USING LANDSCAPE GENETICS TO FORM A PORTRAIT OF A SUCKER: AN ARGUMENT FOR THE CONSIDERATION OF MULTIPLE FACTORS TO CAPTURE THE WHOLE PICTURE OF A SUBARCTIC DENDRITIC METAPOPULATION

by

Sarah Jane Salisbury

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
October 2015

To the family, friends, fish, and felines who made this possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

I tested the relative importance of life history, environmental barriers, dendritic structure, and historical colonization on the neutral genetic structure of a longnose sucker (*Catostomus catostomus*) metapopulation in the Kogaluk River of northern Labrador. Samples were collected from eight lakes, genotyped with 17 microsatellites, and aged using opercula. Lakes demonstrated varying migration rates and genetic differentiation. Isolation by distance was found only when the two most genetically distinct lakes were removed from the analyses, suggesting a lack of migration-drift equilibrium and the importance of historical and contemporary factors in shaping metapopulation structure. Lower allelic richness in the headwaters due to the dendritic structure of the watershed contrasted with high effective population sizes of the south-western headwaters, potentially due to their earlier colonization. Recent colonization, variable migration rates between lakes, long generation times, and upstream migration have stalled achievement of a typical dendritic metapopulation structure and its associated elevated effective size.

LIST OF ABBREVIATIONS AND SYMBOLS USED

AIC_c	Corrected Akaike's Information Criterion
AL	Adult Lifespan
A_R	Allelic Richness
BP	Before Present
CI	Confidence Interval
CL	Cabot
cm	Centimetre
DF	Degrees of Freedom
DIYABC	Do It Yourself Approximate Bayesian Computation
DM2	$(d\mu)^2$ Distance (two sample summary statistic)
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
E	Pairwise Elevation Differences (metres per kilometre)
ESK	Esker
f	Fecundity (number of eggs)
FL	Fork Length
$\hat{F}_{ST}$	A measure of genetic diversity among subpopulations
G	Pairwise Genetic Distances ($\hat{F}_{ST}$)
GH	Genetics H
H2P	Mean Genic Diversity (two-sample summary statistic)
H_e	Expected Heterozygosity
HET	Mean Genic Diversity (one-sample summary statistic)
H_o	Observed Heterozygosity
IBD	Isolation by Distance
K	Number of Populations
km	Kilometre
L1	Lake 1
l_x	Probability of survival to age x
m	Metre

m	Fraction of Migrants
MAS	Metres Above Sea Level
MCMC	Markov Chain Monte Carlo
meta-$\widehat{N}_{\mathrm{e(T+H)}}$	Estimate of Effective Population Size using the Tufto and Hindar (2003) method
mm	Millimetre
mM	Millimolar
m_x	The number of offspring produced by an individual of age x
n	Sample Size
N	DIYABC Effective Population Size
N	Population Size
N2P	Number of Alleles (two-sample summary statistic)
NAL	Mean Number of Alleles (one-sample summary statistic)
$\widehat{N}_{\mathrm{b}}$	Estimate of Number of Breeders
$\widehat{N}_{\mathrm{b(adj2)}}$	Estimate of Number of Breeders accounting for age structure bias
N_{e}	Effective Size
$\widehat{N}_{\mathrm{e}}$	Estimate of Effective Size
$\widehat{N}_{\mathrm{e(adj2)}}$	Estimate of Effective Size accounting for age structure bias
$\widehat{N}_{\mathrm{e(DIYABC)}}$	Estimate of Effective Size calculated using DIYABC
$\widehat{N}_{\mathrm{e(LDNe)}}$	Estimate of Effective Size calculated using LDNe
PCA	Principal Components Analysis
P_{crit}	Minimum Allele Frequency
PCoA	Principal Coordinates Analysis
r	Admixture Rate (DIYABC)
RNA	Ribonucleic Acid
S	Pairwise Elevation Differences (metres)
$\hat{S}$	Robson-Chapman Annual Survivorship Estimate
STG	Strange
SLU	Slushy
t	Time in Generations (DIYABC)

T	Generation Time
TB	T-Bone
TL	Total Length
V2P	Mean Size Variance (two-sample summary statistic)
VAR	Mean Size Variance (one-sample summary statistic)
W	Number of Intermediate Waterfalls
WF	Waterfall
WP	WP152
x	Age
X_{whole}	Number of visible annuli
$Y_{section}$	Age, corrected for annuli hidden under the dense bone region
α	Age at 50% Maturity
ω	Maximum Age
τ	Time since Divergence between Populations
µL	Microlitre
µ	Mutation rate (DIYABC)

ACKNOWLEDGEMENTS

There are many people without whom this work would not have been possible. I first want to thank my supervisor Dr. Daniel Ruzzante for welcoming me into his lab and giving me the opportunity to fly around in helicopters, collecting fish in remote regions of the world. I really appreciate your continued support throughout this project. I would also like to thank Dr. Paul Bentzen and Dr. Mark Johnston for serving on my committee, providing helpful suggestions for the project, and for reviewing my thesis. Thanks also to my external supervisor Dr. Sean Rogers for his review of my thesis, Dr. Christophe Herbinger for serving on my ATC committee and chairing my defense, and Dr. Brian Hall for chairing my ATC committee.

I want to particularly thank Rob Perry and Don Keefe for not only organizing years of sampling in the Kogaluk but for allowing me to experience such a unique and beautiful landscape in person rather than as labels in a vial. Your training in the ways of aging opercula and making toast was also invaluable. Thanks to Shane Hann and Jerry Callahan for your help in the field and the lab and Lorne Pike and Reuben Solomon of Universal Helicopters for flying us around despite not letting me drive.

I am also grateful to the members of our lab. Thank you Greg McCracken for training me in all things lab and glassware related as well as for being a constant sounding board for my ideas on this project. Thanks to Ivan Vera-Escalona for your insights into the mysteries of DIYABC, helping me edit this thesis, and for always being up for a brainstorming session. Angela Fuentes-Pardo, Anahí Jorquera, Hilary Brewis, and Connor Booker; thanks for your continued support, helping me brainstorm, and making our lab such a fun place to work. Thanks also to Ian Paterson and Meghan McBride for your help in the lab, especially when Li-Cors, thermocyclers, and centrifuges were acting up, Kristen Wilson for extracting longnose sucker DNA, and all the other members of the Marine Gene Probe Lab.

Thanks to Chelsea Boaler and Helen McConnell who, despite equivocal feelings for longnose suckers (they are just as cool as whales), were always sympathetic when I started ranting about a poor gel or a stubborn analysis.

I would like to acknowledge NSERC and the Department of Environment and Conservation of Newfoundland and Labrador for providing funding for me to complete this project.

Finally I want to thank Mom, Dad, Daniel, Dave, and Ducati for all of your love and support during this process. Thanks Mom for your help with editing. I especially want to thank Mom and Dad for encouraging me to work hard for the things I love. I couldn't have done it without you.

CHAPTER 1: INTRODUCTION

Landscape genetics provides a powerful theoretical framework for assessing the genetic structure of a metapopulation and identifying its vulnerable subpopulations (Manel et al. 2003, Storfer et al. 2007). River systems are ideal for landscape genetics studies because fully aquatic organisms are confined to the boundaries of the river (Hughes et al. 2009). Since rivers are one-dimensional (Baguette et al. 2013) and have fewer redundant pathways than other metapopulation structures (Peterson et al. 2013) they are less complex and easier to study than two-dimensional terrestrial landscapes. Fish can also demonstrate significant genetic differentiation over a small spatial scale within river landscapes, allowing for decreased sampling effort required to detect those factors that are driving genetic differentiation (Kanno et al. 2011).

Despite their suitability for such studies, river systems have only recently been considered unique from the typical terrestrial landscape within a landscape genetics theoretical framework (Fagan 2002, Campbell Grant et al. 2007). Rivers are often dendritic, composed of a series of bifurcating branches that radiate from a single node (Altermatt 2013) creating a hierarchically-arranged network of habitat areas (Fullerton et al. 2010) which isolates headwaters (Fagan 2002). This isolation can be exacerbated by a downstream bias in migration (Morrissey and de Kerckhove 2009) and the reduced carrying capacities associated with upstream habitats (Carrara et al. 2014). Headwaters are predicted to have a reduced effective population size (N_e) (Wright 1931) and experience a greater fixation of alleles due to drift (Araki et al. 2007, Charlesworth 2009, McCracken et al. 2013a). The allelic diversity of headwater habitats is therefore predicted to be low but highly distinct from other headwater habitats (Hughes et al. 2009, Morrissey and de Kerckhove 2009). The unique alleles fixed in each headwater collect in downstream confluences leading to the increased genetic diversity and allelic richness of these confluences (Morrissey and de Kerckhove 2009, Paz-Vinas and Blanchet 2015). Dendritic systems therefore are unique from most other metapopulations in that genetically homogeneous subpopulations (i.e. the headwaters) are the source of the metapopulation's genetic diversity (Morrissey and de Kerckhove 2009). Unlike most metapopulations subject to asymmetric gene flow and fragmentation (which typically

have a reduced genetic diversity (Whitlock and Barton 1997, Waples 2010, Palstra and Ruzzante 2011, Gomez-Uchida et al. 2013) asymmetric gene flow and fragmentation increase genetic diversity in dendritic metapopulations due to the greater genetic differentiation between the headwaters (Morrissey and de Kerckhove 2009). Therefore, the metapopulation N_e of dendritic systems is predicted to be greater than the sum of the subpopulation N_e values, in contrast to most other fragmented metapopulations where the opposite is observed (Morrissey and de Kerckhove 2009, Gomez-Uchida et al. 2013). Although dendritic metapopulations have been widely studied using computer simulations, the predictions from these models have rarely been tested empirically (Campbell Grant et al. 2007, Labonne et al. 2008, Campbell Grant 2011, Perkin and Gido 2012, Altermatt 2013).

There are a number of other environmental factors that can affect gene flow within a river depending upon the life history and behaviour of the species in question (Fullerton et al. 2010). Isolation by distance (IBD) (Wright 1943) will be observed in those organisms whose maximal dispersal distance during their entire lifetime is less than the maximal distance between subpopulations within the system (Orsini et al. 2013b). River slope and elevation difference between subpopulations, may also limit dispersal if an organism is hindered by travelling upstream (Lowe et al. 2006, Caldera and Bolnick 2008, Hughes et al. 2009, McCracken et al. 2013a). Large physical barriers such as waterfalls or dams can be total or partial barriers to migration between segments of the river system, fracturing the metapopulation (Cote et al. 2009, Horreo et al. 2011). The positioning of these barriers within a dendritic system will also dictate their effect on an organism's dispersal (Altermatt 2013). For example, barriers located closer to the outlet of the river system will have a greater fragmenting effect on anadromous species whereas barriers located closer to the headwaters will have a greater effect on potadramous species (Cote et al. 2009). River systems can also contain a number of habitats with different environmental conditions which can lead to within-metapopulation genetic variability (Carvalho 1993). For example, headwater and downstream habitats are known to be environmentally distinct in a number of variables including depth and habitat area (Vannote et al. 1980, Schlosser 1990). If subpopulations are subject to different selection pressures due to their different environments, maladaptive immigrants may not become

incorporated into the subpopulation's gene pool; leading to isolation by adaptation (Orsini et al. 2013b).

Besides contemporary environmental factors, historical colonization patterns can also have significantly shaped the observed genetic structure of a metapopulation particularly if colonization was recent (Costello et al. 2003, Castric and Bernatchez 2003, Poissant et al. 2005). When a landscape is invaded in accordance with the stepping stone model a reduction in genetic diversity with each subsequent founding event can occur (Austerlitz et al. 1997, Le Corre and Kremer 1998). Rapid population expansion and high gene flow between recently founded subpopulations can ameliorate this effect (Nei et al. 1975, Dlugosch and Parker 2008, Greenbaum et al. 2014). After the colonization of the landscape, local environmental factors will begin to dictate gene flow and drift within the system (Costello et al. 2003). With time and unchanging landscape connectivity, metapopulations should approach migration-drift equilibrium, which results in a strong pattern of IBD (Slatkin 1993, Poissant et al. 2005). While IBD is initially limited to the geographically closest subpopulations due to their higher gene flow, it will eventually extend to include more distant subpopulations (Slatkin 1993, Hutchison and Templeton 1999). Time to achievement of migration-drift equilibrium will be lengthened by large subpopulation $\widehat{N}_{\mathrm{e}}$ values and low migration rates (Slatkin 1993, Turgeon and Bernatchez 2001). Many metapopulations have not achieved migration-drift equilibrium meaning that traces of historical patterns will remain in the observed contemporary structure of a metapopulation (Castric and Bernatchez 2003). While identifying departures from equilibrium is relatively easy, determining its cause can be difficult given that contemporary environmental factors can maintain or erode genetic patterns due to historical processes (Poissant et al. 2005). Disentangling contemporary and historical effects on observed genetic structure is therefore essential in those populations that are not in migration-drift equilibrium.

The purpose of this study is to identify the relative effects of life history traits, environmental barriers, dendritic structure, and historical colonization patterns on the genetic structure of a riverine fish within a single watershed. While each of these factors has been studied extensively they are rarely considered together despite the potential for interaction (Costello et al. 2003, Orsini et al. 2013a). This study assesses the importance

of each of these factors individually and together in a dendritic metapopulation in northern Labrador of longnose suckers (*Catostomus catostomus* (Forester 1773)) which, like most Arctic fishes, are generally understudied despite their susceptibility to climate change (Reist et al. 2006, Harris et al. 2012).

Longnose suckers are benthic invertivores (McPhail and Lindsey 1970, Ryan 1980, Scott and Crossman 1998) found in clear, deep waters including freshwater lakes, rivers, and streams (except when streams are frozen over in winter (Craig 1989)) but may also be found in brackish river mouths (McPhail and Lindsey 1970, Scott and Crossman 1998). Lake-dwelling longnose suckers rarely leave their lake except when spawning (Harris 1962, Walton 1980), which usually occurs in spring (April to June) after the temperature has risen to about 5^{o}C and the ice cover has thawed (McPhail and Lindsey 1970, Ryan 1980, Scott and Crossman 1998). Adults spawn over gravel in shallow, fast-moving water and then typically return downstream to the lake (Geen et al. 1966, Walton 1980, Scott and Crossman 1998). The eggs, which are adhesive and demersal, attach to the gravel substrate and hatch 1 - 2 weeks later (Scott and Crossman 1998). Larval fish then drift downstream until late summer (August) (Walton 1980) and take up residence in the shallow vegetation of lakes (Edwards 1983). Longnose suckers may also spawn in the outlets and shallow areas of lakes (Ryan 1980, Scott and Crossman 1998). The high dispersal potential of this fish should result in observations of elevated migration. Upstream migration of this species could potentially overcome downstream bias in gene flow which contributes to the isolation of headwaters in dendritic systems (Morrissey and de Kerckhove 2009).

Historical migration has also likely shaped the genetic structure of longnose suckers within the Kogaluk. With the withdrawal of the Laurentide Ice Sheet from the Quebec peninsula 14 000 to 5 000 years ago a series of proglacial lakes were left behind (Legendre and Legendre 1984, Bernatchez and Wilson 1998). These lakes were utilized by longnose suckers and other species to migrate east across Quebec to colonize Labrador (Legendre and Legendre 1984, Black et al. 1986, Griffiths 2010). It is thought that longnose suckers first invaded the Churchill River in southern Labrador and from there migrated overland and along the coast (though the latter is contested Dillinger et al. 1991) as far north as the Kogaluk River (Black et al. 1986). The Kogaluk River was covered by

the Laurentide Ice Sheet up until it began to thaw approximately 9 000 years BP (Bryson et al. 1969, Short and Nichols 1977). The glacial lake Naskaupi was formed to the west of the Kogaluk River between 8 400 and 7 500 years BP (Jansson and Kleman 2004) from the glacial runoff that was prevented from draining through Ungava Bay by the retreating glaciers (Ives 1960, Barnett and Peterson 1964, Jansson 2003). Multiple spillover events occurred from Lake Naskaupi into the Kogaluk River (Ives 1960, Barnett and Peterson 196, Jansson 2003) between 8 400 to 7 000 years BP before the final drainage of this glacial lake (i.e., before the final deglaciation of the Ungava peninsula approximately 6 400 years BP (Jansson and Kleman 2004)). This increased connectivity is thought to have facilitated the introduction of lake chub (*Couesius plumbeus*) into the Kogaluk (Michaud et al. 2010) and may also be a potential colonization route of longnose suckers into the system. Given the relatively recent colonization of this system this metapopulation is unlikely to have achieved migration-drift equilibrium.

The Kogaluk River in northern Labrador is an ideal system for the study of dendritic systems. The isolation of this river system means that it is subject to minimal anthropogenic influence (Anderson 1985) allowing for the characterization of natural connectivity within a dendritic system. The Kogaluk contains a number of hierarchically arranged lakes and the system is punctuated by several waterfalls and more gradual changes in elevation which may have shaped gene flow within the system. A waterfall at the easternmost extent of the drainage prevents immigration (Anderson 1985), resulting in a closed system.

Despite the relative simplicity of this system there are several factors that could potentially shape neutral genetic structure of longnose suckers within this metapopulation. I expect that longnose suckers will demonstrate high migration rates between lakes due to their high dispersal ability. However, this dispersal may be thwarted by the shallow streams connecting lakes resulting in isolated lake subpopulations that experience a high degree of drift. This is what was observed in lake trout within this watershed (McCracken et al. 2013a). Highly genetically distinct populations, a correlation between $\widehat{N}_{\mathrm{e}}$ values and lake areas, and a lack of IBD would support the same conclusion for longnose suckers. Several environmental factors may have also shaped neutral genetic structure of longnose suckers in this metapopulation. The waterfalls in this

system in particular are expected to provide significant barriers to gene flow. Distance, slope, and elevation may also limit gene flow between lakes. The dendritic structure of the system itself is anticipated to shape genetic structure by directing genetic diversity towards the confluences and genetic diversity should decrease with distance from the most downstream lake. I also expect that the dendritic structure of the metapopulation will confer upon it an elevated metapopulation N_e that exceeds the sum of the individual lake's $\widehat{N}_e$ values. However, given that this watershed was colonized only as early as 9 000 years ago (Bryson et al. 1969, Short and Nichols 1977) I predict that this metapopulation is not in migration-drift equilibrium and therefore will not demonstrate a pattern of IBD (Slatkin, 1993). If this is the case, there will be a reduction in genetic diversity from the first lake colonized (according to the most likely colonization route) to subsequently colonized lakes due to loss of genetic diversity from founder effects (Austerlitz et al. 1997, Le Corre and Kremer 1998). By determining the relative effects of life history traits, contemporary environmental barriers, the system's dendritic structure, and the historical colonization of the system on this metapopulation's genetic structure, this research will provide an understanding of the factors that are most influential in dictating connectivity in an undisturbed, subarctic river system.

CHAPTER 2: MATERIALS AND METHODS

2.01 Study Site and Sampling

The Kogaluk River (Figure 2.01.1, Table 2.01.1) comprises a number of hierarchically-arranged barren ground lakes that drain into Cabot Lake, a 2 440 ha fjord lake (Anderson 1985). The Kogaluk River runs east from Cabot and drains into Voisey's Bay and the Atlantic Ocean. A 9.2 m high waterfall located 6.4 km upstream from the river mouth at the southern end of Voisey's bay is expected to prevent fish immigration (Anderson 1985). Upstream of Cabot are four more waterfalls, all greater than 5 m in height (and all likely to be complete barriers to upstream fish migration), each marking the downstream end of a major tributary of the Kogaluk River (Anderson 1985). Barren ground lakes in the northwestern part of the system are relatively unproductive and are frozen up until July of each year; they remain ice-free for less than 150 days each year (Lopoukhine 1978). The system's remote location has resulted in minimal anthropogenic influence within the region (Anderson 1985).

Fish species that are present and abundant in the Kogaluk River include: lake trout (*Salvelinus namaycush*), round white fish (*Prosopium cylindaceum*), lake chub (*Couesius plumbeus*), Arctic char (*Salvelinus alpinus*), and the longnose sucker (McCracken et al. 2013a) the latter being the species of focus for this study.

Genetic samples were taken from a total of $n = 1315$ longnose sucker individuals collected from 11 lakes within the Kogaluk River between 2006 and 2014. Mistastin Lake, Hawk Lake and Genetics B did not yield sufficient samples (<15 per lake) for a population genetics study and these samples were removed from subsequent analysis leaving $n = 1297$ individuals. Lakes were sampled in at least two locations during each sampling event using variably-sized standardized nylon monofilament gillnets (1.27 cm to 8.89 cm diagonal) and/or electrofishing to ensure sampling of a wide variety of age classes. Fish were weighed, their fork length measured, and their sex and maturity assessed. Pectoral fin clips were extracted and either immediately stored in 95% ethanol or stored as dried samples. Opercula were extracted, boiled and stored dry.

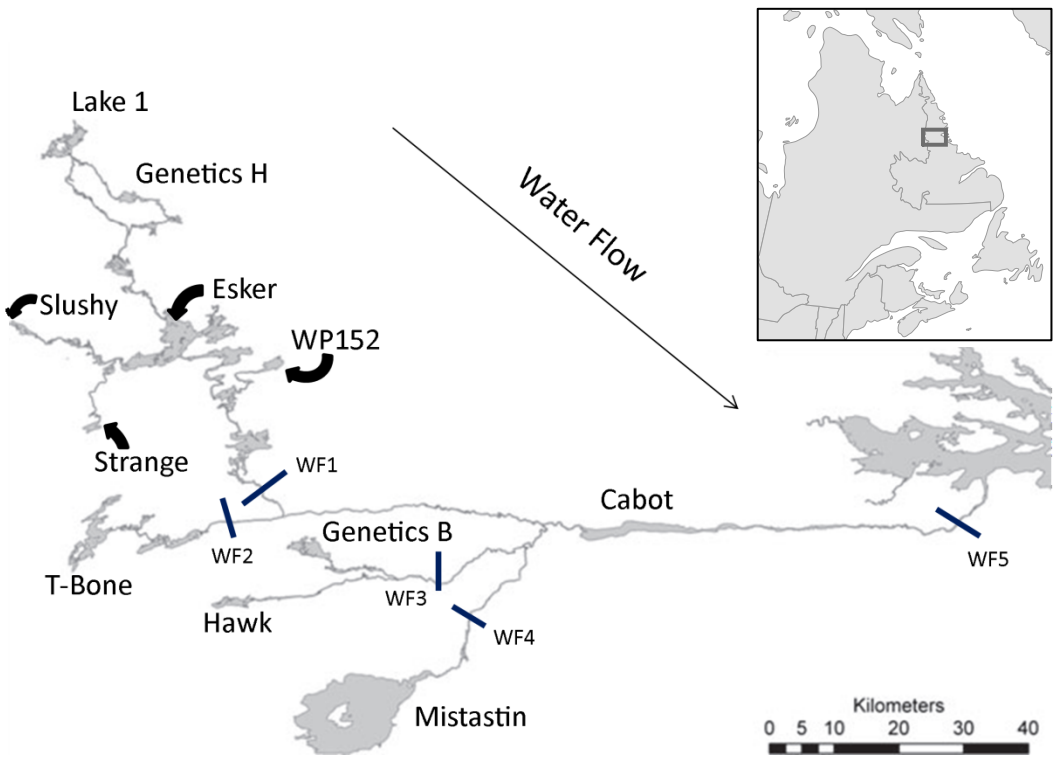

Figure 2.01.1 The Kogaluk River in northern Labrador. The map inset demonstrates the location of the system within Labrador. The Kogaluk River drains into the Atlantic Ocean by Voisey's Bay at the easternmost extent of the system. Five waterfalls (WF1-WF5) are present within the system at the following heights: WF1 is 15 m high, WF2 is 12 m, WF3 is 5.4 m, WF4 is 5.4 m, and WF5 is 9 m (Anderson 1985). Figure adapted from Fig.1 of McCracken et al. (2013a), map inset created using data courtesy of the Ministry of Natural Resource's Geogratis website (http://geogratis.cgdi.gc.ca) in ArcGIS Desktop: Release 10.1 (ESRI, 2012).

2.01.1 Environmental, life history, and genetic characteristics of eight lakes within the Kogaluk River (Labrador, Canada). MAS is metres above sea level, N/A is not available, ω is the maximum age observed in each lake, AL is the adult life span for each lake, T is the generation time, A_R is allelic richness, H_o is observed heterozygosity, H_e is the expected heterozygosity.

Lake name	Latitude, Longitude	Elevation (MAS)	Habitat area (km^2)	Depth (m)	ω	AL	Robson-Chapman $\widehat{S}$ (95% Confidence Interval)	T	Number of Genetic Samples	A_R	H_o	H_e
Lake 1	56°40′ 31.7″N, 64°00′ 07.5″W	525	11.3	3.9	32	22.3	0.704 (0.613 - 0.794)	13.3	59	8.28	0.66	0.65
Genetics H	56°36′ 13.7″N, 63°52′ 09.1″W	512	2.81	6.5	25	15.3	0.706 (0.654 - 0.759)	13.2	201	9.45	0.63	0.64
Slushy	56°24′ 56.2″N, 64°06′ 08.1″W	464	2.99	15.3	26	16.3	0.602 (0.518 - 0.685)	12.0	103	9.92	0.65	0.64
Strange	56°17′ 24.8″N, 63°56′ 53.4″W	487	2.09	N/A	22	12.3	0.716 (0.666 - 0.766)	13.1	122	8.48	0.61	0.61
Esker	56°24′ 53.4″N, 63°40′ 15.1″W	431	53.98*	N/A	51	41.3	0.910 (0.894 - 0.927)	24.2	138	11.19	0.65	0.65
WP152	56°22′ 08.7″N, 63°29′ 30.5″W	445	53.98*	16.1	36	26.3	0.795 (0.740 - 0.850)	15.6	74	11.01	0.68	0.67
T-Bone	56°09′ 09.7″N, 63°56′ 21.2″W	468	19.76	N/A	16	6.3	0.692 (0.609 - 0.775)	12.0	115	9.62	0.63	0.64
Cabot	56°08′ 27.9″N, 62°37′ 52.4″W	60	25.39	N/A	21	11.3	0.623 (0.553 - 0.693)	12.1	57	10.52	0.65	0.65

* Habitat area estimate for Esker Lake and WP152 lake combined.

