comparative Immunology*, 15, S37.

Graser, R., O'hUigin, C., Vincek, V., Meyer, A., & Klein, J. (1996). Trans-species polymorphism of class II MHC loci in danio fishes. *Immunogenetics*, 44, 36–48.

Grewe, P. M., & Hebert, P. D. N. (1988). Mitochondria DNA diversity among broodstocks of the lake trout, *Salvelinus namaycush. Canadian Journal of Fisheries and Aquatic Sciences*, 45, 2114–2122.

Grewe, P. M., Krueger, C. C., Aquadro, C. F., Bermingham, E., Kincaid, H. L., & May, B. (1993). Mitochondrial DNA variation among lake trout (*Salvelinus namaycush*) strains stocked into Lake Ontario. *Canadian Journal of Fisheries and Aquatic Sciences*, 50, 2397–2403.

Grimholt, U., Hordvik, I., Fosse, V. M., Olsaker, I., Endresen, C., & Lie, O. (1993). Molecular cloning of major histocompatibility complex class I cDNAs from Atlantic salmon (*Salmo salar*). *Immunogenetics*, 37, 469–473.

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Stet, R. J. M. (2003). MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics*, 55, 210–219.

Halbisen, M., & Wilson, C. (2009). Variable introgression from supplemental stocking in southern Ontario populations of lake trout. *Transactions of the American Fisheries Society*, 137, 699–719.

Hall, A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleotide Acids Symposium Series,* 41, 95-98.

Harstad, H., Lukacs, M. F., Bakke, H. G., & Grimholt, U. (2008). Multiple expressed MHC class II loci in salmonids; details of one non-classical region in Atlantic salmon (*Salmo salar*). *BMC Genomics*, 9, 193.

Hashimoto, K., Nakanishi, T., & Kurosawa, Y. (1990). Isolation of carp genes encoding major histocompatibility complex antigens. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 6863–6867.

Hemmer, B., Kondo, T., Gran, B., Pinilla, C., Pascal, J., Tzou, A., Martin, R. (2000). Minimal peptide length requirements for CD4(+) T cell clones: implications for molecular mimicry and T cell survival. *International Immunology*, 12, 375–83.

Hill, A. V. S. (1991). HLA associations with malaria in Africa: some implications for MHC evolution. J Klein & D. Klein (Eds.), *Molecular Evolution of the Major Histocompatibility Complex*. Berlin: Springer. pp. 403–419.

Hordvik, I., Grimholt, U., Fosse, V. M., Lie, O., & Endresen, C. (1993). Cloning and sequence analysis of cDNAs encoding the MHC class II beta chain in Atlantic salmon (*Salmo salar*). *Immunogenetics*, 37, 437–441.

Hughes, A. L., & Nei, M. (1988). Patterns of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, *335*.

Hurst, L. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends in Genetics*, 18, 486-487.

Ihssen, P. E., Casselman, J., & Martin, G. (1988). Biochemical genetic differentiation of lake trout (*Salvelinus namaycush*) stocks of the Great Lakes region. *Canadian Journal of Fisheries and Aquatic Sciences*, 45, 1018–1029.

Juul-Madsen, H. R., Glamann, J., Madsen, H. O., & Simonsen, M. (1992). MHC class II beta-chain expression in the rainbow trout. *Scandinavian Journal of Immunology*, 35, 687–94.

Kasaharaa, M., Flajnikb, M. F., & Natoric, T. (1995). Evolution of the major histocompatibility complex : a current overview. *Transplant Immunology,* 3, 1-20.

Kerr, S. (2001). *Lake Trout Stocking in Inland Lakes: An Annotated Bibliography and Literature Review*. Peterborough: Ontario Ministry of Natural Resources.

Khan, N. Y., & Qadri, S. U. (1971). Intraspecific variations and postglacial distribution of lake char (*Salvelinus namaycush*), 28, 465–476.

Klein, J. (1986). *Natural history of the major histocompatibility complex*. New York: Wiley.