2.02 Opercula Dating

Opercula were collected from a total of n = 1175 longnose suckers from 11 lakes within the Kogaluk River between 2006 and 2014. Opercula were aged following Perry and Casselman (2012) by observing them concave side up under a dissection microscope. Each operculum was placed with the articulation socket proximal to the viewer and annuli (as denoted by a translucent region immediately adjacent to an opaque region) arranged horizontally within the field of view. Annuli were counted by eye using a dissecting light microscope with bottom lighting. Perry and Casselman (2012) suggest an age correction to compensate for the number of annuli covered by the dense bone growth region. However, the dense bone growth region was found to only affect aging of those individuals aged 10 or older (Perry and Casselman, 2012). Analyses were calculated using both the uncorrected and corrected ages and results were not found to differ greatly. However, generation time (T), adult lifespan (AL), maximum age (ω), and age at 50% maturity (α) estimates were generally one to two years greater when using corrected ages in comparison to uncorrected ages. The true value for these parameters likely lies between these two extremes. For simplicity, only the uncorrected ages will be subsequently discussed, however, the results of analyses conducted using the corrected ages are presented in Appendix A.

Given the low number of opercula collected from Mistastin Lake, Hawk Lake and Genetics B (<15 per lake), these opercula samples were removed from subsequent analysis leaving n = 1158 samples. Opercula for young of the year fish were not collected due to their fragility and difficulty in removing the operculum. Fish <70 mm and caught using electrofishing were assumed to be young of the year (n = 195). Ages were converted to year cohorts.

2.03 Life History Analyses

i) Age Structure

Age at 50% maturity (α) was estimated using a binomial logistic regression (Harry et al. 2013) with age as the independent variable and maturity as the binomial

dependent variable (n = 1072) (where mature individuals were designated as 1, and immature individuals were designated as 0) in R (R Core Team, 2013). Adult lifespan (AL) was estimated at AL = $\omega - \alpha + 1$, where ω is the maximum age (Waples et al. 2014) and was made equal to the age of the oldest individual sampled in each respective lake.

ii) Generation Time Estimation

Age distributions of gillnet-caught samples were compared between years within each lake using Kolmogorov-Smirnov test with a Bonferonni corrected α-value. Samples were pooled for each lake when yearly age distributions were not significantly different. The Robson-Chapman annual survivorship estimate ($\hat{S}$) (Chapman and Robson 1960, Robson and Chapman 1961) was calculated for each lake using only the ages of those individuals caught using gillnets (n = 1050). $\hat{S}$ was used to determine l_x (probability of survival to age x) for each age class (x), assuming $l_0 = 1$, $l_1 = \hat{S}^1 \text{x} l_0$, $l_2 = \hat{S}^2 \text{x} l_0$, … $l_\omega = \hat{S}^\omega \text{x} l_0$, where ω is the maximum age observed in the respective lake.

The ages of mature females were $\log_e$-transformed and correlated with fork length using a linear model (n = 204). In order to relate age with fecundity, fork length in the derived equation was then equated to total length in the following equation relating fecundity (number of eggs) with total length (Childress et al. 2015) for longnose suckers in the Great Lakes.

[1] $$f = 0.016 \text{ x } (\text{TL})^{3.799}$$

Where f is the number of eggs produced and TL is the total length of the fish in cm. It is assumed that fecundity and its relationship with TL in the Kogaluk is the same for those in the Great Lakes. This assumption is supported by the fact that fecundity was not found to significantly correlate with latitude in 20 species of European freshwater fish (Blanck and Lamouroux 2007).

Given that fork length and total length are often linearly related (Carlander and Smith 1945), equivocating total length with fork length for the purposes of relating

fecundity with age was justified as resulting fecundities for each age class would be biased by a constant factor for each age class, which would cancel out with the calculation of generation time (see Equation 2 below).

Fecundity was estimated for each age class using the resulting equation from age class α (rounded down to the nearest age) to age class ω. The f for each age class was divided by 2 to approximate m_x the number of offspring produced by an individual of age x to account for the fact that only half of the population is female.

Generation time (T) was calculated according to Birch (1948) as:

[2] $$T = \frac{\sum x l_x m_x}{\sum l_x m_x}$$

2.04 DNA Extraction, Amplification, and Genotyping

Fin clips were digested at 55^{o}C for eight or more hours using Proteinase K (Bio Basic Inc., Markham, ON, Canada). DNA was then extracted using a Multiprobe II plus liquid handling system (Perkin Elmer, Waltham, MA, U.S.A.) according to the glassmilk protocol (Elphinstone 2003).

Twenty microsatellite loci with demonstrated polymorphisms that could be accurately and consistently scored were selected for amplification (see Appendix B for primer sequences).

Samples were amplified using PCR (see Appendix C for PCR reaction contents and Appendix D for thermocycler programs).

PCR products were diluted with formamide at a ratio between 1:1 and 1:20 depending on PCR product quality. Li-COR 4200/4300 machines (Li-COR Biosciences, Lincoln, Nebraska) were then used to visualize the PCR product. The resulting images were analyzed using SAGA Automated Microsatellite Software 3.3 (Li-COR Biosciences, Lincoln, NE, U.S.A.) to determine the genotypes of individual samples, which were checked manually for accuracy.

2.05 Genetic Quality Control Analyses

The presence of scoring errors and null alleles was assessed using MICROCHECKER 2.2.3 (van Oosterhout et al. 2004). Linkage disequilibrium and departures from Hardy-Weinberg equilibrium were assessed for each marker within each population using Arlequin 3.5.1.3 (Excoffier and Lischer 2010). Evidence of pairwise linkage between markers was calculated using 10 000 permutations. Departures from Hardy-Weinberg equilibrium were assessed for each marker using 1 000 000 Markov chain steps and 100 000 dememorization steps. *P*-values from both tests were adjusted using the false discovery rate correction (Benjamini and Hochberg 1995). Microsatellite markers that were potentially under selection were identified using LOSITAN (Antao et al. 2008) using 1 000 000 simulations of the stepwise mutation model, a subsample size of 50, and a false discovery rate of 0.05.

2.06 General Genetic Characteristic Analyses

FSTAT (Goudet 2001) was used to determine each lake's allele frequencies and allelic richness and Arlequin 3.5.1.3 (Excoffier and Lischer 2010) was used to determine each lake's observed and expected heterozygosity. Allelic richness, observed heterozygosity and expected heterozygosity were correlated with waterway distance to the most downstream lake (Cabot). Negative correlations would be expected if genetic diversity was pooling in the downstream lakes. $\hat{F}_{ST}$s were calculated in MSA 4.05 using 100 000 MCMC permutations (Dieringer and Schlötterer 2003). A principal coordinates analysis (PCoA) using $\hat{F}_{ST}$s was conducted in GenAlEx 6.501 (Peakall and Smouse 2006) to visualize genetic relationships among lakes.

2.07 Genetic Population Structure Analyses

Population structure was assessed using hierarchical analyses conducted in STRUCTURE 2.3.4 (Pritchard et al. 2000, Hubisz et al. 2009). Each analysis was run using the admixture model, 10 replications, 5 000 000 MCMC permutations and

2 000 000 burn-in steps. The STRUCTURE analysis was performed hierarchically, first with all eight lakes and subsequently with each of the clusters identified in the first step. This procedure was repeated until no more population structure was found among or within lakes. The most likely number of clusters (*K*) was determined using the Evanno method (Evanno et al. 2005) and calculated using STRUCTURE HARVESTER (Earl and vonHoldt 2012). For each analysis, the 10 replicates for the most-likely K-value were aggregated into a single file using CLUMP 1.1.2 (Jakobsson and Rosenberg 2007) and then visualized using DISTRUCT 1.1 (Rosenberg 2004).

AMOVAs were calculated in Arlequin 3.5.1.3 (Excoffier and Lischer 2010) using 50 175 permutations to determine genetic variation among and within 1) the lakes that are north of the Kogaluk River (Lake 1, Genetics H, Slushy, Strange, Esker, WP152) and those that are not (T-Bone, Cabot), 2) groups of lakes separated by waterfalls, 3) clusters identified by STRUCTURE. The lakes north of the Kogaluk River were predicted to be different from the remaining lakes based on the results of McCracken et al. (2013a) who found that the main genetic division in lake trout populations within the Kogaluk was between the northern and southern lakes. Lakes separated by waterfalls were also predicted to exhibit relatively high genetic differentiation due to the presumed barrier to gene flow posed by these waterfalls.

The genetic distances associated with each particular stream section were assessed using STREAMTREE (Kalinowski et al. 2008). This program uses a modification of the least-squares analysis used in phylogenetic tree construction to determine the genetic distances associated with a particular stream section by assuming that the sum of genetic distances attributed to each stream section between two subpopulations is equivalent to those subpopulation's pairwise $\hat{F}_{ST}$ (Kalinowski et al. 2008). Because STREAMTREE requires that there be only one possible pathway between populations, and Lake 1 and Genetics H are connected by two different stream sections (see Figure 2.01.1, a separate analysis was conducted for each possible connection between these two lakes). STREAMTREE analyses were conducted for both the full system and only the northern six lakes.

2.08 Migration Rate Analysis

Gene flow between lakes was estimated using BayesAss version 3.0 which uses a Bayesian analysis to estimate contemporary migration rates (Wilson and Rannala 2003). Gene flow between lakes was estimated for the whole system (eight lakes) and again, for only the six northern lakes (Lake 1, Genetics H, Slushy, Strange, Esker, WP152). Each run had 2 000 000 burn-in steps and 20 000 000 iterations. For all runs the mixing parameters for migration rate, allele frequencies, and inbreeding coefficients were set at 0.07, 0.15, and 0.15 respectively to ensure that acceptance rates for changes in these variables were between 20% and 60% (Rannala 2007). Results were assessed by visually checking tracer plots.

2.09 Estimating Causes of Genetic Differentiation

A series of Mantel tests each using 9 999 iterations was conducted in GenAlEx 6.501(Peakall and Smouse 2006) to test for the effects of a number of landscape factors on gene flow and genetic differentiation. $\hat{F}_{ST}$s were linearized ($\hat{F}_{ST}$ /(1- $\hat{F}_{ST}$)) since linearized $\hat{F}_{ST}$s are thought to correlate more closely with distance when testing for IBD (Rousset 1997) and were used for all Mantel tests for the sake of consistency. Linearized $\hat{F}_{ST}$s were compared to the following: waterway distances between lakes (IBD), differences in elevation between lakes to test for isolation by elevation, and differences in average waterway slope between lakes (calculated according to Stelkens et al. 2012) to determine the effect of isolation by slope. The effect of waterfalls on gene flow was tested by using a matrix of the number of waterfalls between lakes. Mantel tests were conducted for the full system of eight lakes and also for the six northern lakes with the exception of the Mantel test for waterfall effects since there are no waterfalls between any of the six northern lakes.

Given that multiple landscape factors could limit gene flow in this system, the results of each Mantel test were subject to a decomposed pairwise regression (Koizumi et al. 2006, McCracken et al. 2013a) to identify and remove potential outlier lakes which could be masking the effects of the tested landscape variable. Lakes with 95% confidence

intervals of their residuals not including 0 were considered putative outliers and subsequently removed from the Mantel test until no putative outlier lakes remained. The best model for each test was determined based on the lowest corrected Akaike's Information Coefficient (AIC_c).

To test whether genetic drift due to lake isolation was responsible for genetic differentiation among lakes, mean pairwise allelic richness was correlated with pairwise $\hat{F}_{ST}$s. A negative correlation would indicate that populations with high allelic richness were not as genetically distinct as those with low allelic richness due to the effects of drift (Raeymaekers et al. 2008). Additionally, mean pairwise allelic richness was correlated with standardized residual pairwise $\hat{F}_{ST}$s after pairwise $\hat{F}_{ST}$s had been regressed against geographic distance to determine if drift could explain underlying variation in any detected pattern of IBD for the entire system.

2.10 Identification of Migrants

Potential first-generation immigrants were identified using GENECLASS2 (Piry et al. 2004). The L_home/L_max (the likelihood of each individual's genotype based upon the genotypes of all other individuals within the sampled population divided by that individual's genome's maximum likelihood in comparison with any population (including the population where the individual was sampled)) was used to determine the residence likelihood of each individual (Paetkau et al. 2004). These likelihood values were compared with those of 1 000 simulated genotypes created using a Bayesian analysis (Rannala and Mountain 1997) and a Monte Carlo resampling method (Paetkau et al. 2004) to determine the probability of the observed likelihood. First-generation immigrants were identified in each lake using a Type-I error of 0.01 and were removed from subsequent analysis when calculating $\widehat{N}_e$ to meet the assumption of closed populations (Waples and Do 2008).

2.11 Estimating $\widehat{N}_e$ and $\widehat{N}_b$

$\widehat{N}_e$ was calculated for each lake using LDNe; a single-point estimator that assesses N_e based on linkage disequilibrium (Waples and Do 2008). Each lake's $\widehat{N}_e$was calculated with and without migrants identified in GENECLASS2. For all LDNe analyses a P_{crit} (minimum allele frequency) of 0.02 was used as advised by Waples and Do (2010) since all lakes had a sample size >25.

The N_e of the entire metapopulation was calculated using the method described by Tufto and Hindar (2003) (*meta*-$\widehat{N}_{e(T+H)}$). This technique has fewer assumptions than other techniques and is often more precise when metapopulation $N_e$is relatively large (Gomez-Uchida et al. 2013). *meta*-$\widehat{N}_{e(T+H)}$ were calculated using $\widehat{N}_e$ calculated in LDNe ($\widehat{N}_{e(LDNe)}$) that were generated with and without migrants for both the whole system and only the northern six lakes. Migration rates for each *meta*-$\widehat{N}_{e(T+H)}$ were calculated in BayesAss using the same parameters as described above.

The sums of the lake $\widehat{N}_{e(LDNe)}$ values were compared with the *meta*-$\widehat{N}_{e(T+H)}$ values to determine if dendritic structure conferred an elevated genetic diversity as predicted by Morrissey and de Kerckhove (2009).

Correlation analyses between lake $\widehat{N}_e$ and elevation as well as lake $\widehat{N}_e$ and lake area were conducted using R (R Core Team, 2013). A negative correlation between $\widehat{N}_e$ and elevation would suggest that genetic diversity is accumulating in the downstream subpopulations, indicating unidirectional dispersal (Castric et al. 2001, McCracken et al. 2013a). A positive correlation between $\widehat{N}_e$ and lake area would suggest that habitat area and therefore genetic drift dictates genetic diversity in the absence of migration (McCracken et al. 2013a).

The effective number of breeders ($\widehat{N}_b$) was estimated for each lake using LDNe. A combination of two to three age cohorts was used to estimate $\widehat{N}_b$ when a single cohort did not yield a positive $\widehat{N}_b$ (Waples et al. 2014).

To correct for age structure bias in the raw $\widehat{N}_b$ these values were adjusted based on the life history traits α and AL according to Waples et al. (2014):

[3] $$\widehat{N}_{\mathrm{b(adj2)}} = \frac{raw\ \widehat{N}_{\mathrm{b}}}{1.103-0.245 \times \log(\frac{\mathrm{AL}}{\alpha})}$$

The adjusted effective number of breeder estimates ($\widehat{N}_{\mathrm{b(adj2)}}$) were then used to estimate adjusted effective population size estimates ($\widehat{N}_{\mathrm{e(adj2)}}$):

[4] $$\widehat{N}_{\mathrm{e(adj2)}} = \frac{\widehat{N}_{\mathrm{b(adj2)}}}{0.485+0.758 \times \log(\frac{\mathrm{AL}}{\alpha})}$$

$\widehat{N}_{\mathrm{e(adj2)}}$ was compared with $\widehat{N}_{\mathrm{e(LDNe)}}$ for each lake.

2.12 Historical Colonization Assessment

DIYABC v2.0 (Cornuet et al. 2014) was used to assess the relative likelihood of three potential colonization routes taken by longnose suckers into the Kogaluk. Because DIYABC assumes no migration occurs between populations after divergence (Cornuet et al. 2008) and given the very high migration rate and high genetic similarities (as seen in STRUCTURE results) between Esker and WP152, these two lakes were combined for this analysis.

The first colonization route assumes longnose suckers invaded the Kogaluk from the south and east and is based on the prediction that longnose suckers colonized northern Labrador by migrating up the coast after an initial colonization of the Churchill River (Black et al. 1986). In this scenario Cabot was first colonized, followed by a subsequent invasion of Esker/WP152 at time te from which all other lakes were colonized at later time intervals (ta, tb, tc, td) (Figure 2.12.1). T-Bone was independently colonized from Cabot at some other time tf. (Accordingly, tf ≥ te, tf ≥ td ≥ tc ≥ tb ≥ ta, and te > td ≥ tc ≥ tb ≥ ta.)

The second colonization route was based on the STRUCTURE results, where Lake 1, Genetics H, Slushy, Strange, Esker/WP152, and T-Bone were each initially colonized at the same time th. Cabot was subsequently colonized from Esker/WP152 at

time tg in accordance with the location priors STRUCTURE results. (Accordingly, th > tg.)

A third colonization assumed that longnose suckers invaded the Kogaluk system from the south and west using the large paleolake Naskaupi which formed to the west of the Kogaluk between 8 400 and 7 500 years BP. This is the colonization route that lake chub are thought to have taken into the Kogaluk (Michaud et al. 2010). This scenario included an initial colonization of T-Bone, Strange, Slushy at time tm. These lakes were all either covered by lake Naskaupi or immediately adjacent to this lake between 8400 and 7 000 years BP (Jansson and Kleman 2004). An admixture of migrants from Strange and Slushy invaded Esker/WP152 at time tl from which the northernmost lakes, Genetics H and Lake 1 were subsequently invaded (at times tk and tj, respectively). Cabot was the result of an admixture event from the upstream Esker/WP152 and T-Bone at time ti. (Accordingly, tm > tl > tk ≥ tj ≥ ti.)

All time periods were assumed to take place after the retreat of the Laurentide Ice Sheet from the Kogaluk around 9 000 years BP (Bryson et al. 1969, Short and Nichols 1977). Given that the calculated generation time for most lakes in this system varied between 12 and 13.5 years (with the exception of Esker and WP152 which had very high generation times see Table 2.01.1 and section 3.01 of results and section 4.01 of discussion) all time periods were modelled to occur between 10 and 750 generations ago. In total, 3×10^6 simulations were run for all three models. A Stepwise Mutation Model was assumed and mean mutation rate was assigned a uniform distribution and was allowed to range between 1×10^{-4} and 1×10^{-3}, while individual mutation rate was allowed to range between 1×10^{-5} and 1×10^{-2}. N_e values were allowed to be variable among lakes, to range uniformly between 1×10^1 and 5×10^4, and were assumed to be equal through time. The one-sample summary statistics employed for generation of simulated datasets included mean number of alleles, mean genic diversity, and mean size variance. Two-sample summary statistics included mean genic diversity, mean size variance, and $(d\mu)^2$ distance.

A Principal Components Analysis (PCA) was used to pre-evaluate the similarity between the distribution of the datasets generated using each of the three scenarios with the observed dataset. The relative posterior probabilities of all three scenarios were

assessed with the logistic regression method using a subset of 1% of the closest simulated data. Linear regression was used to determine the distribution of the posterior parameters from the most-likely scenario after taking a logit-transformation of the parameters, again using a subset of 1% of the closest simulated data sets. Bias and precision were estimated for each scenario using 500 pseudo-observed test data sets simulated using the original parameters from the 1% subset of the closest simulated data sets. Type I and type II error rates were generated for each scenario using confidence estimates derived from 500 pseudo-observed test data sets simulated using each scenario and the original parameters. Type I error rate was calculated as the proportion of data sets where a scenario other than the focal scenario (the one used to generate the data set) was selected as best. Type II error rate was calculated as the proportion of data sets where the focal scenario was falsely chosen as the best scenario (averaged over both alternative scenarios). Model checking was completed for each scenario using two summary statistics not used in the initial data set generation as encouraged by Cornuet et al. (2010). The summary statistics used for model checking were the two-sample mean number of alleles and $\hat{F}_{ST}$.

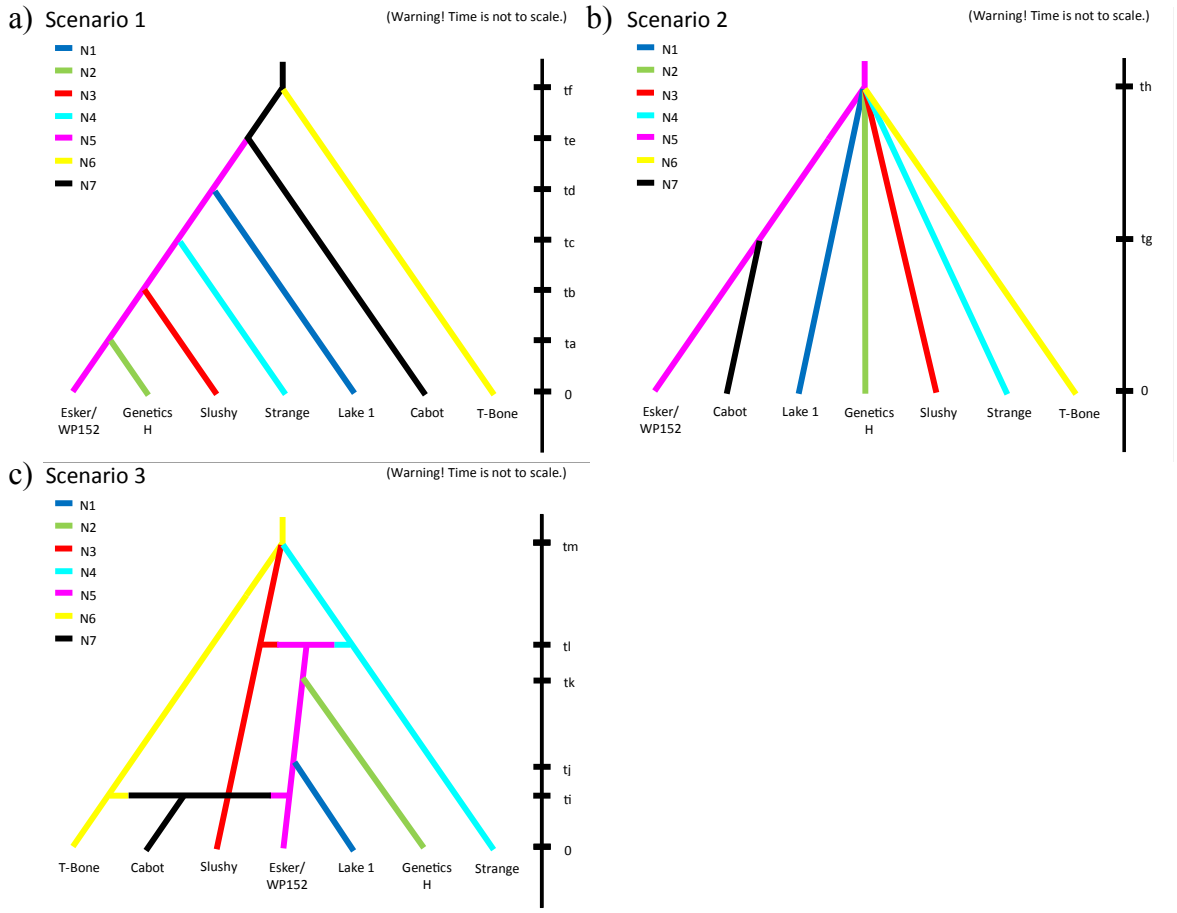

Figure 2.12.1 Three scenarios outlining colonization of longnose suckers (*Catostomus catostomus*) into the Kogaluk River that were assessed using DIYABC. t stands for a time value in generations. a) Scenario 1, where Cabot was colonized first followed by the colonization of Esker from which all other lakes in the northern arm of the Kogaluk were colonized. T-Bone was colonized from Cabot at some independent time. ($tf \geq te$, $tf \geq td \geq tc \geq tb \geq ta$, and $te > td \geq tc \geq tb \geq ta$.) b) Scenario 2 is based on the STRUCTURE results where all lakes were colonized at the same time with the exception of Cabot which was colonized from Esker/WP152 at some more recent time. ($th > tg$.) c) Scenario 3, where colonization occurred first in the westernmost lakes (T-Bone, Slushy and Strange) from lake Naskaupi. Esker/WP152 was colonized from an admixture event of migrants from Slushy and Strange. Genetics H and Lake 1 were colonized from Esker subsequently. Cabot was colonized as an admixture event of migrants from T-Bone and Esker/WP152. ($tm > tl > tk \geq tj \geq ti$.)

CHAPTER 3: RESULTS

3.01 Life History

i) Age Structure

Ages ranged from young of the year (0) to 51 years old over all individuals that were aged (n = 1353) (Figure 3.01.1). Age at 50% maturity (α) was 10.7 (Fig 3.01.2). AL ranged from 6.3 in T-Bone to 41.3 in Esker (Table 2.01.1).

ii) Generation Time

The natural logarithm transformation of the ages of mature females was significantly correlated with fork length ($P < 0.001$, adjusted R^2 value of 0.62, Figure 3.01.3) as follows:

[5] $FL = 13.94 \times \log_e(Age) + 0.98)$

Where FL is the fork length in cm.

FL in equation [5] was equivocated with total length (TL) in equation [1]:

[6] $f = 0.016 \times [13.94 \times (\log_e(Age) + 0.98)^{3.799}]$

Where f is the number of eggs produced. Equation [6] was used to generate f for each age class.

There was no significant difference in age distribution within lakes between collection years based on Kolmogorov-Smirnov tests (all $P > 0.05$, Appendix E). Annual samples within lakes were thus pooled. Robson-Chapman estimates of annual survival varied from 0.602 in Slushy to 0.910 in Esker (Table 2.01.1). Generation time varied from 12.0 in T-Bone and Slushy to 24.2 in Esker (Table 2.01.1).

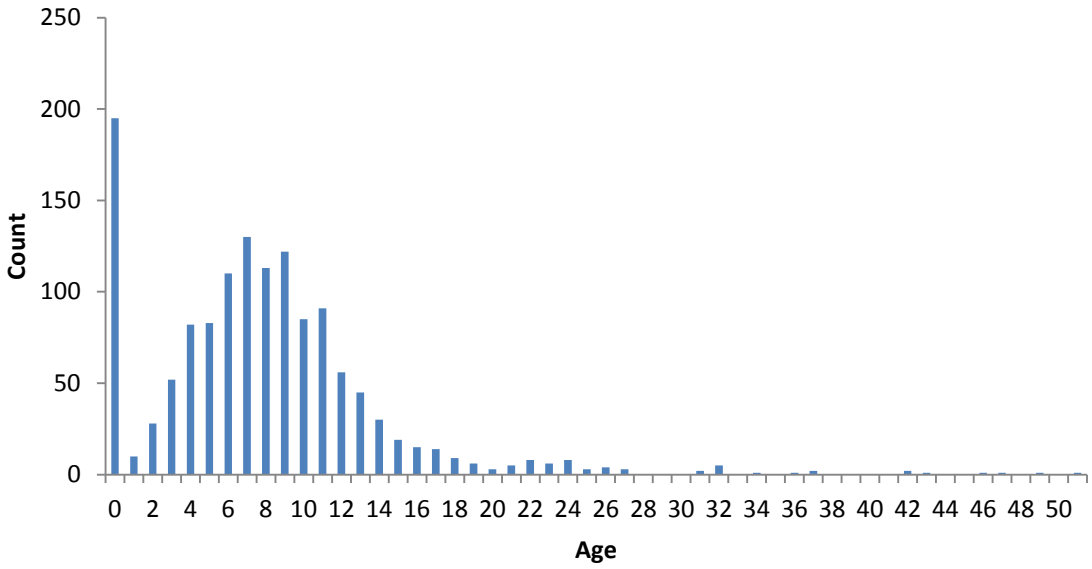

Figure 3.01.1 Number of longnose suckers (*Catostomus catostomus*) sampled for each age class.

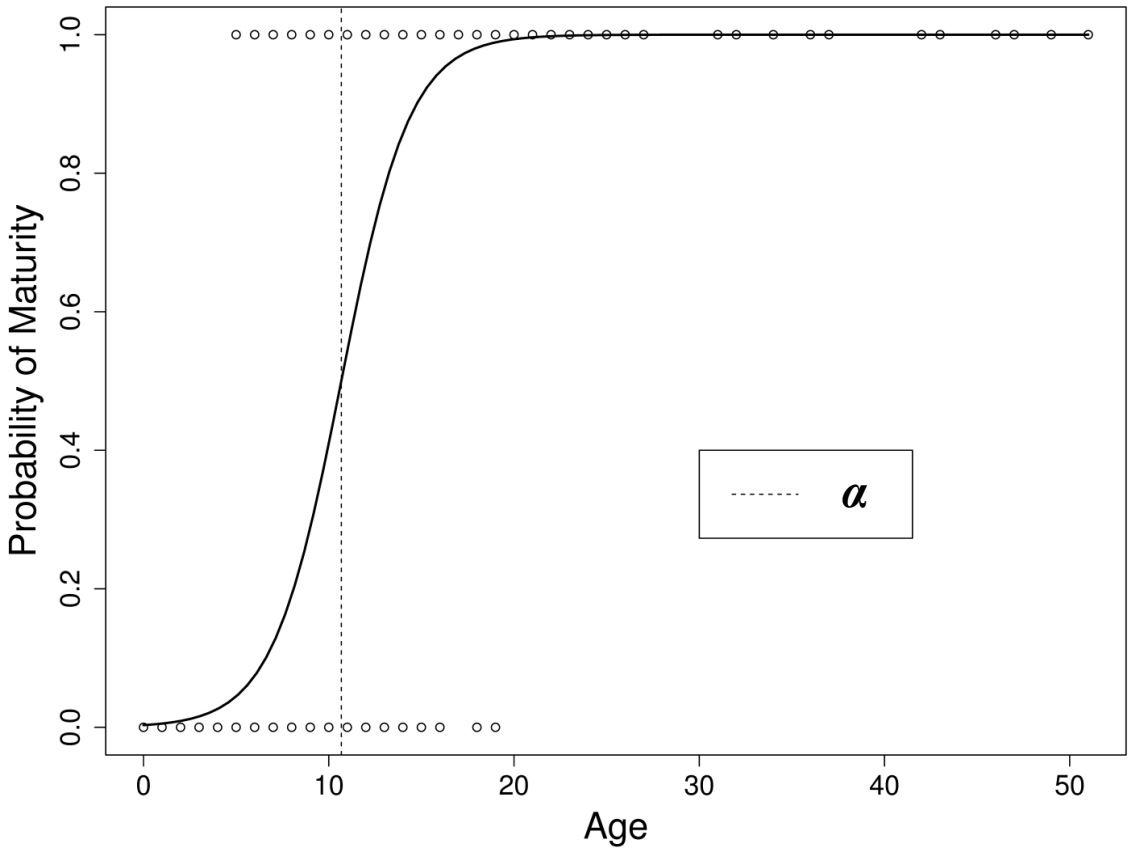

Figure 3.01.2 The age at 50% maturity (α) for longnose suckers (*Catostomus catostomus*) from the Kogaluk River estimated using a binomial logistic regression of n = 1072 samples of age versus maturity.

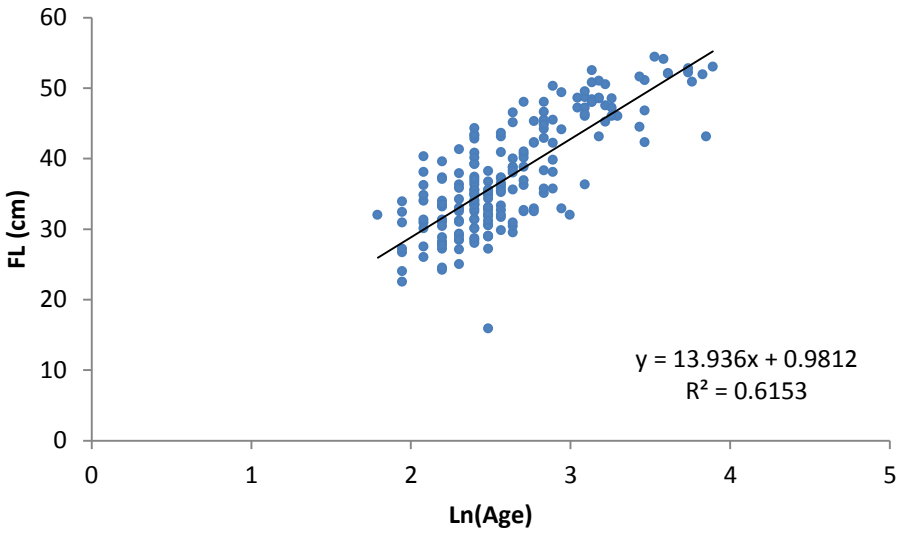

Figure 3.01.3 Correlation between fork length in cm (FL) with the natural logarithm transformation of ages of mature, female longnose suckers (*Catostomus catostomus*) caught in the Kogaluk River.

3.02 Genetic Quality Control

A total of n = 869 individuals were successfully amplified for at least 13 out of 20 markers (Table 2.01.1). Two of these markers (DLU4259 and MOHU268) exhibited low amplification success (>10% missing over all individuals) and a third marker (DLU4183) exhibited evidence of potential null alleles in six of eight lakes (see Appendix F). All three markers were excluded from further analysis. Two other markers (DLU439 and CCAT16), exhibited evidence of null alleles in one lake each; these loci were retained given the lack of consistency of null alleles across lakes in these loci, leaving a total of 17 markers for the study. Overall, 248 (28.5% of 869), 267 (30.7%), 144 (16.6%) and 84 (9.7%) individuals had scores for all 17, 16, 15 and 14 markers, respectively. A further 57 (6.6%), 32 (3.7%), 20 (2.3%) and 17 (1.96%) individuals were missing 4, 5, 6 and 7 markers, respectively. Overall therefore, 98.0% of the individuals had scores for at least 13 of 17 loci and fewer than 2% of the individuals were missing scores for 7 markers. No locus exhibited >10% missing data over all lakes, and missing data per lake over all loci was <10%.

Two loci (CCAT43 and CCAT16) exhibited evidence of departure from Hardy Weinberg equilibrium in one lake (Strange) after controlling for false discovery rate (Benjamini and Hochberg 1995) but this was only slightly greater than the 0.8 significant comparisons expected within a population given an α-value of 0.05 and 17 comparisons. There was also evidence of linkage disequilibrium between DLU439 and CCAT32 in Esker which was fewer than the 6.8 expected number of significant comparisons within a population given an α-value of 0.05 and 136 comparisons. No other evidence of departures from Hardy Weinberg equilibrium or linkage disequilibrium was observed.

Two markers (US3 and CCAT16) were identified as putatively under positive selection by LOSITAN (Fig 3.02.1). However given the high type I error rate associated with this outlier detection program (Narum and Hess 2011), and the fact that microsatellites are generally considered to be neutral markers (Jarne and Lagoda 1996), these two markers were included in subsequent analysis.

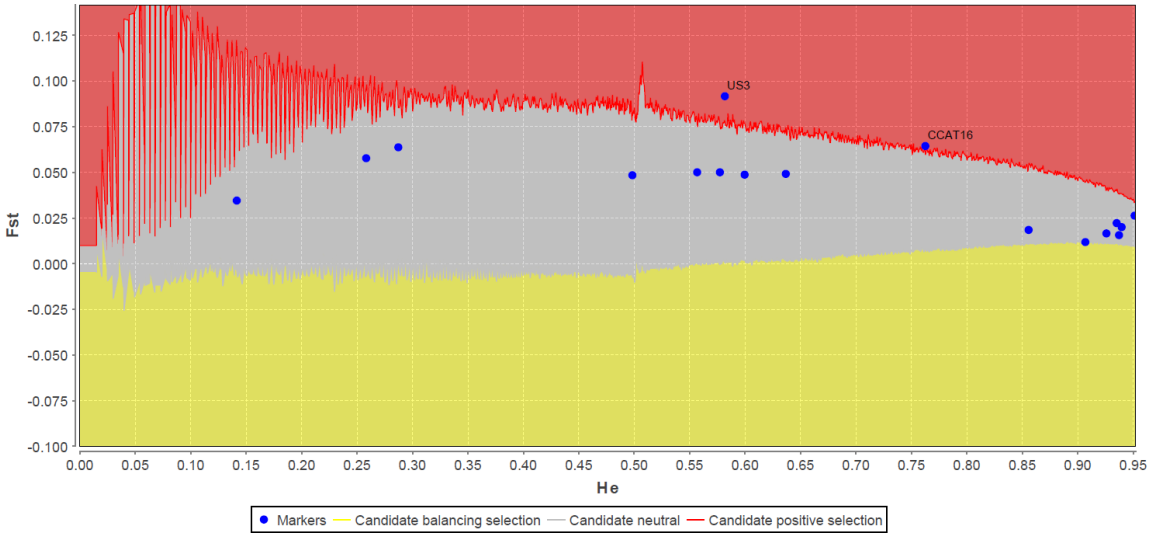

Figure 3.02.1 LOSITAN output relating $\hat{F}_{ST}$ with heterozygosity for 17 neutral microsatellite markers.

3.03 Genetic Characteristics

Allelic richness ranged from 8.28 in Lake 1 to 11.19 in Esker (Table 2.01.1) and correlated significantly with distance to the most downstream lake (Cabot) in the northern six lakes ($R^2 = 0.77$, $P \leq 0.021$) but not when considering all eight lakes ($R^2 = 0.35$, $P \geq 0.121$) (see Appendix G). Observed heterozygosity ranged from 0.61 in Strange to 0.68 in WP152 while expected heterozygosity ranged from 0.61 in Strange to 0.67 in WP152 (Table 2.01.1) and neither correlated with distance to the most downstream lake in either the full system or in the northern six lakes (all $P > 0.05$) (see Appendix G). $\hat{F}_{ST}$s ranged from 0.004 between Esker and WP152 and 0.070 between Lake 1 and Strange (see Appendix H, Table H1) and are visualized in a PCoA biplot (Figure 3.03.1). Linearized $\hat{F}_{ST}s$ ranged from 0.004 between Esker and WP152 and 0.076 between Lake 1 and Strange (see Appendix H, Table H2).

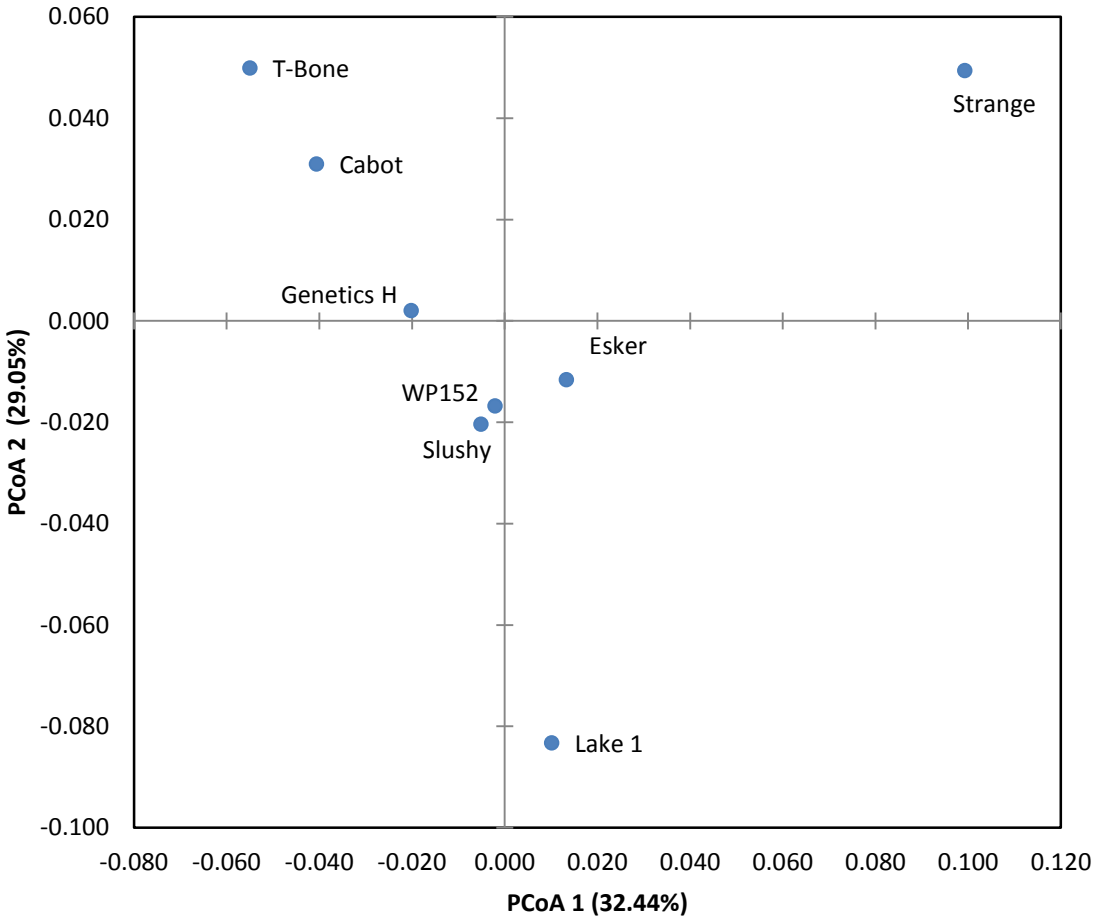

Figure 3.03.1 Principal coordinates analysis based on pairwise linearized $\hat{F}_{ST}$s between longnose suckers (*Catostomus catostomus*) samples from eight lakes within the Kogaluk River.

3.04 Genetic Population Structure

Both the Evanno et al. (2005) and Pritchard et al. (2000) methods suggested $K = 6$ for the initial STRUCTURE analysis (see Appendix I, Figure I1 for Evanno and ΔK plots). All lakes formed genetically distinguishable clusters with the exception of Esker, WP152 and Cabot which formed the sixth cluster (Figure 3.04.1 a). Subsequent analysis of this last cluster revealed no further genetic differentiation among these lakes (Figure 3.04.1 b) unless the analysis was conducted with the use of location priors. The use of location priors is justified in this instance because the great waterway distance and intermediate waterfall separating Cabot from Esker and WP152 suggests that these three lakes are not likely to form a panmictic population. With the use of location priors for only these three lakes $K = 2$, with Cabot lake being genetically distinguishable from both Esker and WP152 (Figure 3.04.1 c see Appendix I, Figure I2 for Evanno and ΔK plots)). Esker and WP152 were found to be genetically indistinguishable even when using location priors (Figure 3.04.1 d).

A hierarchical AMOVA that grouped the northern lakes (Lake 1, Genetics H, Slushy, Strange, Esker, WP152) separately from the southern lakes (Cabot, T-Bone) found 1.06% of genetic variation was explained among groups and 3.07% was explained among populations within groups, meaning that longnose suckers inhabiting the northern lakes were on average 4.14% genetically different from those inhabiting the southern lakes (Table 3.04.1). When lakes were grouped according to waterfall separation 0.71% of genetic variation was explained among groups and 3.20% was explained among populations within groups, for a total of 3.91% average genetic distinction between lakes in groupings separated by waterfalls. An AMOVA that grouped lakes according to the initial STRUCTURE resulted in $K = 6$, explained 2.17% of the total genetic variation with a further 1.47 % explained among populations within groups. When grouping lakes according to the STRUCTURE results using location priors (where all lakes were in a different group with the exception of Esker and WP152), 3.22% of genetic variation was explained among groups and 0.36% was explained among populations within groups. The average total genetic distance between lakes in these groupings was 3.59%. All AMOVA tests were significant ($P < 0.05$) with the exception of the within-population

genetic variation calculated when comparing lake groupings delineated by waterfall presence ($P \geq 0.100$).

Genetic distances associated with stream sections varied from 0 to 0.0409 over all STREAMTREE models. The highest genetic distances over all models were associated with the stream sections between Strange and Esker and between Lake 1 and Esker. In general the stream section between Esker and WP152 and the stream section between Cabot and the node connecting this lake with the rest of the system were the sections with the lowest associated genetic distances. The stream section between Esker and the node located at the point where Lake 1 and Genetics H connect (in the models where Lake 1 and Genetics H drain into a common downstream point) and the stream section between Esker and Genetics H (in the models where Lake 1 flows directly into Genetics H) also had low genetic distances associated with them. The first model (where Lake 1 and Genetics H were connected at a fork downstream from both lakes) had more support ($R^2 = 0.948$ for the whole system (Figure 3.04.2a), $R^2 = 0.969$ for only the northern six lakes (Figure 3.04.2b)) than the second model (where Lake 1 drained directly into Genetics H ($R^2 = 0.879$ for the whole system (Figure 3.04.2c), $R^2 = 0.810$ for only the northern six lakes (Figure 3.04.2d)). Therefore the distance between Lake 1 and Genetics H was calculated using this stream section (as measured using Google Maps (2015, Mountain View, California)) for the IBD Mantel tests (see Section 3.06). Genetic distances associated with stream sections did not differ considerably between the models containing all lakes and the models containing only the northern six lakes.

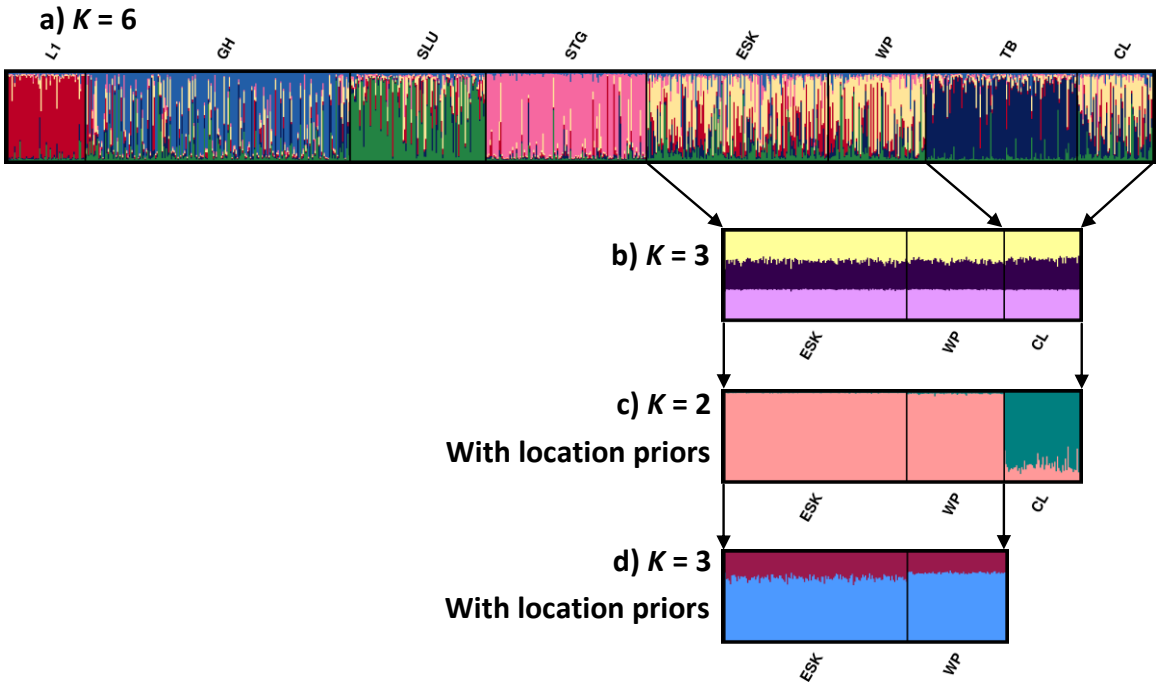

Figure 3.04.1 Hierarchical STRUCTURE plot based on the genotypes of 17 loci for longnose suckers (*Catostomus catostomus*) collected from the Kogaluk River. Most likely *K* (number of clusters) value was determined using the Evanno method (Evanno et al. 2005). a) The initial analysis differentiated Lake 1 (L1), Genetics H (GH), Slushy (SLU), Strange Lake (STG), and T-Bone (TB) as genetically distinct clusters with Esker (ESK), WP152 (WP), and Cabot Lake (CL) forming the final cluster. b) STRUCTURE analysis of only ESK, WP and CL revealed no further substructure. c) Analyzing these three lakes using location priors revealed that the genetic distinctiveness of CL from ESK and WP at $K = 2$ (the most likely *K*-value based on the Evanno method). d) Analyzing only ESK and WP with location priors revealed no further substructure.

Table 3.04.1 Summary of AMOVA results comparing: a) northern lakes (Lake 1 (L1), Genetics H (GH), Slushy (SLU), Strange (STG), Esker (ESK), WP152 (WP)) with southern lakes (T-Bone (TB), Cabot (CL)), b) northern lakes with Cabot and with T-Bone c) a single grouping of Esker, WP152 and Cabot with each other lake (the groupings identified from STRUCTURE K = 6), d) a single grouping of Esker and WP152 with each other lake. DF is degrees of freedom.

Source of variation	DF	Sum of Squares	Variance components	Percentage variation	*P*
a) Northern Lakes, South Lakes					
Among groups	1	70.86	0.06056	1.0642	<0.001
Among populations within groups	6	239.868	0.17483	3.07213	<0.001
Within populations	1604	8711.448	5.45547	95.86367	0.003
Total	1611	9022.176	5.69086		
b) Northern Lakes, CL, TB					
Among groups	2	94.152	0.04015	0.70713	<0.001
Among populations within groups	5	216.576	0.18166	3.19971	<0.001
Within populations	1604	8711.448	5.45547	96.09316	0.100
Total	1611	9022.176	5.67727		
c) L1, GH, SLU, STG, TB, ESK and WP and CL					
Among groups	5	274.088	0.12264	2.16635	<0.001
Among populations within groups	2	36.64	0.0832	1.46963	<0.001
Within populations	1604	8711.448	5.45547	96.36402	<0.001
Total	1611	9022.176	5.66132		
d) L1, GH, SLU, STG, TB, CL, ESK and WP					
Among groups	6	301.564	0.18266	3.22792	<0.001
Among populations within groups	1	9.164	0.02051	0.36246	<0.001
Within populations	1604	8711.448	5.45547	96.40961	<0.001
Total	1611	9022.176	5.65864		

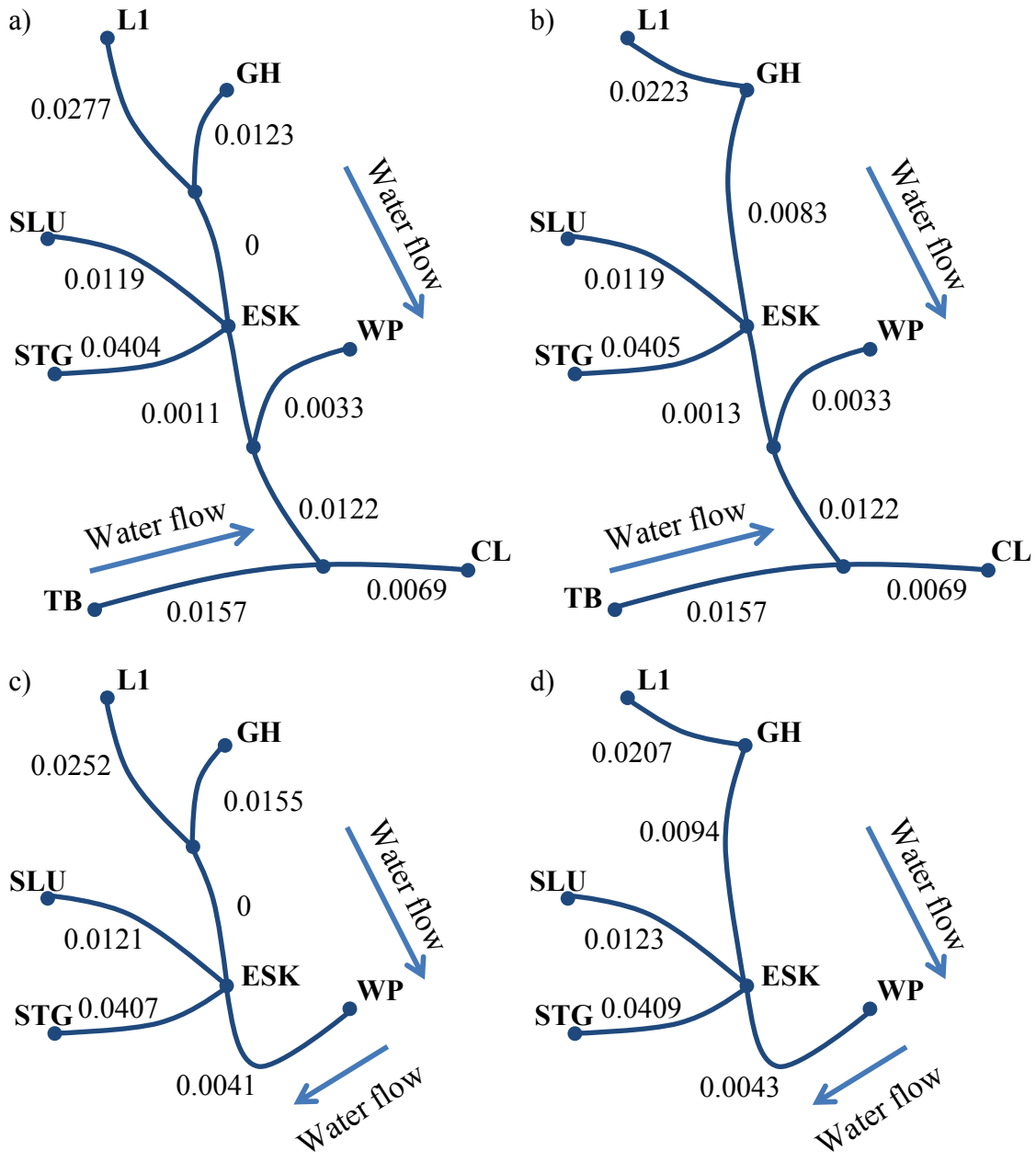

Figure 3.04.2 Longnose sucker (*Catostomus catostomus*) genetic distances associated with stream sections calculated using STREAMTREE between eight lakes in the Kogaluk River: Lake 1 (L1), Genetics H (GH), Slushy (SLU), Strange (STG), Esker (ESK), WP152 (WP), T-Bone (TB), and Cabot (CL). Since Lake 1 and Genetics H are potentially connected by two alternative stream routes L1 and GH were modeled as being connected at a common downstream fork for a) the entire system ($R^2 = 0.948$) and c) only the northern six lakes ($R^2 = 0.969$). Alternatively, Lake 1 was modelled as draining directly into Genetics H for b) the entire system ($R^2 = 0.879$) and d) the northern six lakes ($R^2 = 0.810$).