Klein, J. (1989). The MHC trans-species hypothesis: in the discussion period. *Immunology Supplement*, *2*, 36–39.

Klein, J, & Sato, A. (1998). Birth of the major histocompatibility complex. *Scandinavian Journal of Immunology*, 47, 199–209.

Klein, J, Satta, Y., & O'Huigin, C. (1993). The molecular descent of the major histocompatibility complex. *Annual Review of Immunology*, 11, 269–295.

Klein, Jan. (1987). Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Human Immunology*, 19, 155–162.

Kruiswijk, C. P., Hermsen, T., Fujiki, K., Dixon, B., Savelkoul, H. F. J., & Stet, R. J. M. (2004). Analysis of genomic and expressed major histocompatibility class I α and class II genes in a hexaploid Lake Tana African "large" barb individual (*Barbus intermedius*). *Immunogenetics*, 55, 770–781. 7

Kruiswijk, C. P., Hermsen, T., van Heerwaarden, J., Dixon, B., Savelkoul, H. F. J., & Stet, R. J. M. (2005). Major histocompatibility genes in the Lake Tana African large barb species flock: evidence for complete partitioning of class II B, but not class I, genes among different species. *Immunogenetics*, 56, 894–908.

Landry, C., & Bernatchez, L. (2001). Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, 10, 2525–2539.

Landry, C., Garant, D., Duchesne, P., & Bernatchez, L. (2001). 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proceedings of the Royal Society Biological Sciences*, 268, 1279-1285.

Langefors, A, Lohm, J., Grahn, M., Andersen, O., & von Schantz, T. (2001). Association between major histocompatibility complex class II β alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proceedings of the Royal Society Biological Sciences*, 268, 479–485.

Langefors, Å., Lohm, J., & von Schantz, T. (2001). Allelic polymorphism in MHC class II β in four populations of Atlantic salmon (*Salmo salar*). *Immunogenetics*, 53, 329–336.

Lenz, T., & Becker, S.(2008). Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci - Implications for evolutionary analysis. *Gene*, 427, 117-123.

Librado, P., Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452.

Lindsey, C. (1964). Problems in zoogeography of the lake trout *Salvelinus namaycush*. *Journal of the Fisheries Research Board of Canada*, 21, 977–994.

Lohm, J., Grahn, M., Langefors, A., Andersen, Ø., Storset, A., & von Schantz, T. (2002). Experimental evidence for major histocompatibility complex-allele-specific resistance to a bacterial infection. *Proceedings of the Royal Society Biological Sciences*, 269, 2029–2033.

Martin, N. V., & Olver, C. H. (1980). The lake charr, *Salvelinus namaycush*. In E. K. Balon (Ed.), *Charrs: Salmonid Fishes of the Genus Salvelinus*. Springer. pp. 205-277.

Mayer, W. E., Jonker, M., Klein, D., Ivanyi, P., Van Seventer, G., & Klein, J. (1988). Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution. *The EMBO Journal*, 7, 2765–2774.

McAllister, D. E., Platania, S. P., Schueler, F. W., Baldwin, M. E., & Lee, D. S. (1986). Ichthyofaunal patterns on a geographic grid. In C. H. Hocutt & E. O. Wiley (Eds.), *The Zoogeography of North American Freshwater Fishes*. John Wiley & Sons Inc. pp. 53–104.

McConnell, T. J., Talbot, W. S., McIndoe, R. a, & Wakeland, E. K. (1988, April 14). The origin of MHC class II gene polymorphism within the genus *Mus*. *Nature* 332, 651-654.

Miller, K. M., Kaukinen, K. H., Beacham, T. D., & Withler, R. E. (2001). Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. *Genetica*, 111, 237–257.

Miller, K. M., & Withler, R. E. (1996). Sequence analysis of a polymorphic Mhc class II gene in Pacific salmon. *Immunogenetics*, 43, 337–351.

Miller, K. M., Withler, R. E., & Beacham, T. D. (1997). Molecular evolution at Mhc genes in two populations of chinook salmon *Oncorhynchus tshawytscha*. *Molecular Ecology*, 6, 937–954.

Miller, M. M., Goto, R., Bernot, a, Zoorob, R., Auffray, C., Bumstead, N., & Briles, W. E. (1994). Two Mhc class I and two Mhc class II genes map to the chicken Rfp-Y system outside the B complex. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 4397–401.

Moritz, C. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, 18, 269–292.

Murphy, K., Travers, P., & Walport, M. (2008). *Janeway's Immunobiology* (7th ed.). New York: Garland Science.

Nakamura, T., Sekizawa, A., & Fujii, T. (1986). Cosegregation of the polymorphic C4 with the MHC in the frog, *Xenopus laevis*. *Immunogenetics*, 23, 181–186.

Nei, M., Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3, 418-426.

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nei, M., & Hughes, A. L. (1991). Polymorphism and evolution of the major histocompatibility complex loci in mammals. In R. K. Selander, A. G. Clark, & T. S. Whittam (Eds.), *Evolution at the Molecular Level*. Sunderland Massachusetts: Sinauer.

Noakes, M. A, Reimer, T., & Phillips, R. B. (2003). Genotypic characterization of an MHC class II locus in lake trout (*Salvelinus namaycush*) from Lake Superior by single-stranded conformational polymorphism analysis and reference strand-mediated conformational analysis. *Marine Biotechnology*, 5, 270–278.

O'Brien, S. J., & Yuhki, N. (1999). Comparative genome organization of the major histocompatibility complex: lessons from the Felidae. *Immunological Reviews*, 167, 133–44.

Olsen, K. H., Grahn, M., Lohm, J., & Langefors, A. (1998). MHC and kin discrimination in juvenile arctic charr, *Salvelinus alpinus*. *Animal Behaviour,* 56, 319-327.

Olsen, K. H., Grahn, M., & Lohm, J. (2002). Influence of MHC on sibling discrimination in arctic char, *Salvelinus alpinus*. *Journal of Chemical Ecology*, 28, 783-795.

OMNR. (2005). Fish stocks catalog. Fish Culture Section, Ontario Ministry of Natural Resources.

Ono, H., O'Huigin, C., Tichy, H., & Klein, J. (1993). Major-histocompatibility-complex variation in two species of cichlid fishes from Lake Malawi. *Molecular Biology and Evolution*, 10, 1060–1072.

Ottova, E., Simkova, A., Martin, J. F., Gouy de Bellocq, J., Gelnar, M., Allienne, J.F., & Morand, S. (2005). Evolution and trans-species polymorphism of MHC class II β genes in cyprinid fish. *Fish & Shellfish Immunology*, 18, 199–222.

Penn, D. J., & Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *American Naturalist*, 153, 145-164.

Pielou, E. C. (1991). *After the Ice Age: The Return of Life to Glaciated North America*. Chicago: University of Chicago Press.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253-1257.

Radforth, I. (1944). Some considerations on the distribution of fishes in Ontario. *Contributions of the Royal Ontario Museum of Zoology*, 25, 1–116.

Rajakaruna, R. S., Brown, J. A., Kaukinen, K. H., & Miller, K. M. (2006). Major histocompatibility complex and kin discrimination in Atlantic salmon and brook trout. *Molecular Ecology,* 15, 4569-4575.

Robinson, J., Mistry, K., McWilliam, H., Lopez, R., & Parham, P. (2011). Marsh SGE. The IMGT/HLA database. *Nucleic Acids Research*, 39, 1171–1176.

Saltercid, L., Kasahara, M., & Flajnikl, M. E. (1994). Hsp70 genes are linked to the *Xenopus* major histocompatibility complex. *Immunogenetics*, 39, 1–7.

Sato, A., Figueroa, F., Murray, B. W., Málaga-Trillo, E., Zaleska-Rutczynska, Z., Sültmann, H., Klein, J. (2000). Nonlinkage of major histocompatibility complex class I and class II loci in bony fishes. *Immunogenetics*, 51, 108–16.