3.05 Migration Rates

Significant levels of gene flow were detected between some lakes: ranging from 2% per generation from Esker to Genetics H to 30% from Esker to WP152 (Figure 3.05.1). Most movement in the system was emigration from Esker including emigration to Slushy (8%), Strange (3%), Genetics H (2%), WP152 (30%), and Cabot (24%). There was also evidence of migration from Genetics H to Esker (3%). No other lake pairings demonstrated significant migration.

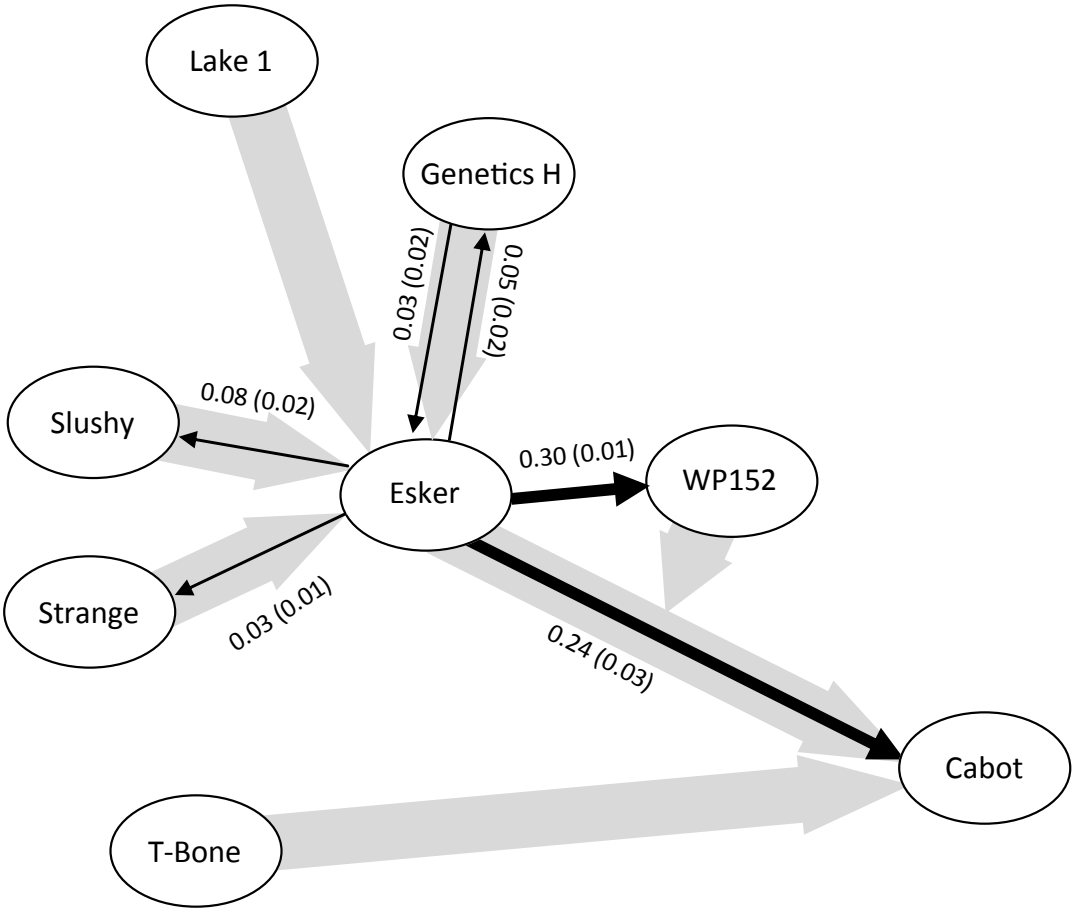

Figure 3.05.1 Black arrows indicate significant migration rates (as proportion of individuals per generation) between lakes based on 95% confidence intervals (Rannala 2007) as calculated in BayesAss ver. 3.0 (Wilson and Rannala 2003). Thick arrows indicate migration rates greater than 10% per generation; thin arrows indicate migration rates less than 10% per generation. Standard deviations are in parentheses. Light grey arrows indicate direction of water flow.

3.06 Causes of Genetic Differentiation

A Mantel test of pairwise genetic and waterway distances initially showed no support for IBD within the whole system ($R^2 = 0.13$, $P \geq 0.122$, Table 3.06.1, Figure 3.06.1 a, see Appendix J, Table J1 for distance matrix). However, there was evidence of a correlation between genetic and waterway distances after the removal of Strange ($R^2 = 0.40$, $P \leq 0.007$, Figure 3.06.1 c) and subsequently Lake 1 ($R^2 = 0.77$, $P \leq 0.003$, Figure 3.06.1 e), which were each identified as outlier lakes through decomposed pairwise regression. The removal of both Strange and Lake 1 also resulted in the lowest AIC_c. Similarly there was initially no evidence of IBD within the northern six lakes ($R^2 = 0.02$, $P \geq 0.397$), even after the removal of Strange which was identified as an outlier lake ($R^2 = 0.02$, $P \geq 0.394$). However when Lake 1 was removed from analysis there was significant evidence of a correlation ($R^2 = 0.58$, $P \leq 0.043$). Lake 1 was not identified as an outlier based on decomposed regression but its removal is justified given its genetic distinctiveness as revealed in the PCoA (Figure 3.03.1) and its removal resulted in the lowest AIC_c. (See Appendix K for all IBD Mantel tests and outlier lake detection among the northern six lakes).

Neither elevation nor slope had an effect on genetic distance regardless of whether the entire system or only the northern six lakes were considered; even after the removal of outlier lakes (all P values > 0.05). (See Appendix J, for slope and elevation matrices and see Appendix L and Appendix M for all tests and outlier lake detection for isolation by elevation and isolation by slope, respectively.)

The Mantel test correlating the number of waterfalls with genetic distance was significant when Strange and Lake 1 were removed and this model had the lowest AIC_c ($R^2 = 0.70$, $P \leq 0.027$, Figure 3.06.2 e). (See Appendix N for all tests.)

Given the collinearity between the number of waterfalls between lakes and the pairwise waterway distances when Lake 1 and Strange lake were removed ($R^2 = 0.675$, $P = 0.033$), two partial Mantel tests were conducted to see which of these variables correlated most with genetic distance. While waterway distance correlated with genetic distance after controlling for the number of waterfalls ($R^2 = 0.60$, $P \leq 0.032$), the

correlation between the number of waterfalls and genetic distance was high but not significant ($R^2 = 0.432$, $P \leq 0.067$) after controlling for waterway distance.

Mean pairwise allelic richness was significantly negatively correlated with both pairwise $\hat{F}_{ST}$s and the standardized residual $\hat{F}_{ST}$s after $\hat{F}_{ST}$s were correlated with pairwise geographic distances when all lakes were included in the analysis ($R^2 = 0.69$, $P \leq 0.001$; $R^2 = 0.69$, $P \leq 0.001$ respectively). However, the correlation between mean pairwise allelic richness and standardized residual $\hat{F}_{ST}$s became insignificant when Lake 1 and Strange were removed from the analysis ($R^2 = 0.02$, $P \geq 0.636$) see Appendix O, Figure O1 and Figure O2). The correlation between mean pairwise allelic richness and $\hat{F}_{ST}$s remained significant with the removal of Strange and Lake 1 ($R^2 = 0.39$, $P \leq 0.01$) but became insignificant with the removal of one visually-identified outlier point, that for Esker and WP152 ($R^2 = 0.20$, $P \geq 0.109$, see Appendix O, Figure O2c). Similar results were observed within the northern six lakes (see Appendix O, Figure O3 and Figure O4).

Table 3.06.1 Results of Mantel tests between pairwise genetic distances ($\hat{F}_{ST}$) (G) and pairwise elevation differences (E), slopes (S), and number of waterfalls (W) between lakes when considering all lakes and only the northern six lakes. Potential outlier lakes were identified and sequentially removed using a decomposed pairwise regression (Koizumi et al. 2006). Lake exclusions were verified using AIC_c and the best model (with the lowest AIC_c) is shown for each comparison.

	All Lakes			Northern Six Lakes		
Comparison	**Excluded Lakes**	R^2	*P*	**Excluded Lakes**	R^2	*P*
D vs. **G**	Strange, Lake 1	0.77	0.003	Strange, Lake 1	0.58	0.043
E vs. **G**	Strange	<0.01	0.308	Strange	0.01	0.311
S vs. **G**	Strange	<0.01	0.437	Strange	<0.01	0.558
W vs. **G**	Strange, Lake 1	0.70	0.027			
D vs. **G|W**	Strange, Lake 1	0.60	0.032			
W vs. **G|D**	Strange, Lake 1	0.43	0.067			

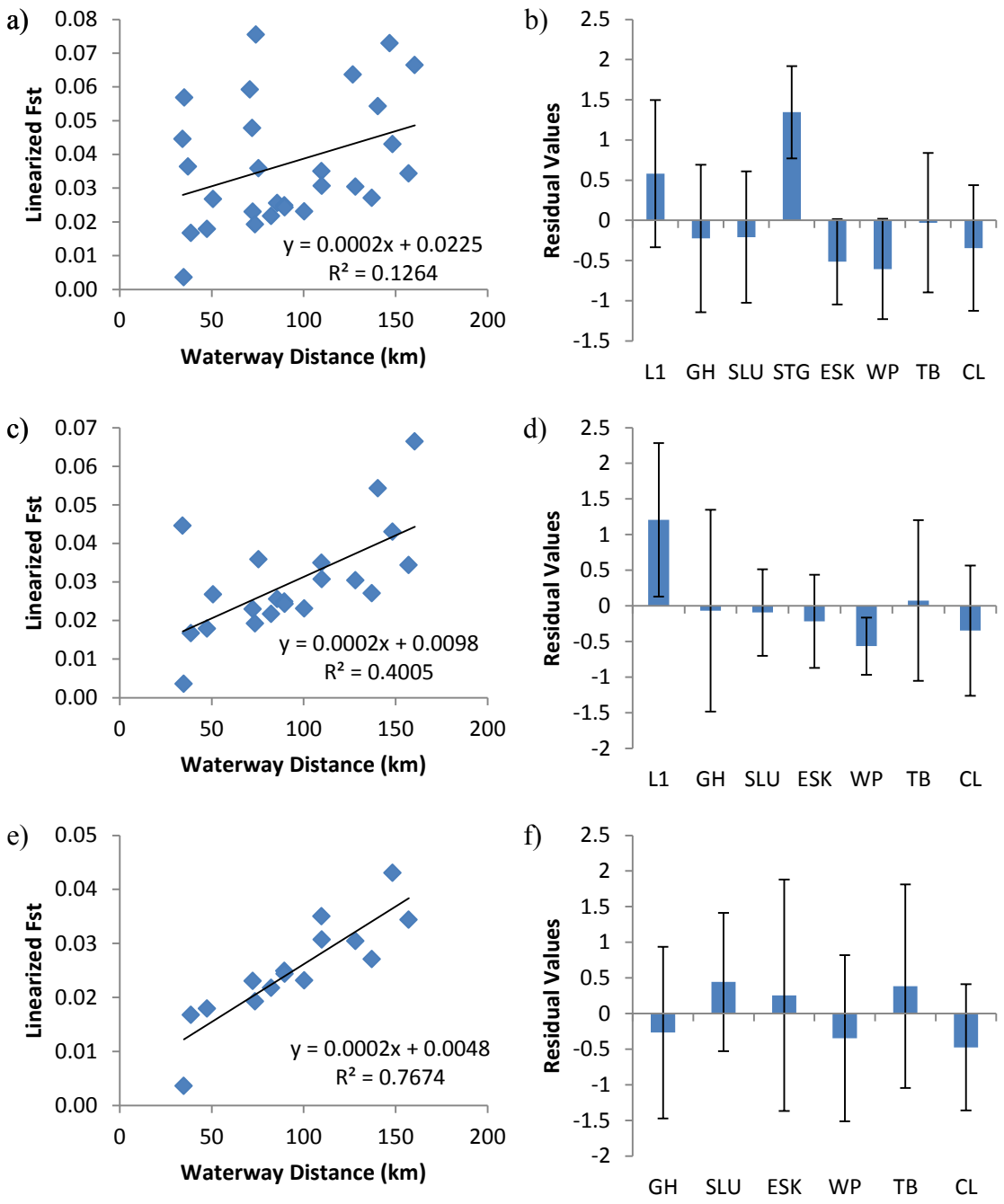

Figure 3.06.1 Correlation between pairwise linearized $\hat{F}_{ST}$ values and waterway distance (km) between samples of longnose suckers (*Catostomus catostomus*) collected from eight lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) with Lake 1 then identified as an outlier lake in d) a barplot of the residuals. e) The correlation was repeated with Lake 1 removed and f) a plot of the residuals revealed no further outlier lakes.

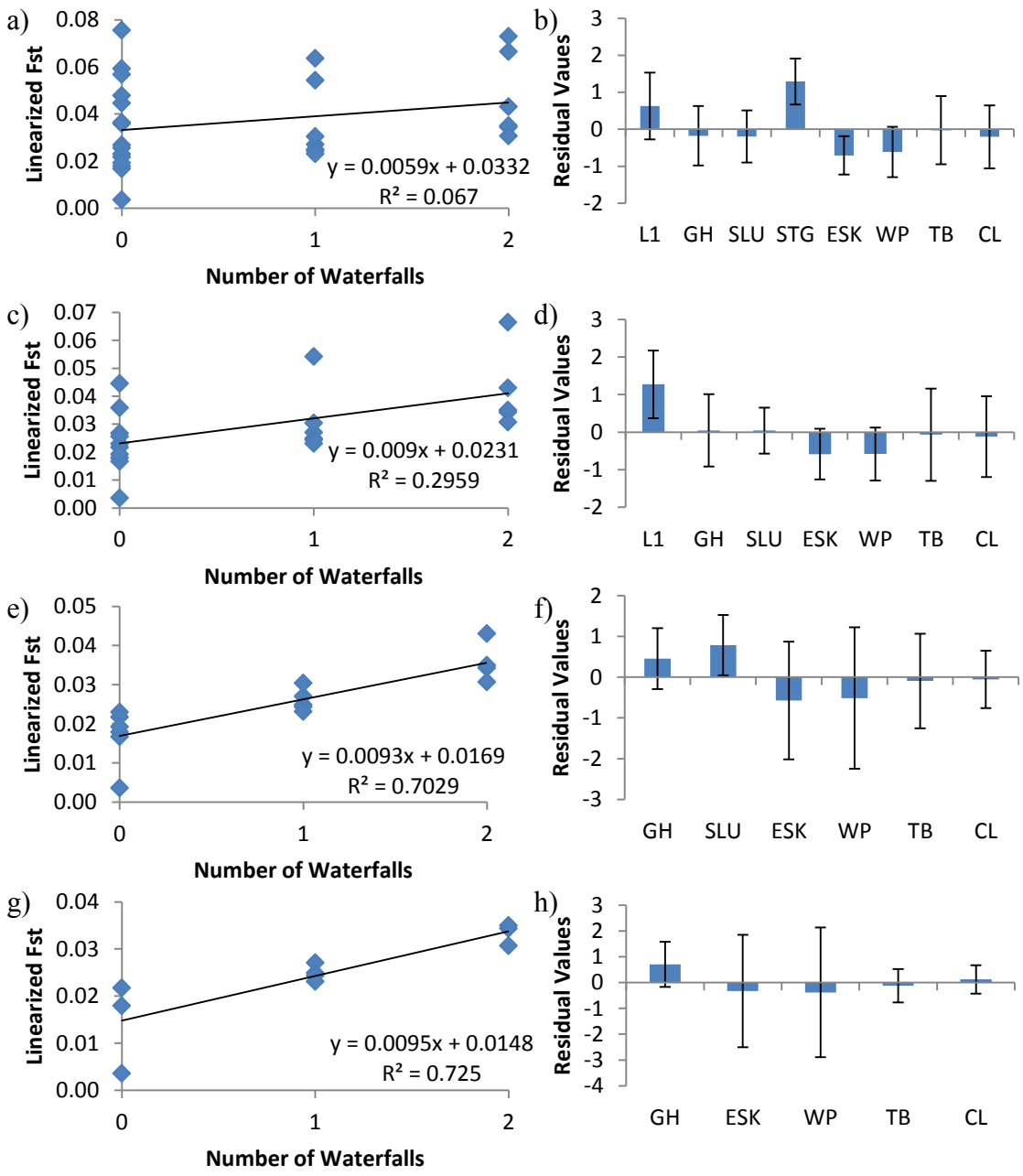

Figure 3.06.2 Correlation between pairwise linearized $\hat{F}_{ST}$ values and the number of intermediate waterfalls between samples of longnose suckers (*Catostomus catostomus*) collected from eight lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) with Lake 1 then identified as an outlier lake in d) a plot of the residuals. e) The correlation was repeated with Lake 1 removed and f) a plot of the residuals where Slushy was identified as an outlier lake. g) The correlation was repeated with Slushy removed and h) a plot of the residuals revealed no further outlier lakes.

3.07 Identified Migrants

A total of 23 individuals were identified as potential migrants using GENECLASS2: four in Lake 1, two in Genetics H, five in Slushy, five in Strange, three in WP152, one in T-Bone, and three in Cabot (see Appendix P for putative origins of migrants). However, the putative origins for some of the migrants were not concordant with the structure of the landscape. For example the migrant in Genetics H was not likely from its identified putative origin of T-Bone due to the two waterfalls and substantial geographic distance between these lakes. In these cases, I suspect migrants are likely from a lake genetically similar to the putative origin since the greatest likelihood for the lake of origin in most migrants was close to that of the second most likely lake of origin. The removal of these migrants for $\widehat{N}_{\mathrm{e}}$ estimation is justified since they are still sufficiently different from the lake they were captured in to be considered migrants despite uncertainty as to their lake of origin.

3.08 Effective Size

$\widehat{N}_{\mathrm{e}}$ was estimated for each lake with and without the individuals identified as potential migrants. Esker and WP152 samples were combined for the purpose of estimating $\widehat{N}_{\mathrm{e}}$ given their observed genetic similarity. Migrants sampled in WP152 but with a putative origin of Esker, as determined using GENECLASS2, and vice-versa, were treated as residents of this combined population and not removed when excluding migrants.

$\widehat{N}_{\mathrm{e}}$ values estimated using all samples were negative for Esker and WP152 when each were considered individually, regardless of whether or not migrants were included, indicating little linkage disequilibrium was present, and suggesting large N_{e} values in these lakes (Macbeth et al. 2013)(Table 3.08.1). $\widehat{N}_{\mathrm{e}}$ for T-Bone was negative with the inclusion of migrants and near the upper limit of values that can be accurately estimated by LDNe when excluding migrants (>10 000, Macbeth et al. 2013). For the remaining lakes the removal of migrants did not considerably affect the $\widehat{N}_{\mathrm{e}}$ values and generally Cabot, Esker and WP152 had $\widehat{N}_{\mathrm{e}}$ values that were about an order of magnitude greater

than those of Lake 1, Genetics H and Slushy with Strange falling between these two groups.

Metapopulation $\widehat{N}_e$ calculated using the Tufto-Hindar method (*meta*-$\widehat{N}_{e(T+H)}$) for both the whole system and only the northern six lakes, with and without potential migrants, were found to be lower than the sum of the lake $\widehat{N}_e$s calculated using LDNe ($\widehat{N}_{e(LDNe)}$) (Table 3.08.2). For all calculations of *meta*-$\widehat{N}_{e(T+H)}$ a single $\widehat{N}_{e(LDNe)}$ was used for Esker and WP152. The negative $\widehat{N}_{e(LDNe)}$ for T-Bone was excluded when the *meta*-$\widehat{N}_{e(T+H)}$ was determined using the $\widehat{N}_{e(LDNe)}$ value for each lake that was calculated with the inclusion of migrants. Migration rates between lakes calculated for each *meta*-$\widehat{N}_{e(T+H)}$ did not differ greatly from migration rates calculated for the whole system (Appendix Q).

$\widehat{N}_e$ values estimated without migrants did not correlate with lake area ($P > 0.05$)(Table 3.08.3 and see Appendix R for scatterplot). $\widehat{N}_e$ values estimated with potential migrants did significantly correlate with lake area ($R^2 = 0.920$, $F_{1,5} = 51.484$, $P \leq 0.002$) after the negative $\widehat{N}_e$ value for T-Bone was removed from the correlation analysis. In both cases Esker and WP152 were treated as a single population with a single $\widehat{N}_e$ value.

There were no significant correlations between lake elevation and any of the $\widehat{N}_e$ values calculated with or without migrants (both $P > 0.05$, Table 3.08.3, and see Appendix S for scatterplot). Again, Esker and WP152 were treated as a single population with a single $\widehat{N}_e$ value.

Positive $\widehat{N}_b$ were estimated for Genetics H, Slushy and Strange (Table 3.08.4). For all three lakes the fewest number of cohorts (between one and three) that resulted in a positive $\widehat{N}_b$ were used. The cohorts used to determine $\widehat{N}_b$ for Genetics H and Slushy did not contain any putative migrants and a positive value of $\widehat{N}_b$ for Strange only occurred with the inclusion of migrants. $\widehat{N}_{e(adj2)}$ derived from these $\widehat{N}_b$ were equal or greater to each lake's $\widehat{N}_{e(LDNe)}$.

Table 3.08.1 $\widehat{N}_e$ calculated using LDNe for each lake the inclusion and removal of migrants identified using GENECLASS2. Esker and WP152 were each analyzed separately and as a single population due to genetic similarity. The minimum allele frequency used in the analysis (P_{crit}) was 0.02. 95% confidence intervals (CI) were calculated using jackknifing between loci pairs.

Lake	With Migrants		Number Of Migrants Removed	Without Migrants	
	$\widehat{N}_e$	95% CI		$\widehat{N}_e$	95% CI
Lake 1	557.6	202.5 - ∞	4	704.3	215.4 - ∞
Genetics H	168.1	134.9 - 216.8	2	168.0	133.6 - 219.5
Slushy	313.7	209.2 - 590.2	5	256.8	178.4 - 436.1
Strange	821.0	382.8 - ∞	5	1301.1	451.7 - ∞
Esker	-3563.0	1792.5 - ∞	0	-3563.0	1792.5 - ∞
WP152	-925.2	1325.0 - ∞	3	-1185.3	992.0 - ∞
Esker/WP152	2740.4	1017.8 - ∞	1	2802.6	1060.7 - ∞
T-Bone	-39298.7	820.3 - ∞	1	11014.7	762.5 - ∞
Cabot	1196.6	302.1 - ∞	3	4287.2	332.3 - ∞

Table 3.08.2 Metapopulation $\widehat{N}_e$ values calculated using the Tufto and Hindar method (*meta*-$\widehat{N}_{e(T+H)}$, Tufto and Hindar 2003) and from the sum of lake effective size estimates calculated in LDNe ($\widehat{N}_{e(LDNe)}$). *Meta*-$\widehat{N}_{e(T+H)}$ values were calculated based on $\widehat{N}_{e(LDNe)}$ values calculated with and without migrants and for both the whole system and only the northern six lakes. The negative $\widehat{N}_{e(LDNe)}$ for T-Bone was removed from analysis when calculating the *meta*-$\widehat{N}_{e(T+H)}$ values based on those $\widehat{N}_{e(LDNe)}$ values calculated with the inclusion of migrants.

	***meta*-$\widehat{N}_{e(T+H)}$**	$\sum \widehat{N}_{e(LDNe)}$
All lakes, with migrants	5018.0	5797.4
All lakes, without migrants	2557.5	20534.7
Northern Six lakes, with migrants	2345.8	4600.8
Northern Six lakes, without migrants	2485.6	5232.8

Table 3.08.3 Correlation between $\widehat{N}_e$ values calculated with and without the exclusion of migrants identified in GENECLASS2 with lake area (km^2) and lake elevation above sea level (m). DF is degrees of freedom.

Correlation	DF	F-Statistic	*P*	R^2
$\widehat{N}_e$ with lake area without migrants	1, 5	0.687	0.445	0.121
$\widehat{N}_e$ with lake area with migrants	1, 4	51.484	0.002	0.920
$\widehat{N}_e$ with lake elevation without migrants	1, 5	0.245	0.641	0.046
$\widehat{N}_e$ with lake elevation migrants	1, 4	0.260	0.637	0.061

Table 3.08.4 The adjusted effective number of breeders $\widehat{N}_{\mathrm{b(adj2)}}$ and adjusted effective population size $\widehat{N}_{\mathrm{e(adj2)}}$ of three lakes with positive effective number of breeders ($\widehat{N}_{\mathrm{b}}$) calculated according to Waples et al. (2014). Values were calculated assuming the age at 50% maturity (α) was 10.7 and adult lifespan (AL) was equal to the maximum age of each particular lake minus 10.7 (α).