Schaschl, H., Suchentrunk, F., Hammer, S., & Goodman, S. J. (2005). Recombination and the origin of sequence diversity in the DRB MHC class II locus in chamois (*Rupicapra spp.*). *Immunogenetics*, 57, 108–115.

Scott, W. B., Scott, W. B., & Crossman, E. J. (1985). *Freshwater Fishes of Canada*. Fisheries Research Board of Canada.

Shuter, B. J., Jones, M. L., Korver, R. M., & Lester, N. P. (1998). A general, life history based model for regional management of fish stocks: the inland lake trout (*Salvelinus namaycush*) fisheries of Ontario. *Science*, 55, 2161–2177.

Spurgin, L. G., & Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society Biological Sciences*, 277, 979–88.

Steinmetz, M., & Hood, L. (1983). Genes of the major histocompatibility complex in mouse and man. *Science*, 222, 727–733.

Stet, R. J., Johnstone, R., Parham, P., & Wiegertjes, G. F. (1997). The UNMHC of teleostean fish: segregation analyses in common carp and Atlantic salmon. *Hereditas*, 127, 169–170.

Stet, R. J. M., de Vries, B., Mudde, K., Hermsen, T., van Heerwaarden, J., Shum, B. P., & Grimholt, U. (2002). Unique haplotypes of co-segregating major histocompatibility class II α and class II β alleles in Atlantic salmon (*Salmo salar*) give rise to diverse class II genotypes. *Immunogenetics*, 54, 320–31.

Swanson, H. K., Kidd, K. A., Babaluk, J. A., Wastle, R. J., Yang, P. P., Halden, N. M., & Reist, J. D. (2010). Anadromy in Arctic populations of lake trout (*Salvelinus namaycush*): otolith microchemistry, stable isotopes, and comparisons with arctic char (*Salvelinus alpinus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 67, 842–853.

Swofford, D. L. (2003). PAUP*. Phylogenetic analysis using parsimony (and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

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

Tamura, K., Peterson, D., Peterson N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGAS5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731-2739.

Ujvari, B., & Belov, K. (2011). Major histocompatibility complex (MHC) markers in conservation biology. *International Journal of Molecular Sciences*, 12, 5168–5186.

Vitiello, A., Potter, T. A., & Sherman, L. A. (1990). The role of beta 2-microglobulin in peptide binding by class I molecules. *Science*, 250, 1423–6.

Walters, V. (1955). Fishes of western arctic america and eastern Arctic Siberia. *Bulletin of the American Museum of Natural History*, 106, 255–368.

Watts, C. (1997). Capture and processing of exogenous antigens for presentation on MHC molecules. *Annual Review of Immunology*, 15, 821–850.

Wiegertjes, G. F., Egberts, E., & Stet, R. J. (1997). Segregation of MHC class I and class II genes in the common carp (*Cyprinus carpio L.*). *Developmental & Comparative Immunology*, 21, 105.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, 26, 375–400.

Wilson, Chris C, & Hebert, P. D. N. (1996). Phylogeographic origins of lake trout (*Salvelinus namaycush*) in eastern North America. *Canadian Journal of Fisheries and Aquatic Sciences*, 53, 2764–2775.

Wilson, C. C., & Mandrak, N. E. (2004). History and evolution of lake trout in shield lakes: past and future challenges. In J. M. Gunn, R. J. Steedman, & R. Ryder (Eds.), *Boreal Shield Watersheds: lake trout ecosystems in a changing environment*. Boca Raton, Florida: CRC Press. pp. 21–36.

Wilson, Christopher C., & Hebert, P. D. N. (1998). Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. *Canadian Journal of Fisheries and Aquatic Sciences*, *55,* 1010–1024.

Yeager, M., & Hughes, A L. (1999). Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. *Immunological Reviews*, 167, 45–58.

York, I. A., & Rock, K. L. (1996). Antigen processing and presentation by the class I major histocompatibility complex. *Annual Review of Immunology*, 14, 369–396.