Lake	Cohort(s) (by year of birth)	*n*	$\widehat{N}_{\mathrm{b}}$	95% CI	$\widehat{N}_{\mathrm{b(adj2)}}$	$\widehat{N}_{\mathrm{e(adj2)}}$
Genetics H	2013	78	173.1	113.3 - 334.3	162.6	269.2
Slushy	2002, 2003	31	815.5	140.7 - ∞	770.9	1234.0
Strange (with migrants)	2003, 2004, 2005	37	1737.0	184.1 - ∞	1596.9	3000.6

3.09 Historical Colonization

The PCA used to pre-evaluate all three colonization scenarios in the DIYABC analysis revealed only six summary statistics over all three colonization scenarios where the proportion of values that were lower than that for the observed data set was lower than 5% or higher than 95% (but greater than 1% and lower than 99%) (see Appendix T). The models were concluded to fit well enough for further DIYABC analysis. DIYABC identified scenario 3 which depicted colonization from the west as the most likely colonization scenario with a posterior probability of 0.7801 (95% CI = (0.5487, 1.0000)) (Figure 3.09.1). This scenario included initial colonization of T-Bone, Slushy and Strange from the glacial lake Naskaupi 615 generations ago (7 380 – 8 303 years BP assuming a generation time of 12 - 13.5 (the generation time of most lakes in the Kogaluk excluding WP152 and Esker)). This date corresponds well with the period when significant overflow events occurred from lake Naskaupi through the Kogaluk between 8 400 to 7 000 years BP (Jansson and Kleman 2004). This was followed by an admixture of migrants from Strange and Slushy invading Esker 560 generations ago (6 720 -7 560 years BP) but with little input from Strange (admixture rate = 0.208). Colonization of Lake 1 and Genetics H from Esker occurred at approximately the same time (432 generations or 5 184 – 5 832 years BP for Lake 1, 466 generations or 5 592 – 6 291 years BP for Genetics H). Cabot was colonized as an admixture of migrants from T-Bone and Esker 210 generations ago (2 520 – 2 835 years BP) with approximately equal proportion of migrants from each lake (admixture rate for Esker = 0.449). $\widehat{N}_{\mathrm{e}}$ values estimated using DIYABC ($\widehat{N}_{\mathrm{e(DIYABC)}}$) tended to exceed $\widehat{N}_{\mathrm{e(LDNe)}}$ values but, in general, the relative relationships among lakes remained similar (Table 3.09.1). Type I and average type II error remained relatively low for the colonization from the west scenario at 0.168 and 0.055 respectively. Ten out of 42 summary statistics in this scenario were identified as having a proportion of values that were lower than that for the observed data set that were <5% or >95% (see Appendix U).

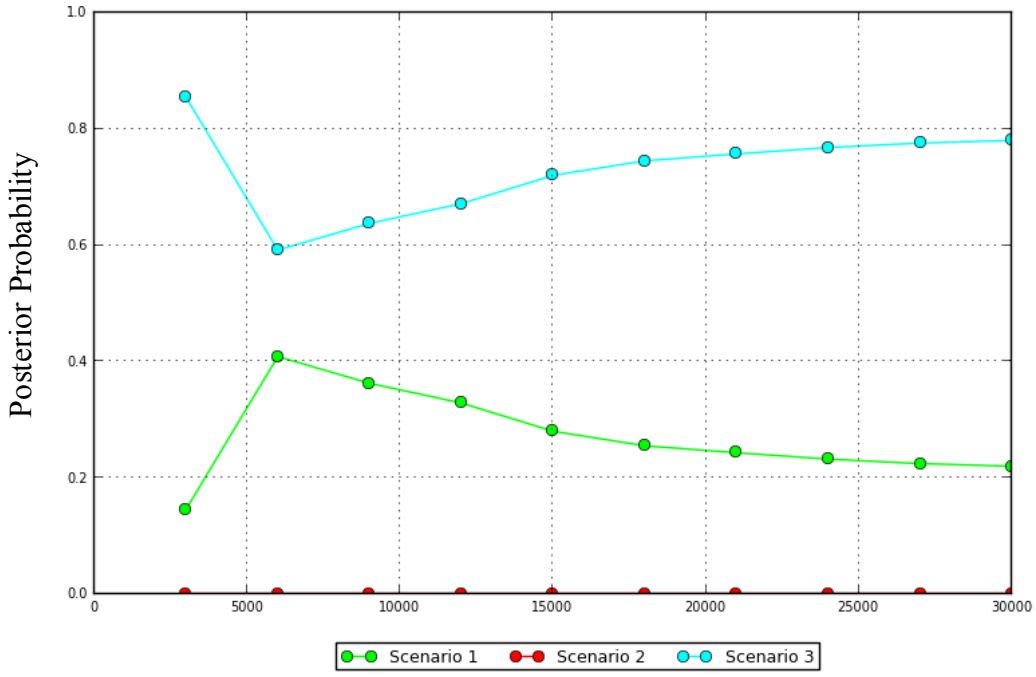

Figure 3.09.1 Posterior probabilities of three colonization scenarios based on a subset of the data sets generated in DIYABC that were closest to the observed data set.

Table 3.09.1 Prior ranges and mean posterior values with 0.025 and 0.975 quantiles in brackets for the parameters estimated for the best colonization scenario ("colonization from the west") attributed to a metapopulation of longnose suckers (*Catostomus catostomus*) in the Kogaluk River using DIYABC. N is the effective population size, t is time in generations, r is admixture rate, μ is mutation rate.

Parameter	Prior	Mean Posterior (Q0.025 - Q0.975)
N1	$1 \times 10^{1} – 5 \times 10^{4}$	5.87×10^{3} ($3.75 \times 10^{2} – 3.35 \times 10^{4}$)
N2	$1 \times 10^{1} – 5 \times 10^{4}$	8.87×10^{3} ($9.00 \times 10^{2} – 3.88 \times 10^{4}$)
N3	$1 \times 10^{1} – 5 \times 10^{4}$	3.42×10^{4} ($1.71 \times 10^{4} – 4.92 \times 10^{4}$)
N4	$1 \times 10^{1} – 5 \times 10^{4}$	1.56×10^{4} ($3.49 \times 10^{3} – 4.46 \times 10^{4}$)
N5	$1 \times 10^{1} – 5 \times 10^{4}$	4.95×10^{4} ($4.83 \times 10^{4} – 5.00 \times 10^{4}$)
N6	$1 \times 10^{1} – 5 \times 10^{4}$	7.84×10^{3} ($3.83 \times 10^{3} – 1.46 \times 10^{4}$)
N7	$1 \times 10^{1} – 5 \times 10^{4}$	2.23×10^{4} ($3.94 \times 10^{3} – 4.80 \times 10^{4}$)
ti	$1 \times 10^{1} – 7.5 \times 10^{2}$	2.10×10^{2} ($2.26 \times 10^{2} – 4.91 \times 10^{2}$)
ri	$1 \times 10^{-3} – 9.99 \times 10^{-1}$	4.49×10^{-1} ($3.36 \times 10^{-2} – 9.34 \times 10^{-1}$)
tj	$1 \times 10^{1} – 7.5 \times 10^{2}$	4.32×10^{2} ($6.25 \times 10^{1} – 6.85 \times 10^{2}$)
tk	$1 \times 10^{1} – 7.5 \times 10^{2}$	4.66×10^{2} ($7.74 \times 10^{1} – 6.94 \times 10^{2}$)
tl	$1 \times 10^{1} – 7.5 \times 10^{2}$	5.60×10^{2} ($3.13 \times 10^{2} – 7.22 \times 10^{2}$)
rl	$1 \times 10^{-3} – 9.99 \times 10^{-1}$	2.08×10^{-1} ($4.34 \times 10^{-3} – 8.33 \times 10^{-1}$)
tm	$1 \times 10^{1} – 7.5 \times 10^{2}$	6.15×10^{2} ($3.59 \times 10^{2} – 7.45 \times 10^{2}$)
μ	$1 \times 10^{-4} – 7.5 \times 10^{-3}$	8.12×10^{-4} ($4.46 \times 10^{-4} – 1.00 \times 10^{-3}$)

CHAPTER 4: DISCUSSION

4.01 Life History Implications

Longnose suckers within the Kogaluk system are slow-growing and long-lived. At 51 years, the maximum age observed in this study exceeded that observed in Great Slave Lake and the Willow Creek-Chain Lakes system in Alberta (8 and 19 respectively, Harris 1962, Walton 1980), and also exceeded the maximum age range of 22 - 24 reported by Scott and Crossman (1998). The age at 50% maturity for fish in this system at 10.7 was similar to that of longnose suckers sampled in the Great Slave Lake where no mature fish under 9 were found (Harris 1962). In contrast, longnose suckers matured between 4 - 6 years of age in Lake Superior and the Willow Creek-Chain Lakes system in Alberta (Bailey 1969, Walton 1980). These observations are consistent with the delayed maturation and increased lifespan observed at higher latitudes in other species due to the short seasonal growth period and reduced metabolism associated with cold environments (Blanck and Lamouroux 2007, Munch and Salinas 2009).

Despite the overall trend of slow-growth and long life spans, life history traits varied among lakes. Adult lifespans ranged between 6.3 and 11.3 in T-Bone and Cabot to 41.3 years in Esker. The low lifespans associated with T-Bone, a barren ground lake, could be due to higher levels of predation or competition (Hixon and Jones 2005) or to poorer habitat quality (Sinsch et al. 2007). Cabot is a deep fjord lake providing a better thermal habitat for lake trout (which have an optimal temperature range of 8 – 12 °C (Christie & Regier 1988)), than the shallow barren ground lakes where their thermal optimum could be exceeded. Cabot's depth could therefore result in a greater lake trout population size and increased predation of longnose suckers, reducing their life spans (Hixon and Jones 2005). Alternatively, Esker and WP152, which have the longest-lived fish, also have the largest combined area suggesting that there may be more habitat for suckers in these lakes, reducing competition or predation (Hixon and Jones 2005). The western extent of Esker is very shallow; potentially excluding the more heat-sensitive lake trout but providing good habitat for longnose suckers (Robert Perry, personal communication). The longer lifespans and Robson-Chapman survivorship estimates in

Esker and WP152 have driven up generation time in these lakes to 24.2 and 15.6 respectively which is substantially greater than the generation times between 12 and 13.5 observed in all other lakes. This suggests that Esker and to a lesser extent WP152 are a much more stable and higher quality habitat for longnose suckers than the other lakes which is consistent with the downstream location of these lakes (Moore et al. 2015).

4.02 Genetic Structure of Longnose Suckers within the Kogaluk

Connectivity and genetic differentiation varied widely among lakes. Most lakes in the system harboured genetically distinguishable populations with the exception of Esker and WP152 which were genetically undifferentiated. Similar results were observed in lake trout within the Kogaluk River (McCracken et al. 2013a) where all lakes except Esker and WP152 were found to have formed distinct populations. The similar genetic patterns in these two species suggest that the shallow streams connecting the lakes within this watershed are a major barrier to gene flow. These streams are frozen from late October through April and May (Wheeler 1935) and may also dry out in the summer after the initial spring surge. The shallowness of these streams can also lead to higher temperatures and reduced oxygenation (Matthews 1998, Griffiths 2010). The transient and inhospitable nature of these streams has likely limited gene flow between lakes.

Despite the challenge to migration posed by these streams, movement throughout the system does occur. Upstream migration from Esker into several headwater lakes is consistent with the capacity for upstream migration exhibited by adult individuals of this species during spawning (McPhail and Lindsey 1970, Ryan 1980, Scott and Crossman 1998). Downstream migration in this system is instead probably due to drawdown, which particularly affects juvenile longnose suckers (Ryan 1980).

The genetic differentiation among lakes was clearly influenced by this migration. Strange and Lake 1 were the most genetically distinct populations in the system as visualized in the PCoA (Figure 3.03.1). The genetic distances associated with the stream sections immediately downstream of Strange and Lake 1 were the highest in any STREAMTREE model including those sections that contained waterfalls. In the Mantel tests, Strange and Lake 1 were routinely identified as outlier lakes with higher than

expected genetic distance from the other lakes. Potentially driving this differentiation is the little to no migration experienced by Strange and Lake 1 as revealed by BAYESASS. Strange's reduced migration may be a function of its shallow outlet, which appeared to be completely dry at some points in August 2014. Lake 1 is at a higher altitude and latitude than the other lakes suggesting its connecting streams are the last in the system to thaw and the first to freeze, contributing to a narrower window in which fish migration can occur. These potential causes for reduced migration into and from Strange and Lake 1 have obvious consequences for gene flow and have likely resulted in the genetic isolation of these lakes, increased genetic drift and genetic differentiation.

On the other hand, Esker and WP152 experienced very high migration and were genetically indistinguishable. These two lakes had a pairwise $\hat{F}_{ST}$ that is an order of magnitude lower than any other observed within the system. There was also a very small genetic distance associated with the stream section(s) separating these lakes. Compellingly, these lakes were not identified as genetically separate using STRUCTURE (Figure 3.04.1) even with the use of location priors. The genetic similarity between Esker and WP152 was also observed in lake trout in these same lakes (McCracken et al. 2013a) and is likely due to the very high migration rate of 30% per generation between these lakes. This approaches the upper limits of migration rates that can be accurately estimated by BAYESASS and therefore actual migration between these lakes may be even higher (Wilson and Rannala 2003, Kanno et al. 2011). Higher flow rates as are expected in the downstream portions of a river system (Vannote 1980) potentially facilitates higher migration between these lakes (Ward and Stanford 1995).

Contrary to what was observed in lake trout in the system (McCracken et al. 2013a), Cabot was also genetically similar to Esker and WP152. These three lakes were not genetically distinguishable using STRUCTURE without location priors. The genetic distances associated with the stream sections connecting Cabot and Esker and WP152 were all relatively low (ranging between 0.0202 and 0.0224). BAYESASS also revealed a high amount (24% per generation) of migration from Esker to Cabot. This is surprising given the great geographic distance (~ 90 km) and intervening waterfall separating Cabot from Esker and WP152. However, the expected high flow rates between downstream lakes (Vannote 1980) may sweep juvenile and small fish, downstream and into Cabot

(Ryan 1980). Alternatively, the high genetic similarity between these lakes may be due to the fewer genetic samples collected in Cabot than in Esker and WP152. The high $\widehat{N}_e$ values observed in these lakes may mean that Cabot is isolated from Esker and WP152 but so genetically diverse that I was unable to detect this distinction at the given level of sampling. This is even more likely if colonization of Cabot was recent as is suggested by the DIYABC analysis. Additionally, in contrast to the STRUCTURE results the PCoA (Figure 3.03.1) indicates that Cabot is genetically intermediate to T-Bone, Esker and WP152. This result is more similar to that observed in lake trout where Cabot and T-Bone were initially grouped together and distinct from the northern six lakes in STRUCTURE (McCracken et al. 2013a). This observed similarity between Cabot and T-Bone in the PCoA suggests that Cabot may also be the recipient of migrants from T-Bone as well as from Esker. BAYESASS analysis suggests that there is no migration between T-Bone and Cabot and I observed that the stream section between Cabot and T-Bone was dry in some places in August 2014. However, one of the migrants found in Cabot had a putative origin of T-Bone suggesting a connection between these lakes which may have been greater in the past and contributed to the genetic similarity of these lakes. Additionally, the best DIYABC scenario supported the colonization of Cabot from an admixture of migrants from Esker/WP152 and T-Bone. This supports the migration from T-Bone into Cabot at least historically, if not recently.

4.03 Effects of Physical Features on Subpopulation Genetic Differentiation

The Mantel tests suggest that both waterway distance and the number of waterfalls affect genetic distance between lakes. IBD was significant within the whole watershed and among the northern six lakes but only when Lake 1 and Strange were removed from the analysis. Although only waterway distance remained significantly correlated with genetic distance after using partial Mantels, the effect of the number of waterfalls was only marginally non-significant when distance was taken into account. I therefore suspect that both factors are important in shaping the genetic structure of this system. The influence of the number of waterfalls is clearly tempered by the observed high migration rate between Esker and Cabot, suggesting that this particular waterfall is

not a complete barrier, at least in the downstream direction. The lack of migration from Cabot into Esker suggests that the waterfall between these two lakes is a complete barrier to upstream migration. Similarly, the waterfall between T-Bone and Cabot appears to be a complete barrier to migration in the upstream direction given the lack of observed recent migration in and out of T-Bone as well as the relatively high pairwise $\hat{F}_{ST}$s associated with T-Bone. However, there is some evidence of migration from T-Bone into Cabot (see section 4.02). Although downstream migration into Cabot limits the influence of waterfalls on the genetic structure of this metapopulation, waterfalls seem to pose a significant barrier to gene flow between T-Bone and the northern six lakes. Therefore the genetic divergence of these two groups of lakes should increase in future (albeit slowly due to the large N_{e} of T-Bone), resulting in a more apparent influence of waterfalls on genetic structure in this metapopulation.

Neither elevation nor slope seemed to significantly affect gene flow in this system. Elevation and slope did not differ greatly between the barren ground lakes (as is expected of lakes in this area (Wheeler 1935)) despite significant genetic differentiation. Given the known dispersal ability of longnose suckers (McPhail and Lindsey 1970, Ryan 1980, Scott and Crossman 1998) it is unlikely that the gradual slopes observed between the northern six lakes would cause any difficulty for these fish. Elevation and slope were also not found to hinder lake trout within the Kogaluk River (McCracken et al. 2013a) and slope did not substantially obstruct brook trout within a Connecticut river system (Kanno et al. 2011).

4.04 Effects of Drift on Subpopulation Genetic Differentiation

The significant negative correlations between mean pairwise allelic richness and pairwise $\hat{F}_{ST}$s and residual pairwise $\hat{F}_{ST}$s after $\hat{F}_{ST}$s were correlated with pairwise geographic distance suggest the importance of genetic drift in this metapopulation (Raeymaekers et al. 2008). These correlations are driven by Lake 1 and Strange which are, on average, more genetically distinct than expected given geographic distance (see Figure 3.06.1b and Appendix K, Figure K2b). These two lakes experience little migration and are highly isolated in the system likely causing increased drift which has resulted in a

loss of allelic richness (Table 2.01.1) and a high degree of genetic differentiation (see Figure 3.03.1). This is confirmed by the fact that removal of these two lakes results in an insignificant correlation between mean pairwise allelic richness and $\hat{F}_{ST}$ residuals (see Appendix O). Although there is still a significant correlation between $\hat{F}_{ST}$ and mean pairwise allelic richness with the exclusion of Lake 1 and Strange, this is due to a single outlier point for Esker and WP152. These two lakes have very high allelic richness and are very genetically similar. I suspect that this is due to the high flow rates expected in these downstream lakes (Vannote 1980), which contributes to high migration (Ward and Stanford 1995). Though removal of this point makes the relationship between mean pairwise allelic richness and $\hat{F}_{ST}$ insignificant, there remains an observable negative relationship which may be explained by the fact that the largest distances within a dendritic system should be between headwaters or between headwaters and confluences. This geographic isolation of headwaters contributes to their increased experience of drift, resulting in lower allelic richness. Therefore, once distance is taken into account the negative correlation between allelic richness and $\hat{F}_{ST}$ becomes insignificant. Headwaters are clearly subject to some genetic drift in this system, but not nearly to the extent of Lake 1 and Strange.

The correlation between lake area and $\widehat{N}_{\mathrm{e}}$ (calculated with migrants) when T-Bone is excluded suggests that genetic diversity within lakes is a product of available habitat. This would suggest that there is little migration between lakes and that genetic differentiation is primarily a function of drift. However, this correlation is driven predominately by the large $\widehat{N}_{\mathrm{e}}$ and lake area associated with Esker/WP152 and removing this point makes the correlation insignificant ($R^2 = 0.590$, $P \geq 0.129$, see Appendix R, Figure R2c). The correlation between $\widehat{N}_{\mathrm{e}}$ and lake area is therefore tenuous in this system which is likely due to the fact that the relative importance of drift and migration differs between lakes.

4.05 Migration-Drift Equilibrium

A metapopulation that has achieved migration-drift equilibrium is expected to develop a pattern an IBD pattern when there is at least some migration (Slatkin 1993) but

such a pattern was not initially observed in the Kogaluk. When all lakes are considered, the lack of correlation between pairwise $\hat{F}_{ST}$s and geographic distance is typical of case III described by Hutchison and Templeton (1999). Such a pattern is expected to arise in highly-divided metapopulations where drift is the main contributor to genetic differentiation (Hutchison and Templeton 1999). However, with the removal of Lake 1 and Strange, both identified as subject to a high degree of drift, the resulting relationship between pairwise $\hat{F}_{ST}$s and geographic distance is better typified by Hutchison and Templeton's (1999) case I, or IBD. Clearly migration is of greater importance in some lakes (e.g. Esker and WP152) and drift is more important in other lakes (e.g. Strange and Lake 1) suggesting that this system is not in migration-drift equilibrium.

There are several potential reasons longnose suckers in the Kogaluk have not reached migration-drift equilibrium. Slatkin's (1993) one-dimensional "radiation model" suggests observations of IBD should increase from a small to a large scale as a function of $\sqrt{2Nm\tau}$, where N is subpopulation size, m is the fraction of migrants for each generation, and τ is the time since divergence between populations. Because the Laurentide Ice Sheet retreated from the Kogaluk ~ 9 000 years ago (Bryson et al. 1969, Short and Nichols 1977), longnose suckers must have colonized this watershed only recently. This suggests that this metapopulation has a small τ and has not yet attained migration-drift equilibrium (Slatkin 1993). Alternatively, the relatively low migration rates observed between some lakes in the Kogaluk may have prevented an IBD pattern (Castric and Bernatchez 2003). High subpopulation N_e may also stall achievement of migration-drift equilibrium as was the case for lake cisco populations (*Coregonus artedi*) across Canada which demonstrated a clinal variation in microsatellite allele frequencies due to the historic secondary contact of descendants from two glacial refugia (Turgeon and Bernatchez 2001). This clinal pattern was maintained due to high lake N_e and low modern migration (Turgeon and Bernatchez 2001). Some lakes in this system demonstrated $\widehat{N}_e$ values equal or above the average lake N_e (2 753) observed in these cisco subpopulations (Turgeon and Bernatchez 2001) suggesting large N_e values could similarly be preventing achievement of migration-drift equilibrium in the Kogaluk. Barriers to gene flow can also disrupt an IBD pattern (Crispo and Hendry 2005) and within the northern six lakes there are clearly some stream sections that are more easily

traversed than others, potentially due to degree of water flow, (Ward and Stanford 1995) regardless of distance between lakes. These reasons are not necessarily mutually exclusive, for example, the variable gene flow rates between lakes may have slowed the already delayed achievement of migration-drift equilibrium due to the combination of the relatively recent colonization of the Kogaluk and the long generation time exhibited by longnose suckers in this system.

Most studies seeking to identify the presence of migration-drift equilibrium consider only the overall trend presented in the scatterplot of pairwise $\hat{F}_{ST}$s versus geographic distances (as suggested by Hutchison and Templeton 1999; e.g. Costello et al. 2003, Hänfling and Weetman 2006, Raeymaekers et al. 2008) and do not consider the potential for variation in the importance of drift and migration among lakes within watersheds. However, not excluding Strange and Lake 1 as outliers would have resulted in the erroneous conclusion that this metapopulation is highly fragmented and genetic differentiation is mainly driven by drift. My results demonstrate that it is imperative to consider the relative importance of drift and migration for individual subpopulations within a metapopulation when assessing migration-drift equilibrium.

4.06 Effects of Migrants on $\widehat{N}_e$

$\widehat{N}_e$ values for the most part did not change dramatically with the removal of putative migrants identified in GENECLASS2 with the exceptions of Strange, Lake 1 and Cabot, where $\widehat{N}_e$ values increased substantially. It is possible that the $\widehat{N}_e$ values of these lakes were depressed due to the mixture linkage disequilibrium created from the introduction of highly divergent migrants (Waples and England 2011). A large pulse of non-equilibrium migration (as might explain the large migration from Esker into Cabot) could cause such mixture linkage disequilibrium (Waples and England 2011). In the case of Lake 1 and Strange, migration rates calculated using BAYESASS were low into Strange and non-existent into Lake 1, however, 4 putative migrants in Lake 1 were identified using GENECLASS2 suggesting some gene flow. The low migration rates into these two lakes may still cause mixture linkage disequilibrium despite the high $\widehat{N}_e$ values associated with these lakes because these two lakes are so highly genetically divergent

from the other lakes that any migrants will be genetically different from those subpopulations. T-Bone's $\widehat{N}_{\mathrm{e}}$ also changes from a large negative value to a large positive value with the exclusion of its one migrant. Given that the positive $\widehat{N}_{\mathrm{e}}$ attributed to T-Bone is at the limit of values accurately calculated by LDNe (>10 000, Macbeth et al. 2013), the effect of the removal of this putative migrant only confirms that T-Bone has a very large $\widehat{N}_{\mathrm{e}}$.

4.07 Adjustment of $\widehat{\mathbf{N}}_{\mathbf{e}}$ using $\widehat{\mathbf{N}}_{\mathbf{b}}$

For those lakes where $\widehat{N}_{\mathrm{b}}$ could be estimated (Genetics H, Slushy, Strange), $\widehat{N}_{\mathrm{e(adj2)}}$ were all greater than their $\widehat{N}_{\mathrm{e(LDNE)}}$. $\widehat{N}_{\mathrm{e(adj2)}}$ were inflated over $\widehat{N}_{\mathrm{e(LDNE)}}$ by as much as 481% in the case of Slushy (when migrants were not included in the estimation of $\widehat{N}_{\mathrm{e(LDNe)}}$) and by as little as 160% in the case of Genetics H (regardless of whether migrants were or were not included in the estimation of $\widehat{N}_{\mathrm{e(LDNe)}}$). This is consistent with the results of Waples et al. (2014) who found that in $\widehat{N}_{\mathrm{e}}$ values based on individuals over all sampled age classes were less than the true $\widehat{N}_{\mathrm{e}}$ in 19 species with varying life histories. Estimates of $\widehat{N}_{\mathrm{e}}$ using LDNe are the effective size of the parents of the sampled generation (Waples 2006). This should be a function of the harmonic mean of $\widehat{N}_{\mathrm{b}}$ and $\widehat{N}_{\mathrm{e}}$, but is biased downward due to a Wahlund effect caused by the parents of a particular population having different allele frequencies from being in different cohorts (Waples et al. 2014). Waples et al. (2014) argue that $\widehat{N}_{\mathrm{e(adj2)}}$ calculated using a single cohort (and less optimally two to three cohorts) provided the most accurate estimate of $\widehat{N}_{\mathrm{e}}$ using the linkage disequilibrium method. This suggests that $\widehat{N}_{\mathrm{e}}$ values calculated using LDNe with the use of all sampled individuals are conservative estimates of the actual N_{e} in each of these lakes.

4.08 Effects of Lake Hierarchy on $\widehat{\mathbf{N}}_{\mathbf{e}}$

In general, the more downstream lakes seem to have higher $\widehat{N}_{\mathrm{e}}$ values than headwater lakes with a few notable exceptions. Cabot and the combined Esker and

WP152 have two of the highest $\widehat{N}_e$ values in the system which is consistent with their position as confluences within the system. Cabot's large $\widehat{N}_e$ is due to the pooling of genetic diversity from upstream lakes, most importantly Esker, but also potentially from T-Bone and several upstream lakes from the southern part of the system where not enough samples were collected for a genetic study (e.g. Mistastin and Genetics B). Cabot is also one of the larger lakes in the system and despite its intermediate area is the deepest lake in the system, likely providing more habitat for these longnose suckers allowing for large population sizes to maintain genetic diversity (Frankham 1996). Esker/WP152, despite having more contemporary emigration than immigration, also has a large $\widehat{N}_e$ potentially due to historical downstream gene flow and the large habitat size associated with this lake. The stability associated with downstream river habitats (Moore et al. 2015) likely also contributes to the higher $\widehat{N}_e$ values observed in Cabot and Esker/WP152.

In contrast, as expected from the isolation associated with headwaters (Morrissey and de Kerckhove 2009), Slushy and Genetics H both had the lowest $\widehat{N}_e$ values in the system despite relatively high migration from Esker. However, T-Bone and Strange, both headwaters, have some of the highest $\widehat{N}_e$ values observed in the system. T-Bone had the highest observed $\widehat{N}_e$ in the system when migrants were not included and a negative $\widehat{N}_e$ when migrants were included suggesting a "very large" effective size (Waples and Do 2010). T-Bone experiences no migration and is isolated from all other lakes by a waterfall. However, T-Bone does have a relatively large habitat area which may contribute to its large $\widehat{N}_e$ (Frankham 1996). A similar observation was made in bull trout (*Salvelinus confluentus*) populations in British Columbia, where two large lakes did not demonstrate a reduction in genetic diversity despite being situated upstream of a waterfall (Costello et al. 2003). However, other lakes in the Kogaluk with larger areas have smaller $\widehat{N}_e$ values (i.e. Esker/WP152 and Cabot). Strange has the smallest area but a large $\widehat{N}_e$ suggesting that it may have high quality habitat despite its small size, leading to increased carrying capacity and genetic diversity. Alternatively, T-Bone and Strange may be the lakes first colonized in the system as is supported by the DIYABC analysis (see section 4.11) and therefore have greater genetic diversity than those lakes that were subsequently colonized by less genetically diverse founder populations (Austerlitz et al. 1997, Le Corre and Kremer 1998).

4.09 Effects of Dendritic Structure on Metapopulation Genetic Structure

Evidence of dendricity shaping genetic structure of longnose suckers within the Kogaluk was mixed. The significant negative correlation with allelic richness and distance to Cabot within the northern six lakes is consistent with similar observations in: a waterway-restricted amphipod *Gammarus fossarum* (Alp et al. 2012), Trinidadian guppies (*Poecilia reticulata*) (Crispo et al. 2006, Barson et al. 2009), Brook trout (*Salvelinus fontinalis*) (Torterotot et al. 2014), stickleback (*Gasterosteus aculeatus* L.) (Raeymaekers et al. 2008), and *Cottus gobio* (Hänfling and Weetman 2006). It is often difficult to determine the factor causing a decrease in genetic diversity with distance from the outlet of a watershed since this can be caused not only by dendritic structure (Morrissey and de Kerckhove 2009, Paz-Vinas and Blanchet 2015), but also the loss of genetic diversity due to the founder effects of sequential upstream colonization (Crispo et al. 2006, Caldera and Bolnick 2008). However, in this system there is no ambiguity because colonization likely occurred first in the headwaters meaning that the observed decrease in allelic diversity in headwaters is most likely due to the dendritic structure of this system.

The lack of a significant correlation between allelic richness and upstream distance within the whole system is due to Cabot having a much lower allelic richness than expected given its downstream position. Perhaps this is due to Cabot being recently colonized as newly colonized populations often demonstrate a lowered allelic richness (Nei et al. 1975) due to founder effect (Dlugosch and Parker 2008). Additional support for the influence of dendricity on the metapopulation comes from the finding that the most genetically distinct lakes as revealed in the PCoA (Strange, Lake 1, T-Bone, see Figure 3.03.1) are all headwaters whereas both confluences (Esker and Cabot) are genetically intermediate to all lakes.

There was also evidence that dendricity was not the most important factor driving genetic structure of this system. There was no correlation in observed or expected heterozygosity with upstream distance. This was consistent with a lack of difference in observed and expected heterozygosity between genetic clusters of Brook trout occupying varying positions along a watershed in Connecticut, USA (Kanno et al. 2011).

Additionally, a decrease in total number of alleles but not heterozygosity with distance from the watershed outlet was observed in Trinidadian guppies (Crispo et al. 2006). However these results were inconsistent with a strong negative correlation in heterozygosity and distance from outlet found in mottled sculpin (*Cottus bairdi*) (Lamphere and Blum 2012) and brook trout (Torterotot et al. 2014) that inhabited dendritic river systems. In addition, there was no correlation between elevation and $\widehat{N}_{\mathrm{e}}$ which would indicate a downward flow of genetic diversity (McCracken et al. 2013a). However, this observation may be an indication of the lack of elevation differences in this system rather than a lack of support for downstream flow of genetic diversity in dendritic systems.

Although a strong pattern of IBD is predicted to arise as a result of a dendritic metapopulation structure (Paz-Vinas and Blanchet 2015), a pattern of IBD was only found in the Kogaluk with the exclusion of Strange and Lake 1. This is in contrast to the strong pattern of IBD observed in other species inhabiting dendritic systems including: brook trout (Kanno et al. 2011), *Cottus gobio* (Hänfling and Weetman 2006), *Gammarus fossarum* (Alp et al. 2012), sculpin (Lamphere and Blum 2012), and Quebec brook trout (Torterotot et al. 2014). The observed lack of a consistent IBD pattern in longnose suckers may be due to the absence of migration-drift equilibrium in this metapopulation due to the recent colonization of the Kogaluk. This was the conclusion for a metapopulation of bull trout in Pine River British Columbia which demonstrated little correlation between pairwise $\hat{F}_{ST}$s and geographic distance but a greater variance in $\hat{F}_{ST}$s than a more southernly metapopulation that was colonized more distantly in the past (Costello et al. 2003). Alternatively, an IBD pattern has been simulated to arise in dendritic systems when migration rates are symmetric and constant within the system and subpopulation sizes are identical (Paz-Vinas and Blanchet 2015). My longnose sucker metapopulation does not meet either criterion.

Additionally, neither the *meta*-$\widehat{N}_{\mathrm{e(T+H)}}$ of the northern six lakes nor that of the entire system exceeded the sum of the individual lake's $\widehat{N}_{\mathrm{e}}$ values suggesting that genetic diversity is not elevated in this dendritic system. This could be due to the fact that the Kogaluk is not a very complex dendritic system; lakes are not separated by more than two other lakes. Increased dendritic complexity and branching is expected to maximize

genetic differences between headwaters and confluences and increase metapopulation persistence (Mari et al. 2014, Yeakel et al. 2014). An alternative reason is the upstream dispersal from Esker into the headwater lakes which could ameliorate genetic differences accumulated in the headwaters due to genetic drift (Mills and Allendorf 1996). There was evidence for this in Slushy and Genetics H each of which had lower pairwise $\hat{F}_{ST}$s with Esker and WP152 and had higher allelic richness than Strange and Lake 1 which experienced little to no migration with Esker. This is consistent with previous work which has shown that population fragmentation, asymmetric gene flow and differing N_{e} values among subpopulations will depress a metapopulation's N_{e} (Whitlock and Barton 1997, Palstra and Ruzzante 2011).

4.10 Source/Sink Paradigm

River confluences in the dendritic paradigm are thought to be sinks, receiving genetic diversity from the more isolated headwater populations which are expected to be highly genetically distinguishable due to genetic drift (Morrissey and de Kerckhove 2009, Baguette et al. 2013). This is true in the case of Cabot as it has clearly received a boost in genetic diversity from upstream Esker. This was similar to the most downstream subpopulation of Brook Trout in a river system in Connecticut USA (Kanno et al. 2011) which received an immigration rate of ~ 30% per generation, similar to the migration rate into Cabot from Esker. High admixture and historically high immigration rates were found in Trinidadian guppy populations located in the most downstream locations of the Caroni drainage likely caused by the rapid flow of water in the lower part of the drainage (Barson et al. 2009). However, Esker, which is the main confluence of the northern six lakes, does not follow this paradigm and instead appears to act as a source of immigrants to most of the lakes within the northern arm of the Kogaluk. This is in accordance with what is expected of sink/source dynamics where large populations in rich habitats are sources for habitat-poor areas (Dias 1996) since it is known that downstream habitats are more environmentally stable than their headwater counterparts (Moore et al. 2015). Esker does however show evidence of being a confluence due to the significant correlation of upstream distance and allelic richness within the northern six lakes. Therefore Esker

appears to act as a reservoir of genetic diversity which sometimes reseeds isolated headwater populations with genetic diversity through the upstream migration of the longnose suckers for spawning. This upstream migration likely limits metapopulation N_e by increasing similarity between Esker and the headwater populations but may also contribute to the stability of the system by preventing headwaters from becoming so isolated as to go extinct (Ewers and Didham 2006). This upstream bias in migration is also predicted to increase the time to extinction in dendritic systems (Campbell Grant 2011) and has likely evolved in longnose suckers among other species to compensate for the downstream flow in river systems which would eventually result in extinction if all migration was passive (Barson et al. 2009).

4.11 Historical Colonization Implications

The DIYABC analysis results showed several deviations from the results obtained from LDNe. $\widehat{N}_{e(\mathrm{DIYABC})}$ values were typically an order of magnitude or more larger than $\widehat{N}_{e(\mathrm{LDNe})}$ values. This is not entirely implausible given that the 95% confidence intervals estimated for most $\widehat{N}_{e(\mathrm{LDNE})}$ values had an upper limit of infinity. However the high $\widehat{N}_{e(\mathrm{LDNE})}$ estimated for both Slushy and Genetics H, which each had an upper 95% confidence interval limit in the order of 10^2, contrasts significantly with the $\widehat{N}_{e(\mathrm{DIYABC})}$ values for both of these lakes which was in the order of 10^4. Such deviations may be due to the fact that DIYABC analysis assumes that no migration occurs between populations once diverged (Cornuet et al. 2008), whereas there is clear evidence of modern migration between lakes in violation of this assumption. Similarly, the type I error of 0.168 associated with the best scenario is somewhat concerning as is the model checking analysis which showed about a quarter of summary statistics to only weakly match those for the original data set.

Despite these potential issues with the analysis, the identified best scenario 3, which describes colonization from the west, is supported by previous predictions that longnose suckers invaded northern Labrador by migrating overland, north and east from the Churchill River (Black et al. 1986). The fact that some of this system's headwaters

were the first lakes to be colonized has important consequences for the genetic structure of this system. The populations that are first colonized in a stepping stone model should have the highest genetic diversity due to experiencing fewer founder effects than subsequently colonized lakes (Austerlitz et al. 1997, Le Corre and Kremer 1998). Therefore the elevated $\widehat{N}_{\mathrm{e}}$ values observed in Strange and T-Bone in both the DIYABC and LDNe analyses may be a result of their early colonization. The particularly large $\widehat{N}_{\mathrm{e(LDNe)}}$ for T-Bone may have been due to its prolonged direct connection to lake Naskaupi from 8 400 to 7 000 years BP (Jansson and Kleman 2004) which may have allowed for multiple introductions, increasing genetic diversity in this lake (Dlugosch and Parker 2008). However, Slushy was also one of the first lakes colonized according to the best scenario yet it had a relatively low $\widehat{N}_{\mathrm{e(LDNe)}}$ in contrast with a relatively high $\widehat{N}_{\mathrm{e(DIYABC)}}$. The reason for this discrepancy is not immediately apparent though likely has to do with migration patterns that occurred between colonization and the present which are not taken into account by DIYABC. This is supported by the higher contemporary gene flow observed in Slushy than in T-Bone and Strange.

The concentration of genetic diversity in the headwaters due to their initial colonization counters the expected genetic structure of dendritic systems where genetic diversity is predicted to congregate in the confluences (Morrissey and de Kerckhove 2009).There is evidence of both effects in this system. The effects of dendricity are demonstrated by the high $\widehat{N}_{\mathrm{e}}$ of Esker/WP152 and Cabot and the significant negative correlation between allelic richness and distance to the most downstream lake (Cabot). Evidence of the effects of colonization of the headwaters comes from the observed elevated $\widehat{N}_{\mathrm{e}}$ values of Strange and T-Bone. The recent colonization of the Kogaluk has likely resulted in this metapopulation being in a state of transition from one where genetic diversity is concentrated in the headwaters due to colonization patterns to one where genetic diversity is concentrated in the confluence as is typical of a dendritic metapopulation. However, a reduced genetic diversity in T-Bone may not occur for a very long time due to its lack of connectivity with the rest of the system and currently high N_{e}. Strange, however, which demonstrates modern migration with Esker, may eventually lose its high N_{e} if it begins to lose migrants downstream which would further increase its already high levels of drift.

CHAPTER 5: CONCLUSION

Despite the fact that this metapopulation is relatively small, consisting of only eight lakes, and largely unaffected by anthropogenic influences, multiple factors are clearly at play in shaping its genetic structure. Each lake appears to be subject to a unique combination of factors that dictate its genetic diversity and distinctiveness. Both migration and drift play a part in shaping the genetic structure of this metapopulation though the relative importance of each differs between lakes. Migration generally homogenizes populations and is sometimes high enough to prevent subpopulation divergence by drift (e.g. Esker, WP152, Cabot). In more isolated lakes (e.g. Strange, Lake 1) drift has led to the genetic differentiation of these lakes, though this effect has been tempered in T-Bone likely due to its large N_e. The connectivity of the system has been shaped by contemporary environmental barriers such as distance and waterfalls. However, a lack of migration-drift equilibrium indicates that historical processes have also left their mark on metapopulation structure. The likely historical colonization of the headwaters of this system has resulted in some headwaters (e.g. T-Bone and Strange) with increased genetic diversity than that expected given the isolation associated with headwaters. Yet there is evidence that the dendritic nature of the system is asserting itself, leaching genetic diversity from the headwaters to the confluences. The recent colonization of the Kogaluk means that the metapopulation has not had time to transition to a more typical dendritic system and this transition has been tempered by the long generation time and upstream migration of this species.

These traits make the longnose suckers vulnerable to extinction in this system. A long generation time means it will be difficult for this species to recover from any reduction in genetic diversity (Lippe et al. 2006). Upstream migration prevents genetic divergence between lakes (Kanno et al. 2011), reducing the metapopulation's overall genetic diversity, and ability to respond to a disturbance (Morrissey and de Kerckhove 2009). Alternatively, the long generation time means populations will lose genetic diversity from drift very slowly (Lippe et al. 2006). Upstream migration could rescue headwater populations, which are more prone to extinction due to their isolation (Baguette et al. 2013). Upstream migration was also found to increase time to extinction

for the entire metapopulation in comparison to systems where migration was downstream-biased (Campbell Grant 2011). However, any extrinsic loss of genetic diversity would be difficult for this metapopulation to recover from and because a disproportionate amount of genetic diversity is still locked up in the headwaters, particularly in Strange and T-Bone, reductions in these lakes could greatly reduce metapopulation N_e. Additionally, given that this metapopulation is in a state of transition from a neutral structure dictated by historical processes to one dictated by contemporary processes it may be more vulnerable than its neutral genetic structure suggests if modern isolation in the form of waterfalls or reduced water flow rate has not yet been reflected in its observed neutral structure (Lippe et al. 2006). This potential vulnerability and the fact that Arctic ecosystems are thought to be particularly susceptible to disturbances caused by climate change (Reist et al. 2006, Harris et al. 2012) mean that longnose suckers in this system may be at increased risk of extinction.

From a broader management perspective, this study demonstrates the need to consider not only the physical barriers but also historical colonization of the system and the biology of the species when making management decisions about dendritic riverine systems (Vera-Escalona et al 2015). It is therefore important to consider that multiple factors may be influencing gene flow within a system and the effects of these factors may not be consistent between lakes. As a consequence, not all headwaters are equal bastions of genetic material and should be prioritized accordingly. Similarly some confluences can be sinks, while others are sources, and identifying which is which will be essential given the importance of sources in maintaining genetic diversity in a metapopulation (Dias 1996).

Despite the spatial simplicity of dendritic river systems (Fullerton et al. 2010, Baguette et al. 2013), the results of my study reveal that dendritic metapopulations are clearly subject to a complex interaction of a number of factors in nature, particularly when migration-drift equilibrium has not been achieved. Dendritic systems are being increasingly recognized as an important and distinct metapopulation that requires a unique theoretical framework within the field of landscape genetics (Campbell Grant et al. 2007, Labonne et al. 2008, Campbell Grant 2011, Baguette et al. 2013). However, my results suggest that theoretical models of dendritic metapopulations, which often make

oversimplifying assumptions (Perkin and Gido 2012), do not necessarily materialize in nature. There is a need for more study of this metapopulation type in nature to test the predictions of these models, particularly over a wider range of species and dendritic spatial arrangements.

References

Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology*, *57*(5), 969-981.

Altermatt, F. (2013). Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology*, *47*(3), 365-377.

Anderson, T. C. (1985). *Rivers of Labrador*. Canadian Special Publication of Fisheries and Aquatic Sciences, *81*, 1-389.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC bioinformatics*, *9*(1), 323.

Araki, H., Waples, R. S., Ardren, W. R., Cooper, B., & Blouin, M. S. (2007). Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. *Molecular Ecology*, *16*(5), 953-966.

Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P. H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, *51*(2), 148-164.

Bailey, M. M. (1969). Age, growth, and maturity of the longnose sucker Catostomus catostomus, of western Lake Superior. *Journal of the Fisheries Board of Canada*, *26*(5), 1288-1298.

Baguette, M., Blanchet, S., Legrand, D., Stevens, V. M., & Turlure, C. (2013). Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, *88*(2), 310-326.

Barnett, D. M., & Peterson, J. A. (1964). The significance of glacial Lake Naskaupi 2 in the deglaciation of Labrador-Ungava*. *The Canadian Geographer/Le Géographe canadien*, *8*(4), 173-181.

Barson, N. J., Cable, J., & Van Oosterhout, C. (2009). Population genetic analysis of microsatellite variation of guppies (Poecilia reticulata) in Trinidad and Tobago: evidence for a dynamic source–sink metapopulation structure, founder events and population bottlenecks. *Journal of evolutionary biology*, *22*(3), 485-497.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289-300.

Bernatchez, L., & Wilson, C. C. (1998). Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, *7*(4), 431-452.

Birch, L. (1948). The intrinsic rate of natural increase of an insect population. *The Journal of Animal Ecology*, 15-26.

Black, G. A., Dempson, J. B., & Bruce, W. J. (1986). Distribution and postglacial dispersal of freshwater fishes of Labrador. *Canadian journal of zoology*, *64*(1), 21-31.

Blanck, A., & Lamouroux, N. (2007). Large-scale intraspecific variation in life-history traits of European freshwater fish. *Journal of Biogeography*, *34*(5), 862-875.

Bryson, R.A., Wendland, W.M., Ives, J.D. & Andrews, J.T. (1969). Radiocarbon isochrones on the disintegration of the Laurentide Ice Sheet. *Arctic and Alpine Research*, *1*(1), 1-13.

Caldera, E. J., & Bolnick, D. I. (2008). Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (Gasterosteus aculeatus) populations in a single watershed. *Evolutionary Ecology Research*, *10*(4), 575-598.

Campbell Grant, E. H., Lowe, W. H., & Fagan, W. F. (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, *10*(2), 165-175.

Campbell Grant, E. H. (2011). Structural complexity, movement bias, and metapopulation extinction risk in dendritic ecological networks. *JNABS Journal*, *30*(1), 252-258.

Cardall, B. L., Bjerregaard, L. S., & Mock, K. E. (2007). Microsatellite markers for the June sucker (Chasmistes liorus mictus), Utah sucker (Catostomus ardens), and five other catostomid fishes of western North America. *Molecular Ecology Notes*, *7*(3), 457-460.

Carlander, K.D., & Smith, L.L. Jr. (1945). Some Factors to Consider in the Choice between Standard, Fork, or Total Lengths in Fishery Investigations. *Copeia*, *1945*(1), 7-12.

Carrara, F., Rinaldo, A., Giometto, A., & Altermatt, F. (2014). Complex interaction of dendritic connectivity and hierarchical patch size on biodiversity in river-like landscapes. *The American Naturalist*, *183*(1), 13-25.

Carvalho, G. R. (1993). Evolutionary aspects of fish distribution: genetic variability and adaptation. *Journal of Fish Biology*, *43*(sA), 53-73.

Castric, V., & Bernatchez, L. (2003). The rise and fall of isolation by distance in the anadromous brook charr (Salvelinus fontinalis Mitchill). *Genetics*, *163*(3), 983-996.

Castric, V., Bonney, F., & Bernatchez, L. (2001). Landscape structure and hierarchical genetic diversity in the brook charr, Salvelinus fontinalis. *Evolution*, *55*(5), 1016-1028.

Chapman, D., & Robson, D. S. (1960). The analysis of a catch curve. *Biometrics*, *16*(3), 354-368.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195-205.

Childress, E. S., Papke, R., & McIntyre, P. B. (2015). Spawning success and early life history of longnose suckers in Great Lakes tributaries. *Ecology of Freshwater Fish*. doi: 10.1111/eff.12220.

Christie, G. C., & Regier, H. A. (1988). Measures of optimal thermal habitat and their relationship to yields for four commercial fish species. *Canadian Journal of Fisheries and Aquatic Sciences*, *45*(2), 301-314.

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J., & Estoup, A. (2014). DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, *30*(8), 1187-1189.

Cornuet, J. M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1. 0). *Bmc Bioinformatics*, *11*(1), 401.

Cornuet, J. M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J. M., Balding, D. J., Guillemaud, T., & Estoup, A. (2008). Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, *24*(23), 2713-2719.

Le Corre, V., & Kremer, A. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, *11*(4), 495-512.

Costello, A. B., Down, T. E., Pollard, S. M., Pacas, C. J., & Taylor, E. B. (2003). The influence of history and contemporary stream hydrology on the evolution of genetic diversity within species: an examination of microsatellite DNA variation in bull trout, Salvelinus confluentus (Pisces: Salmonidae). *Evolution*, *57*(2), 328-344.

Cote, D., Kehler, D. G., Bourne, C., & Wiersma, Y. F. (2009). A new measure of longitudinal connectivity for stream networks. *Landscape Ecology*, *24*(1), 101-113.

Craig, P. C. (1989). An introduction to anadromous fishes in the Alaskan Arctic. *Biological Papers of the University of Alaska*, *24*, 27-54.

Crispo, E., & Hendry, A. P. (2005). Does time since colonization influence isolation by distance? A meta-analysis. *Conservation Genetics*, *6*(5), 665-682.

Crispo, E., Bentzen, P., Reznick, D. N., Kinnison, M. T., & Hendry, A. P. (2006). The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, *15*(1), 49-62.

Dias, P. C. (1996). Sources and sinks in population biology. *Trends in Ecology & Evolution*, *11*(8), 326-330.

Dieringer, D., & Schlötterer, C. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, *3*(1), 167-169.

Dillinger Jr, R. E., Birt, T. P., Green, J. M., & Davidson, W. S. (1991). Postglacial dispersal of longnose suckers, *Catostomus catostomus*, in the Mackenzie and Yukon Drainages. *Biochemical systematics and ecology*, *19*(8), 651-657.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431-449.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, *4*(2), 359-361.

Edwards, E. A. 1983. Habitat suitability index models: Longnose sucker. OBS-82/10.35. U.S. Department of the Interior. U.S. Fish and Wildlife Service. Washington, DC, Maryland.

Elphinstone, M. S., Hinten, G. N., Anderson, M. J., & Nock, C. J. (2003). An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology Notes*, *3*(2), 317-320.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, *14*(8), 2611-2620.

Ewers, R. M., & Didham, R. K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews*, *81*(01), 117-142.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, *10*(3), 564-567.

Fagan, W. F. (2002). Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology*, *83*(12), 3243-3249.

Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, *10*(6), 1500-1508.

Fullerton, A. H., Burnett, K. M., Steel, E. A., Flitcroft, R. L., Pess, G. R., Feist, B. E.,Torgersen, C. E., Miller, D. J., & Sanderson, B. L. (2010). Hydrological connectivity for riverine fish: measurement challenges and research opportunities. *Freshwater Biology*, *55*(11), 2215-2237.

Geen, G. H., Northcote, T. G., Hartman, G. F., & Lindsey, C. C. (1966). Life histories of two species of catostomid fishes in Sixteenmile Lake, British Columbia, with particular reference to inlet stream spawning. *Journal of the Fisheries Board of Canada*, *23*(11), 1761-1788.

Greenbaum, G., Templeton, A. R., Zarmi, Y., & Bar-David, S. (2014). Allelic Richness following Population Founding Events–A Stochastic Modeling Framework Incorporating Gene Flow and Genetic Drift. *PloS one*, *9*(12), e115203.

Griffiths, D. (2010). Pattern and process in the distribution of North American freshwater fish. *Biological Journal of the Linnean Society*, *100*(1), 46-61.

Gomez-Uchida, D., Palstra, F. P., Knight, T. W., & Ruzzante, D. E. (2013). Contemporary effective population and metapopulation size (Ne and meta-Ne): comparison among three salmonids inhabiting a fragmented system and differing in gene flow and its asymmetries. *Ecology and evolution*, *3*(3), 569-580.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet J. (1995) FSTAT Version 1.2: A computer program to calculate F-statistics. Heredity, *86*, 485-486.

Hänfling, B., & Weetman, D. (2006). Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the river sculpin, Cottus gobio. *Genetics*, *173*(3), 1487-1501.

Harris, R. H. (1962). Growth and Reproduction of the Longnose Sucker, Catostomus catostomus (Forster), in Great Slave Lake. *Journal of the Fisheries Board of Canada*, *19*(1), 113-126.

Harris, L. N., Taylor, E. B., Tallman, R. F., & Reist, J. D. (2012). Gene flow and effective population size in two life-history types of broad whitefish Coregonus nasus from the Canadian Arctic. *Journal of fish biology*, *81*(1), 288-307.

Harry, A. V., Tobin, A. J., & Simpfendorfer, C. A. (2013). Age, growth and reproductive biology of the spot-tail shark, *Carcharhinus sorrah*, and the Australian blacktip shark, *C. tilstoni*, from the Great Barrier Reef World Heritage Area, north-eastern Australia. *Marine and Freshwater Research*, *64*(4), 277-293.

Hixon, M. A., & Jones, G. P. (2005). Competition, predation, and density-dependent mortality in demersal marine fishes. *Ecology*, *86*(11), 2847-2859.

Horreo, J. L., Martinez, J. L., Ayllon, F., Pola, I. G., Monteoliva, J. A., Héland, M., & Garcia-Vazquez, E. V. A. (2011). Impact of habitat fragmentation on the genetics of populations in dendritic landscapes. *Freshwater Biology*, *56*(12), 2567-2579.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular ecology resources*, *9*(5), 1322-1332.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, *59*(7), 573-583.

Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, *53*(6), 1898-1914.

Ives, J. D. (1960). The deglaciation of Labrador-Ungava–An outline. *Cahiers de géographie du Québec*, *4*(8), 323-343.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801-1806.

Jansson, K. N. (2003). Early Holocene glacial lakes and ice marginal retreat pattern in Labrador/Ungava, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *193*(3), 473-501.

Jansson, K. N., & Kleman, J. (2004). Early Holocene glacial lake meltwater injections into the Labrador Sea and Ungava Bay. *Paleoceanography*, *19*, PA1001.

Jarne, P., & Lagoda, P. J. (1996). Microsatellites, from molecules to populations and back. *Trends in ecology & evolution*, *11*(10), 424-429.

Kalinowski, S. T., Meeuwig, M. H., Narum, S. R., & Taper, M. L. (2008). Stream trees: a statistical method for mapping genetic differences between populations of freshwater organisms to the sections of streams that connect them. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(12), 2752-2760.

Kanno, Y., Vokoun, J. C., & Letcher, B. H. (2011). Fine-scale population structure and riverscape genetics of brook trout (Salvelinus fontinalis) distributed continuously along headwater channel networks. *Molecular Ecology*, *20*(18), 3711-3729.

Koizumi, I., Yamamoto, S., & Maekawa, K. (2006). Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. *Molecular Ecology*, *15*(11), 3175-3189.

Labonne, J., Ravigné, V., Parisi, B., & Gaucherel, C. (2008). Linking dendritic network structures to population demogenetics: the downside of connectivity. *Oikos*, *117*(10), 1479-1490.

Lamphere, B. A., & Blum, M. J. (2012). Genetic estimates of population structure and dispersal in a benthic stream fish. *Ecology of Freshwater Fish*, *21*(1), 75-86.

Legendre, P., & Legendre, V. (1984). Postglacial dispersal of freshwater fishes in the Québec peninsula. *Canadian Journal of Fisheries and Aquatic Sciences*, *41*(12), 1781-1802.

Lippe, C., Dumont, P., & Bernatchez, L. (2004). Isolation and identification of 21 microsatellite loci in the Copper redhorse (Moxostoma hubbsi; Catostomidae) and their variability in other catostomids. *Molecular Ecology Notes*, *4*(4), 638-641.

Lippe, C., Dumont, P., & Bernatchez, L. (2006). High genetic diversity and no inbreeding in the endangered copper redhorse, Moxostoma hubbsi (Catostomidae, Pisces): the positive sides of a long generation time. *Molecular Ecology*, *15*(7), 1769-1780.

Lopoukhine, N., Prout, N. A., & Hirvonen, H. E. (1978). *Ecological land classification of Labrador: a reconnaissance*. Ecological land classification series No. 4, Lands Directorate (Atlantic Region), Environmental Management Service, Fisheries Environment Canada, Halifax, Nova Scotia, Canada.

Lowe, W. H., Likens, G. E., McPeek, M. A., & Buso, D. C. (2006). Linking direct and indirect data on dispersal: isolation by slope in a headwater stream salamander. *Ecology*, *87*(2), 334-339.

Macbeth, G. M., Broderick, D., Buckworth, R. C., & Ovenden, J. R. (2013). Linkage disequilibrium estimation of effective population size with immigrants from divergent populations: a case study on Spanish Mackerel (Scomberomorus commerson). *G3: Genes, Genomes, Genetics*, *3*(4), 709-717.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in ecology & evolution*, *18*(4), 189-197.

Mari, L., Casagrandi, R., Bertuzzo, E., Rinaldo, A., & Gatto, M. (2014). Metapopulation persistence and species spread in river networks. *Ecology letters*, *17*(4), 426-434.

Matthews, W. J. (1998). *Patterns in freshwater fish ecology*. Chap-man & Hall, New York.

McCracken, G. R., Perry, R., Keefe, D., & Ruzzante, D. E. (2013a). Hierarchical population structure and genetic diversity of lake trout (Salvelinus namaycush) in a dendritic system in Northern Labrador. *Freshwater Biology*, *58*(9), 1903-1917.

McCracken, G. R., Wilson, K. L., Paterson, I., Perry, R., Keefe, D., & Ruzzante, D. E. (2013b). Development of 17 novel microsatellite markers for the longnose sucker (Catostomus catostomus) and successful cross-specific amplification of 14 previously developed markers from congeneric species. *Conservation Genetics Resources*, 1-4. doi: 10.1007/s12686-013-0086-3.

McPhail, J. D., & Lindsey, C. C. (1970). *Freshwater fishes of northwestern Canada and Alaska*. Bulletin 173, Ottawa: Fisheries Research Board of Canada.

Michaud, W. K., Perry, R. C., Dempson, J. B., Shears, M., & Power, M. (2010). Occurrence of Lake Chub, Couesius plumbeus, in Northern Labrador. *The Canadian Field-Naturalist*, *124*(2), 113-117.

Mills, L. S., & Allendorf, F. W. (1996). The one-migrant-per-generation rule in conservation and management. *Conservation Biology*, *10*(6) 1509-1518.

Moore, J. W., Beakes, M. P., Nesbitt, H. K., Yeakel, J. D., Patterson, D. A., Thompson, L. A., ... & Atlas, W. I. (2015). Emergent stability in a large, free-flowing watershed. *Ecology*, *96*(2), 340-347.

Morrissey, M. B., & de Kerckhove, D. T. (2009). The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. The American Naturalist, 174(6), 875-889.

Munch, S. B., & Salinas, S. (2009). Latitudinal variation in lifespan within species is explained by the metabolic theory of ecology. *Proceedings of the National Academy of Sciences*, *106*(33), 13860-13864.

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*(s1), 184-194.

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 1-10.

van Oosterhout, C., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4(3), 535-538.

Orsini, L., Mergeay, J., Vanoverbeke, J., & Meester, L. (2013a). The role of selection in driving landscape genomic structure of the waterflea Daphnia magna. *Molecular Ecology*, *22*(3), 583-601.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & Meester, L. (2013b). Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular ecology*, *22*(24), 5983-5999.

Paetkau, D., Slade, R., Burden, M., & Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular ecology*, *13*(1), 55-65.

Palstra, F. P., & Ruzzante, D. E. (2011). Demographic and genetic factors shaping contemporary metapopulation effective size and its empirical estimation in salmonid fish. *Heredity*, *107*(5), 444-455.

Paz-Vinas, I., & Blanchet, S. (2015). Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *Journal of evolutionary biology*, *28*(4), 986-994.

Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288-295.

Perkin, J. S., & Gido, K. B. (2012). Fragmentation alters stream fish community structure in dendritic ecological networks. *Ecological Applications*, *22*(8), 2176-2187.

Perry, R. C., & Casselman, J. M. (2012). Comparisons of precision and bias with two age interpretation techniques for opercular bones of longnose Sucker, a long-lived northern fish. *North American Journal of Fisheries Management*, *32*(4), 790-795.

Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: an integrated network perspective. *Ecology Letters*, *16*(5), 707-719.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of heredity*, *95*(6), 536-539.

Poissant, J., Knight, T. W., & Ferguson, M. M. (2005). Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, *14*(5), 1321-1331.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raeymaekers, J. A., Maes, G. E., Geldof, S., Hontis, I., Nackaerts, K., & Volckaert, F. A. (2008). Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary applications*, *1*(3), 475-488.

Rannala, B., & Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, *94*(17), 9197-9201.

Rannala B. (2007) BayesAss Edition 3.0 User's Manual. Accessed online 20 May 2015 at: http://www.rannala.org/?page_id=245.

Reist, J. D., Wrona, F. J., Prowse, T. D., Power, M., Dempson, J. B., Beamish, R. J., King, J. R., Carmichael, T. J., & Sawatzky, C. D. (2006). General effects of climate change on Arctic fishes and fish populations. *AMBIO: A Journal of the Human Environment*, *35*(7), 370-380.

Robson, D. S., & Chapman, D. G. (1961). Catch curves and mortality rates. *Transactions of the American Fisheries Society*, *90*(2), 181-189.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, *4*(1), 137-138.Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219-1228.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219-1228.

Ryan, P. M. 1980. Fishes of the Lower Churchill River, Labrador. Fisheries and Marine Service. Technical Report No. 922.

Schlosser, I. J. (1990). Environmental variation, life history attributes, and community structure in stream fishes: implications for environmental management and assessment. Environmental Management, 14(5), 621-628.

Scott, W. B., & Crossman, E. J. 1998. *Freshwater fishes of Canada*. Galt House Publications Ltd., Oakville, Ontario, Canada.

Short, S. K., & Nichols, H. (1977). Holocene pollen diagrams from subarctic Labrador-Ungava: Vegetational history and climatic change. *Arctic and Alpine Research*, *9*(3), 265-290.

Sinsch, U., Leskovar, C., Drobig, A., König, A., & Grosse, W. R. (2007). Life-history traits in green toad (Bufo viridis) populations: indicators of habitat quality. *Canadian journal of zoology*, *85*(5), 665-673.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*(1) 264-279.

Stelkens, R. B., Jaffuel, G., Escher, M., & Wedekind, C. (2012). Genetic and phenotypic population divergence on a microgeographic scale in brown trout. *Molecular ecology*, *21*(12), 2896-2915.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., & Waits, L. P. (2007). Putting the 'landscape'in landscape genetics. *Heredity*, *98*(3), 128-142.

Torterotot, J. B., Perrier, C., Bergeron, N. E., & Bernatchez, L. (2014). Influence of Forest Road Culverts and Waterfalls on the Fine-Scale Distribution of Brook Trout Genetic Diversity in a Boreal Watershed. *Transactions of the American Fisheries Society*, *143*(6), 1577-1591

Tranah, G. J., Agresti, J. J., & May, B. (2001). New microsatellite loci for suckers (Catostomidae): primer homology in Catostomus, Chasmistes, and Deltistes. *Molecular Ecology Notes*, *1*(1-2), 55-60.

Tufto, J., & Hindar, K. (2003). Effective size in management and conservation of subdivided populations. *Journal of Theoretical Biology*, *222*(3), 273-281.

Turgeon, J., & Bernatchez, L. (2001). Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of Coregonus artedi (Teleostei: Coregoninae). *Evolution*, *55*(11), 2274-2286.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian journal of fisheries and aquatic sciences*, *37*(1), 130-137.

Vera-Escalona, I., Habit, E., & Ruzzante, D. E. (2015). Echoes of a distant time: effects of historical processes on contemporary genetic patterns in Galaxias platei in Patagonia. *Molecular ecology*, *24*(16), 4112-4128.

Walton, B. D. 1980. The reproductive biology, early life history, and growth of white suckers, *Catostomus commersoni*, and longnose suckers, *C. catostomus*, in the Willow Creek - Chain Lakes System, Alberta. MSc thesis. University of Alberta, Edmonton.

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. *Conservation Genetics*, *7*(2), 167-184.

Waples, R. S., & Do, C. H. I. (2008). LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular ecology resources*, *8*(4), 753-756.

Waples, R. S. (2010). Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Molecular Ecology Resources*, *10*(5), 785-796.

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, *3*(3), 244-262.

Waples, R. S., & England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, *189*(2), 633-644.

Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, *197*(2), 769-780.

Ward, J. V., & Stanford, J. A. (1995). Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. *Regulated Rivers: Research & Management*, *11*(1), 105-119.

Wheeler 2nd, E. P. (1935). The Nain-Okak Section of Labrador*. *Geographical Review*, *25*(2), 240-254.

Whitlock, M. C., & Barton, N. H. (1997). The effective size of a subdivided population. *Genetics*, *146*(1), 427-441.

Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, *163*(3), 1177-1191.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(2), 97.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Yeakel, J. D., Moore, J. W., Guimaraes, P. R., & Aguiar, M. A. M. (2014). Synchronisation and stability in river metapopulation networks. *Ecology letters*, *17*(3), 273-283.

APPENDIX A: LIFE HISTORY ANALYSES USING CORRECTED AGES

Life history traits and $\widehat{N}_b$ values were estimated using operculum ages corrected according to Perry and Casselman (2012) where the age for those individuals with 10 or more annuli was adjusted to account for annuli hidden under the dense bone region using the formula: $Y_{section} = 2.69+0.971X_{whole}$, where $Y_{section}$ represents the age corrected for annuli hidden under the dense bone region and X_{whole} represents the number of visible annuli (Perry and Casselman 2012). Adjusted ages were rounded down to the nearest year.

i) Life History Analysis

Ages ranged from young of the year (0) to 52 years old over all aged individuals (n = 1 353). Age at 50% maturity (α) was 12.1 (Figure A1). Adult lifespan (AL) ranged from 6.9 in T-Bone to 40.9 in Esker (Table A1).

ii) Generation Time

The natural logarithm transformation of the ages of mature, females was significantly correlated with fork length ($P < 0.001$, adjusted R^2 value of 0.58, Figure A2) as follows:

$$FL = 13.72 \times \log_e(Age) + 0.14) \quad [5]$$

Where FL is the fork length in cm.

FL in equation [5] was equivocated with total length (TL) in equation [1]:

$$f = 0.016 \times [13.72 \times (\log_e(Age) + 0.14)^{3.799}] \quad [6]$$

Where f is the number of eggs produced. Equation [6] was used to generate f for each age class.

There was no significant difference in age distribution within lakes between collection years based on Kolmogorov-Smirnov tests (all $P > 0.05$, Table A2). Annual samples within lakes were thus pooled. Robson-Chapman estimates of annual survival varied from 0.62 in Cabot to 0.91 in Esker (Table A3).

Generation time varied from 13.9 in Slushy and T-Bone to 25.3 in Esker (Table A4).

iii) Effective Size

Positive $\widehat{N}_{\mathrm{b}}$ were estimated for Genetics H, Slushy and Strange (Table A5). For all three lakes the fewest number of cohorts (between one and three) that resulted in a positive $\widehat{N}_{\mathrm{b}}$ were used. The cohorts used to determine $\widehat{N}_{\mathrm{b}}$ for Genetics H did not contain any putative migrants and a positive value of $\widehat{N}_{\mathrm{b}}$ for Strange only occurred with the inclusion of migrants. A positive $\widehat{N}_{\mathrm{b}}$ was calculated for Slushy both with and without migrants, however, these values were not significantly different. $\widehat{N}_{\mathrm{e(adj2)}}$ derived from these $\widehat{N}_{\mathrm{b}}$ were equal or greater to each lake's $\widehat{N}_{\mathrm{e(LDNe)}}$.

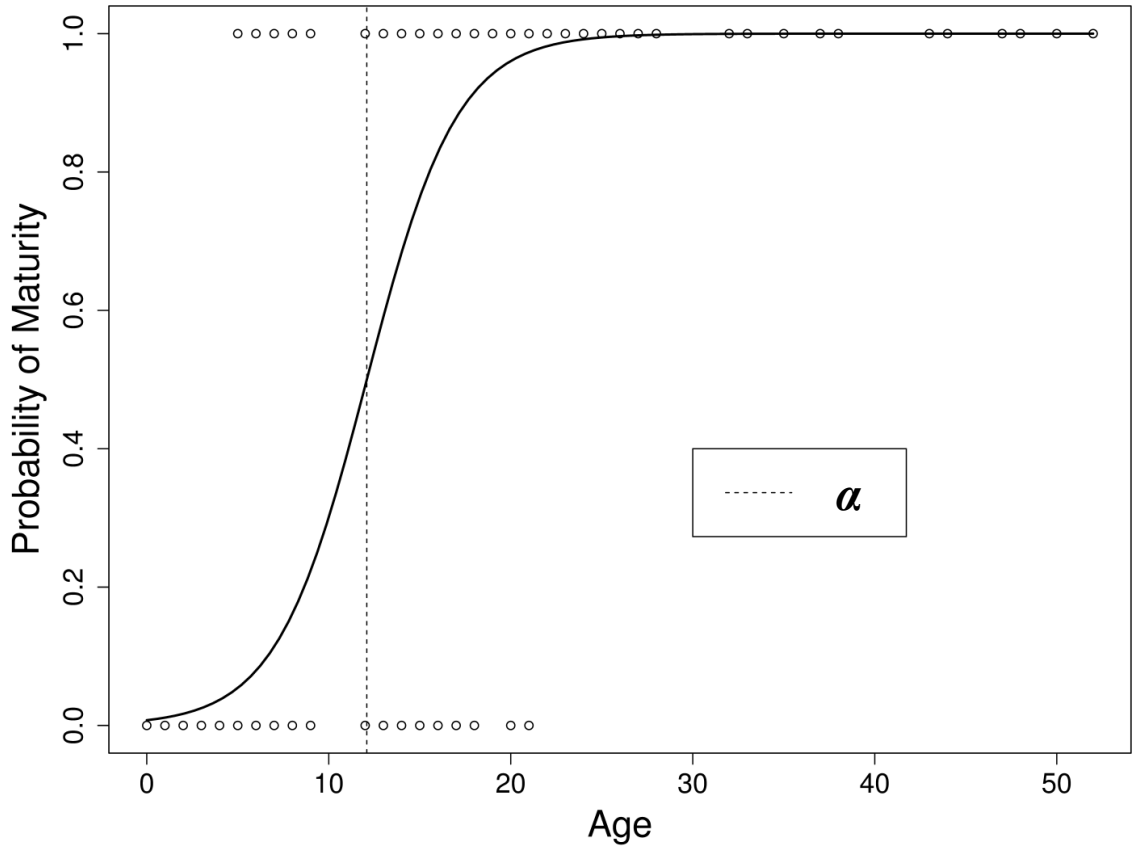

Figure A1 Age at 50% maturity (α) for longnose suckers (*Catostomus catostomus*) estimated using n = 1072 samples from the Kogaluk using a binomial logistic regression. Calculations were done using ages corrected according to Perry and Casselman (2012).

Table A1 Maximum age (ω) and adult life span (AL) of longnose suckers (*Catostomus catostomus*) in eight lakes within the Kogaluk River. Calculations were done using ages corrected according to Perry and Casselman (2012).

Lake	ω	AL
Lake 1	33	21.9
Genetics H	26	14.9
Slushy	27	15.9
Strange	24	12.9
Esker	52	40.9
WP152	37	25.9
T-Bone	18	6.9
Cabot	23	11.9

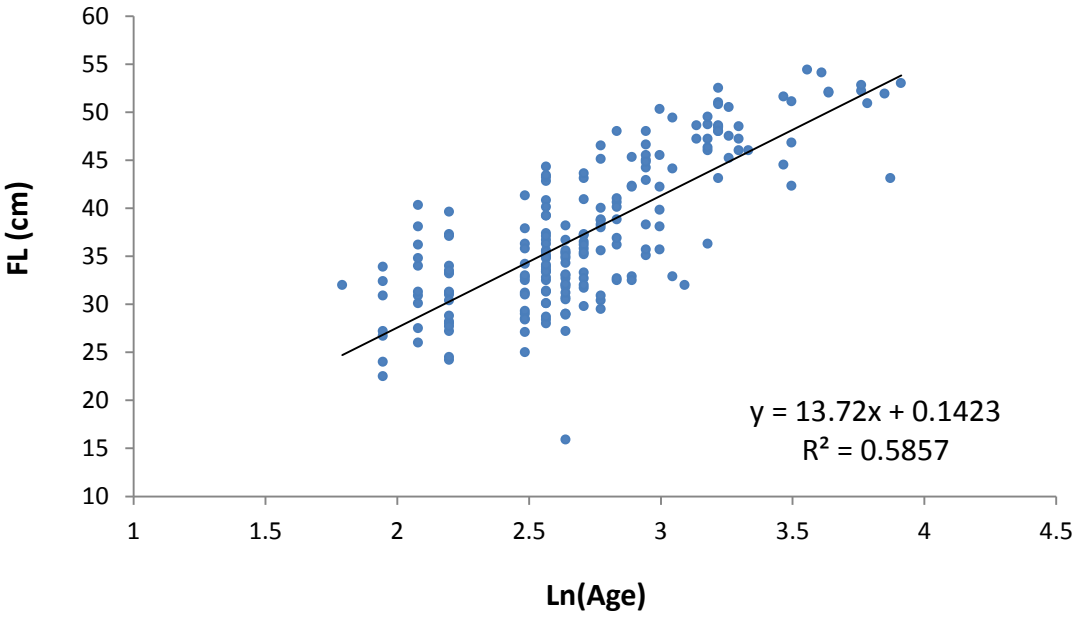

Figure A2 correlation between fork length in cm (FL) with the natural logarithm transformation at age of mature, female longnose suckers caught in the Kogaluk River. Calculations were done using ages corrected according to Perry and Casselman (2012).

Table A2 Kolmogorov-Smirnov test results comparing corrected age class distributions between years in each lake. Calculations were done using ages corrected according to Perry and Casselman (2012).

Comparison	*P*
WP152 2009 vs. WP152 2010	0.82
WP152 2009 vs. WP152 2011	0.67
WP152 2009 vs. WP152 2013	>0.99
WP152 2010 vs. WP152 2011	>0.99
WP152 2010 vs. WP152 2013	0.67
WP152 2011 vs. WP152 2013	0.52
Slushy 2010 vs. Slushy 2011	0.82
Slushy 2010 vs. Slushy 2012	0.99
Slushy 2011 vs. Slushy 2012	0.27
Genetics H 2009 vs. Genetics H 2011	0.82
Genetics H 2009 vs. Genetics H 2013	0.99
Genetics H 2011 vs. Genetics H 2013	0.52
Lake 1 2010 vs. Lake 1 2011	>0.99
Lake 1 2010 vs. Lake 1 2014	0.82
Lake 1 2011 vs. Lake 1 2014	0.52

Table A3 Robson-Chapman annual survival rate ($\hat{S}$) for eight lakes within the Kogaluk. Calculations were done using ages corrected according to Perry and Casselman (2012).

Lake	Robson-Chapman $\hat{S}$	95% CI
Lake 1	0.704	0.613 - 0.795
Genetics H	0.706	0.654 - 0.759
Slushy	0.602	0.518 - 0.685
Strange	0.716	0.666 - 0.765
Esker	0.908	0.891 - 0.925
WP152	0.795	0.741 - 0.850
T-Bone	0.692	0.609 - 0.775
Cabot	0.623	0.553 - 0.693

Table A4 Generation time calculated for eight lakes within the Kogaluk assuming age at 50% maturity (α) is 12 and maximum age (AL) is equal to the oldest fish observed in each particular lake. Calculations were done using ages corrected according to Perry and Casselman (2012).

Lake	Generation Time
Lake 1	15.1
Genetics H	15.0
Slushy	13.9
Strange	15.0
Esker	25.3
WP152	17.3
T-Bone	14.0
Cabot	14.0

Table A5 The adjusted effective number of breeders $\widehat{N}_{\mathrm{b(adj2)}}$ and adjusted effective population size $\widehat{N}_{\mathrm{e(adj2)}}$ of three lakes with positive effective number of breeders ($\widehat{N}_{\mathrm{b}}$) calculated according to Waples et al. (2014). Values were calculated assuming the age at 50% maturity (α) was 12 and AL was equal to the maximum age of each particular lake minus 12 (α). Calculations were done using ages corrected according to Perry and Casselman (2012).

Lake	Cohort(s) (by year of birth)	Sample size	$\widehat{N}_{\mathrm{b}}$	95% CI	$\widehat{N}_{\mathrm{b(adj2)}}$	$\widehat{N}_{\mathrm{e(adj2)}}$
Genetics H	2013	78	173.1	113.3 - 334.3	160.2	289.1
Slushy (with migrants)	2003, 2004	27	232.6	105.0 - ∞	216.6	376.5
Slushy (without migrants)	2003, 2004	24	218.5	88.1 - ∞	203.5	353.7
Strange (with migrants)	2003, 2004, 2005	37	1737.0	184.1 - ∞	1584.8	3128.6

APPENDIX B: PRIMERS USED

Table B1 Primers used to genotype longnose sucker samples from the Kogaluk River in Labrador, Canada.

Locus	Repeat Motif	Sequence	Source
CCAT7	$(GTTT)_4$	F:CTCCGGTGCAGTTTCTTCC R:ACTCTGATACATACTCTGCAAGC	McCracken et al. 2013b
CCAT16	$(AAAG)_{11}$	F:TACCTGGGTTGGTTGCAGG R:GTGACGGGAGGCTGGTAG	McCracken et al. 2013b
CCAT20	$(ACTC)_5$	F:CTTCTCTGTGCTGCCCAAC R:GGGCTTGACAGACTTGTGG	McCracken et al. 2013b
CCAT32	$(CT)_{10}$	F:TCCTTACGTGTGATTATTCTGGC R:AGCGTAAGTCTGATAGGTGTC	McCracken et al. 2013b
CCAT35	$(CT)_{14}$	F:AGGCATCAAATCCTTGGCAG R:CCTGTGAGACTGTGTGAAACC	McCracken et al. 2013b
CCAT43	$(ATC)_7$	F:CGTGCTCTGCTTACATTACCAC R:GATGGAAAGGCACCCGTAG	McCracken et al. 2013b
CCAT44	$(AG)_{13}$	F:CATCGGAATGGCTTCATGGG R:AAATTAAATGAGCCTGAGGTGG	McCracken et al. 2013b
CCAT51	$(AG)_{12}$	F:ATGATGCAGGGCAAACAGC R:CTGTTAAAGTTCCTCTCTACAGC	McCracken et al. 2013b
US3	$(TTAG)_{17}$	F:CCCTGGGAGCATCAGTTAGA R:AAAAGGTTGTGACCCACTGC	Cardall et al. 2007, McCracken et al. 2013b
US6	$(TCTA)_{15}$	F:AAGTGTGTGCCAAAGCATCA R:GCCTTGTTAAGGGCATATGAA	Cardall et al. 2007, McCracken et al. 2013b
DLU409	$(GATA)_{20}$	F:TGCGATCCTAGAAGGAGTAAAACA R:ATTCCATTTGCTGTCAACTTCAAA	Tranah et al. 2001, McCracken et al. 2013b
DLU439	$(ACAG)7(GATA)_{25}$	F:GAGACAGTCCACACTTCACATTGT R:TTCCATAATACACTCTTGGCATAG	Tranah et al. 2001, McCracken et al. 2013b
DLU4183	$(GATA)_{27}$	F:CTGAAAGCACCTCCTCCATTAG R:GTTCTCTTCTCCTGTTTCGCTTAT	Tranah et al. 2001, McCracken et al. 2013b
DLU4201	$(GATA)_{21}$	F:CCAACCTTCTGAACAACTGTAAAT R:GTGGTAAAGAGGTCTGCCTGTAT	Tranah et al. 2001, McCracken et al. 2013b
DLU4235	$(GATA)_{12}$	F:TGGTATTAACCGTTTACTTCCACA R:TAAACTCCGCTTTTGTTATCAGC	Tranah et al. 2001, McCracken et al. 2013b
DLU4243	$(GATA)_{24}$	F:TGGTTGGATGCTGAAATAAAGTAA R:TGAGCCTCATCATAGATGGATAGA	Tranah et al. 2001, McCracken et al. 2013b
DLU4259	$(GATA)_{24}$	F:GGGTGCAGAAACGTATCCAAAAAC R:AAGCATCATTCAACACCACATTCA	Tranah et al. 2001, McCracken et al. 2013b
DLU4314	$(GATA)_{25}$	F:GAGGGTCTGTGGAGAACA R:TTTCACTTCAATGACAAAAATA	Tranah et al. 2001, McCracken et al. 2013b
DLU4339	$(GATA)_{18}$	F:TGTTCCTCGGTCAGCTCTTCATCA R:GGCCAAAGGGGCAGCACATAC	Tranah et al. 2001, McCracken et al. 2013b
Mohu-Lav268	$(GACA)_5(GGCA)_2GGTATA(TAGA)_{23}$	F: CACAACAGCAGAATTAAGACAGG R: TCACCTTCAATCCATCATCAA	Lippe et al. 2004

APPENDIX C: PCR REACTION REAGENTS

Table C1 PCR reaction reagents for amplification of longnose sucker genetic samples.

PCR reaction
2.3µL RNAse free water
0.5 µL 10X reaction buffer (Bio Basi Inc., Markham, Ontario)
0.5 µL $MgSO_4$ (20mM)
0.5 µL dNTPs (Bio Basi Inc., Markham, Ontario)
0.05 µL forward primer
0.05 µL reverse primer
0.05 µL m13 fluorescent tag (700 nm or 800 nm fluorescence)
0.05 µL TSG Polymerase (Qiagen Inc., United States)
1.0 µL of DNA
5 µL Total Volume

APPENDIX D: THERMOCYCLER PROGRAMS

Table D1 Thermocycler programs for amplification of longnose sucker genetic samples. Markers DLU4235 and DLU4314 were amplified using the 55°C annealing program. The MOHU268 marker was amplified using 59°C annealing program. The DLU4339 marker was amplified using the 60°C annealing program. The DLU439 marker was amplified using either the touchdown program or 55°C annealing program depending on extracted DNA quality. All other markers were amplified using the touchdown program.

	55°C Annealing Thermocycle	**59°C Annealing Thermocycle**	**60°C Annealing Thermocycle**		**Touchdown Thermocycle**
	95°C - 5 min	95°C - 5 min	95°C - 5 min		95°C - 15 min
X 35 Cycles	95°C - 45 s	94°C - 45 s	94°C - 45 s	X 15 Cycles	95°C - 45 s
	55°C - 45 s	59°C - 45 s	60°C - 45 s		65°C - 45 s
	72°C - 45 s	72°C - 45 s	72°C - 45 s		72°C - 45 s
	72°C - 5 min	72°C -5 min	72°C - 5 min	X 18 Cycles	95°C -45 s
					55°C - 45 s
					72°C - 45 s
					72°C - 5 min

APPENDIX E: KOLMOGOROV-SMIRNOV TESTS

Table E1 Kolmogorov-Smirnov test results comparing age class distributions between years in each lake.

Comparison	*P*
WP152 2009 vs. WP152 2010	0.82
WP152 2009 vs. WP152 2011	0.67
WP152 2009 vs. WP152 2013	>0.99
WP152 2010 vs. WP152 2011	>0.99
WP152 2010 vs. WP152 2013	0.67
WP152 2011 vs. WP152 2013	0.52
Slushy 2010 vs. Slushy 2011	0.82
Slushy 2010 vs. Slushy 2012	0.99
Slushy 2011 vs. Slushy 2012	0.27
Genetics H 2009 vs. Genetics H 2011	0.82
Genetics H 2009 vs. Genetics H 2013	0.99
Genetics H 2011 vs. Genetics H 2013	0.52
Lake 1 2010 vs. Lake 1 2011	>0.99
Lake 1 2010 vs. Lake 1 2014	0.82
Lake 1 2011 vs. Lake 1 2014	0.52

APPENDIX F: NULL ALLELES DETECTED USING MICROCHECKER

Table F1 Markers identified as potentially containing null alleles in eight lakes by MICROCHECKER.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T Bone	Cabot
CCAT7								
CCAT20								
CCAT16				✓				
CCAT35								
US6								
DLU409								
CCAT43								
CCAT44								
DLU439		✓						
CCAT32								
DLU4243								
DLU4183		✓	✓	✓	✓	✓	✓	
CCAT51								
US3								
DLU4201								
DLU4235								
DLU4314								
DLU4339								

APPENDIX G: DIVERSITY AND DISTANCE CORRELATIONS

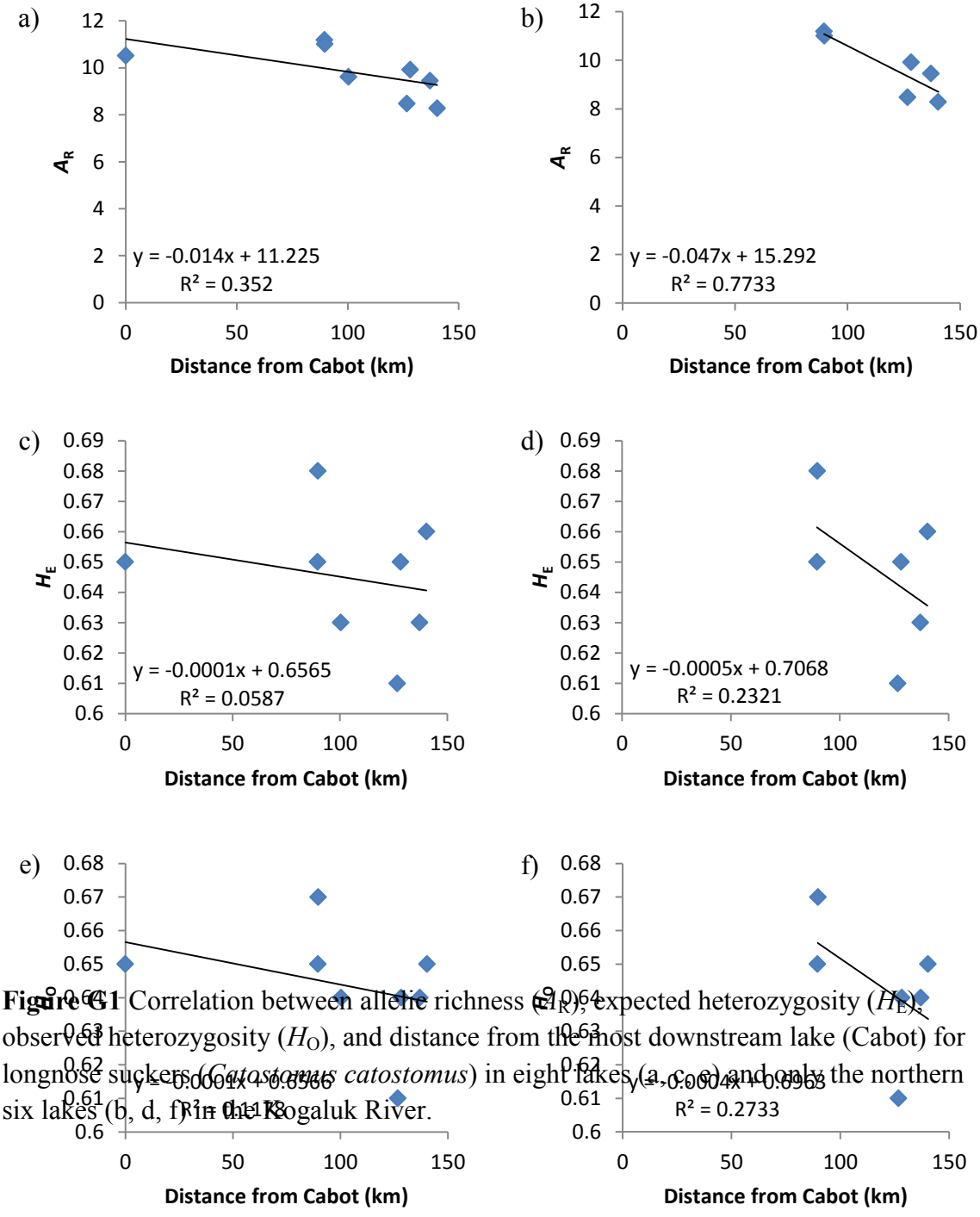

Figure G1 Correlation between allelic richness (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O), and distance from the most downstream lake (Cabot) for longnose suckers (*Catostomus catostomus*) in eight lakes (a, c, e) and only the northern six lakes (b, d, f) in the Kogaluk River.

APPENDIX H: PAIRWISE AND LINEARIZED PAIRWISE $\widehat{F}_{ST}$s

Table H1 Pairwise $\hat{F}_{ST}$ values calculated using 17 loci for longnose suckers (*Catostomus catostomus*) sampled from eight lakes within the Kogaluk River.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T-Bone	Cabot
Lake 1	0							
Genetics H	0.04267	0						
Slushy	0.034639	0.022473	0					
Strange	0.070202	0.055884	0.053757	0				
Esker	0.026051	0.017603	0.016449	0.035134	0			
WP152	0.024922	0.02126	0.018874	0.045632	0.003581	0		
T-Bone	0.062289	0.033197	0.041264	0.067973	0.033804	0.029786	0	
Cabot	0.051459	0.026346	0.029516	0.059827	0.024279	0.023745	0.022593	0

Table H2 Linearized pairwise $\hat{F}_{ST}$ values calculated using 17 loci for longnose suckers (*Catostomus catostomus*) sampled from eight lakes within the Kogaluk River.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T-Bone	Cabot
Lake 1	0							
Genetics H	0.044572	0						
Slushy	0.035882	0.02299	0					
Strange	0.075502	0.059192	0.056811	0				
Esker	0.026748	0.017918	0.016724	0.036413	0			
WP152	0.025559	0.021722	0.019237	0.047814	0.003594	0		
T-Bone	0.066427	0.034337	0.04304	0.07293	0.034987	0.0307	0	
Cabot	0.054251	0.027059	0.030414	0.063634	0.024883	0.024323	0.023115	0

APPENDIX I: STRUCTURE HARVESTER PLOTS

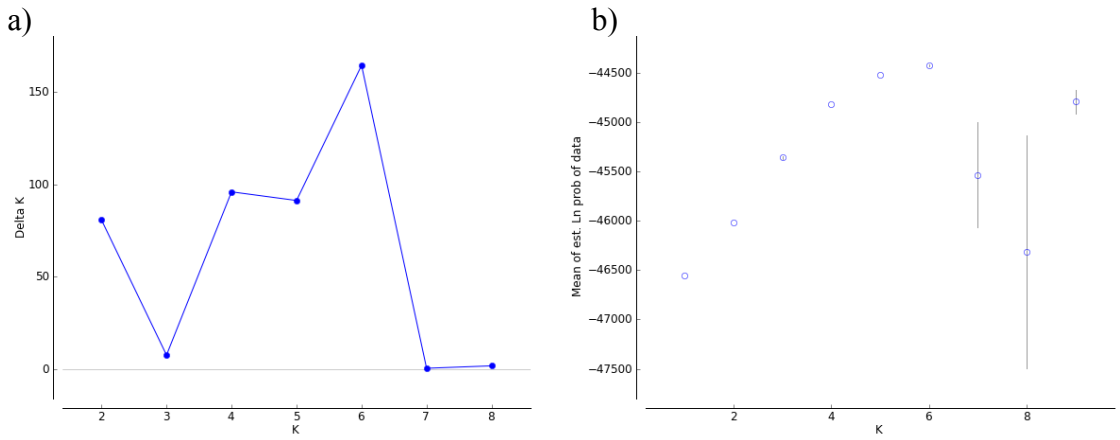

Figure I1 a) Delta *K*-values versus *K*-values for STRUCTURE analyses based on genotypes of 17 loci from longnose sucker samples from all lakes. b) Ln(*K*) values versus *K*-values for STRUCTURE analyses including longnose sucker samples from all lakes.

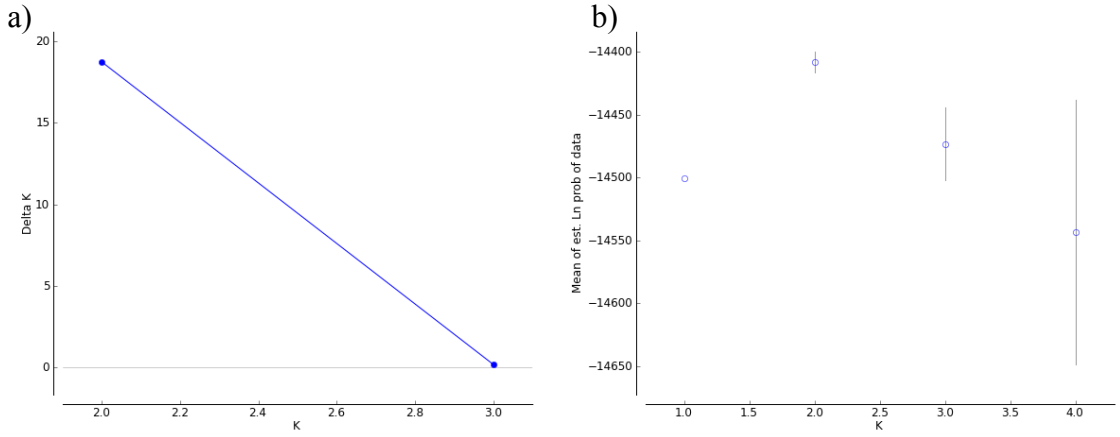

Figure I2 a) Delta *K*-values versus *K*-values and b) Ln(*K*) values versus *K*-values for STRUCTURE analyses including location priors based on genotypes of 17 loci from longnose sucker samples from Esker, WP152 and Cabot only.

APPENDIX J: PAIRWISE DISTANCES, ELEVATIONS, AND SLOPES

Table J1 Pairwise waterway distances in kilometres between eight lakes within the Kogaluk River.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T-Bone	Cabot
Lake 1	0							
Genetics H	34.25	0						
Slushy	75.55	72.29	0					
Strange	74.040	70.770	35.150	0				
Esker	50.75	47.49	38.67	37.150	0			
WP152	85.61	82.35	73.53	72.010	34.86	0		
T-Bone	160.42	157.16	148.34	146.820	109.67	109.84	0	
Cabot	140.32	137.06	128.24	126.720	89.57	89.74	100.35	0

Table J2 Pairwise elevation in metres between eight lakes within the Kogaluk River.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T-Bone	Cabot
Lake 1	0							
Genetics H	13	0						
Slushy	61	48	0					
Strange	38	25	23	0				
Esker	94	81	33	56	0			
WP152	80	67	19	42	14	0		
T-Bone	57	44	4	19	37	23	0	
Cabot	465	452	404	427	371	385	408	0

Table J3 Pairwise slope in metres per kilometre between eight lakes within the Kogaluk River.

	Lake 1	Genetics H	Slushy	Strange	Esker	WP152	T-Bone	Cabot
Lake 1	0							
Genetics H	0.379562	0						
Slushy	0.807412	0.663992	0					
Strange	0.513236	0.353257	0.654339	0				
Esker	1.852217	1.705622	0.853375	1.507402	0			
WP152	0.93447	0.8136	0.258398	0.583252	0.401606	0		
T-Bone	0.355317	0.279969	0.026965	0.12941	0.337376	0.209395	0	
Cabot	3.313854	3.297826	3.150343	3.369634	4.142012	4.290172	4.06577	0

APPENDIX K: PAIRWISE DISTANCE MANTEL TESTS

Table K1 Results of Mantel test correlating pairwise waterway distances and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from eight lakes within the Kogaluk River. Significant *P*-values are in bold (α = 0.05).

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	0.13	0.122	-71.51	16.22
Strange	0.40	**0.007**	-76.77	10.96
Strange, Lake 1	0.77	**0.003**	-87.73	0.00

Table K2 Results of Mantel test correlating pairwise waterway distances and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from the northern six lakes within the Kogaluk River. Significant *P*-values are in bold (α = 0.05).

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	0.02	0.397	-61.49	21.17
Strange	0.02	0.394	-68.29	14.36
Strange, Lake 1	0.58	**0.043**	-82.66	0.00

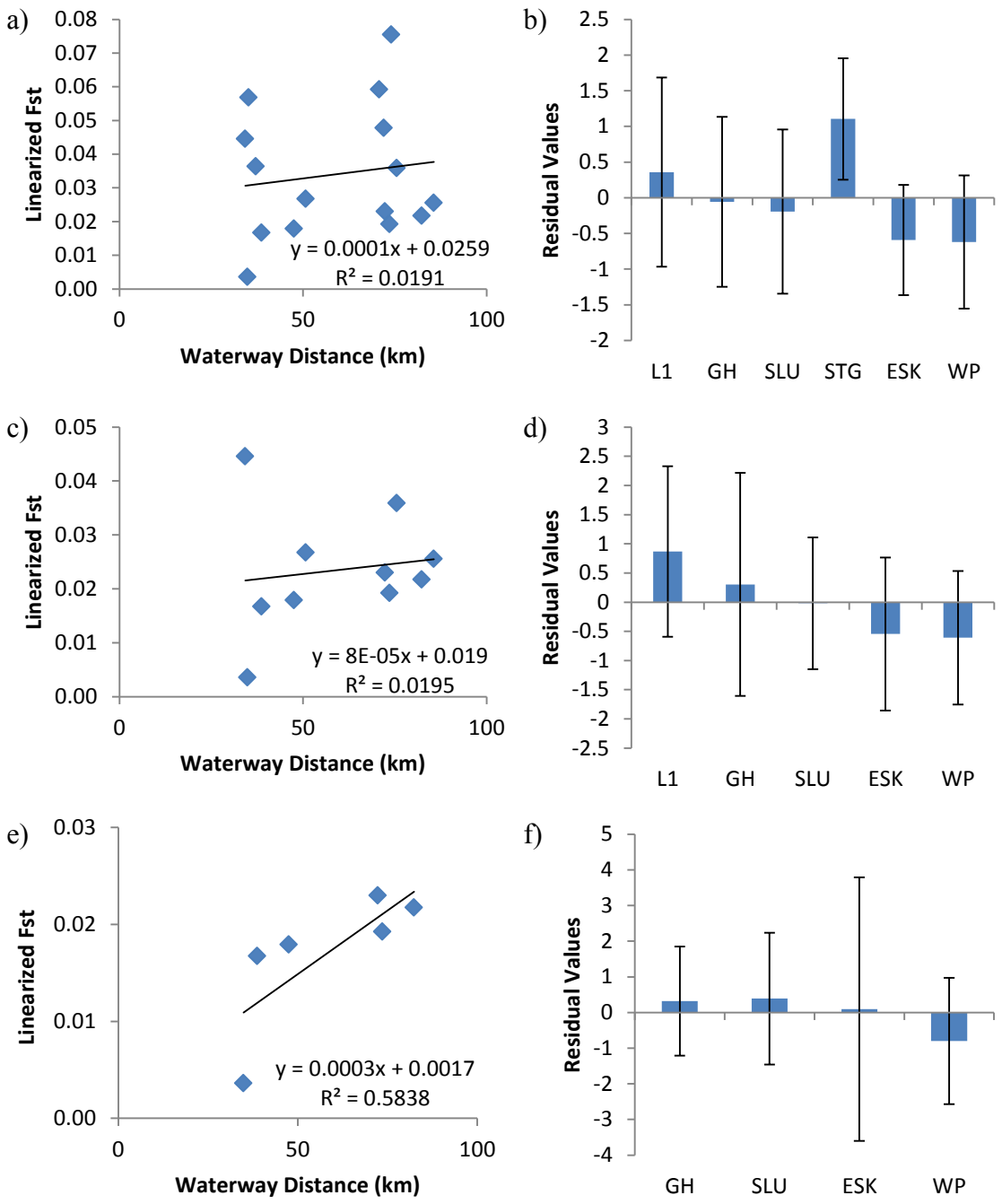

Figure K2 Correlation between pairwise linearized $\hat{F}_{ST}$ values and waterway distance (km) between samples of longnose suckers (*Catostomus catostomus*) collected from the northern six lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) with no further outlier lakes identified in d) a plot of the residuals. Based on the genetic distinctiveness of Lake 1 it was removed and the correlation was repeated e) and f) a plot of the residuals revealed no further outlier lakes.

APPENDIX L: PAIRWISE ELEVATION MANTEL TESTS

Table L1 Results of Mantel test correlating elevation differences and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from eight lakes within the Kogaluk River.

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	<0.01	0.568	-70.11	1.68
Strange	<0.01	0.308	-71.79	0.00

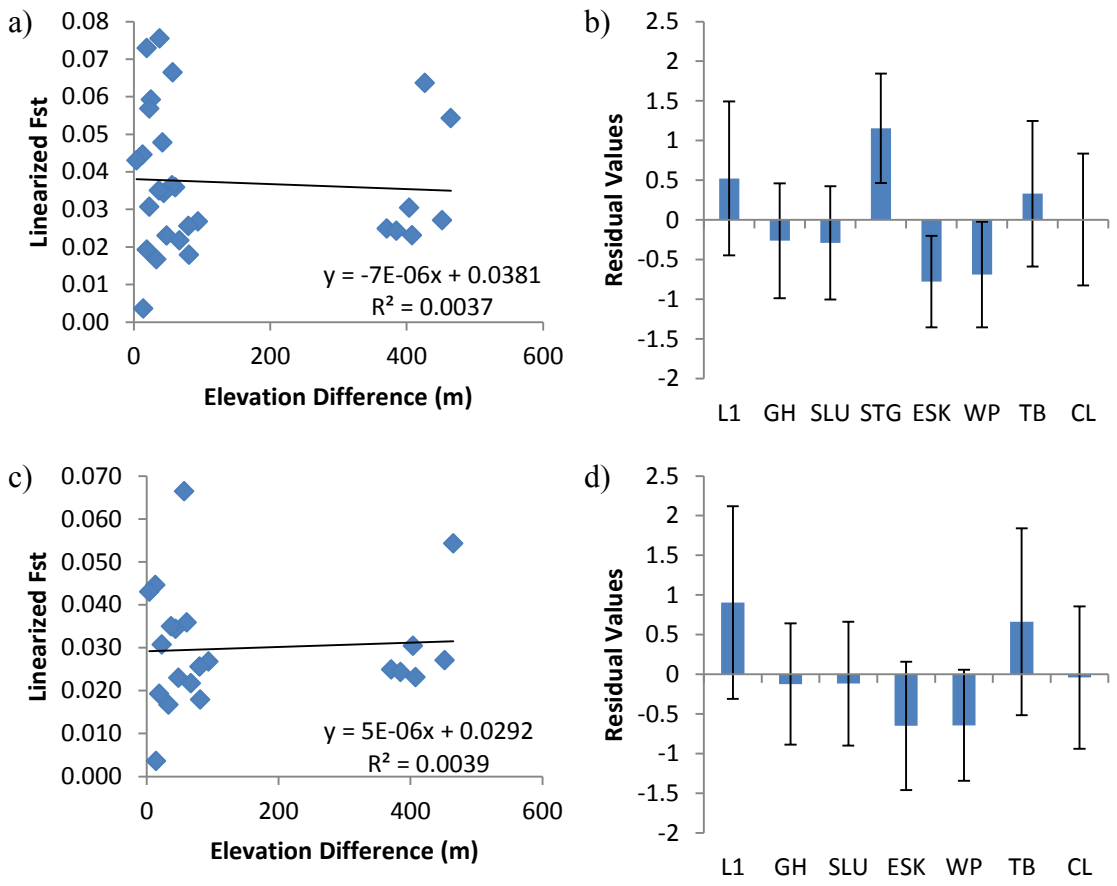

Figure L1 Correlation between pairwise linearized $\hat{F}_{ST}$ values and elevation (m) between samples of longnose suckers (*Catostomus catostomus*) collected from eight lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) and d) a plot of the residuals revealed no further outlier lakes.

Table L2 Results of Mantel test correlating elevation differences and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from the northern six lakes within the Kogaluk River.

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	0.05	0.180	-61.79	6.43
Strange	0.01	0.311	-68.23	0.00

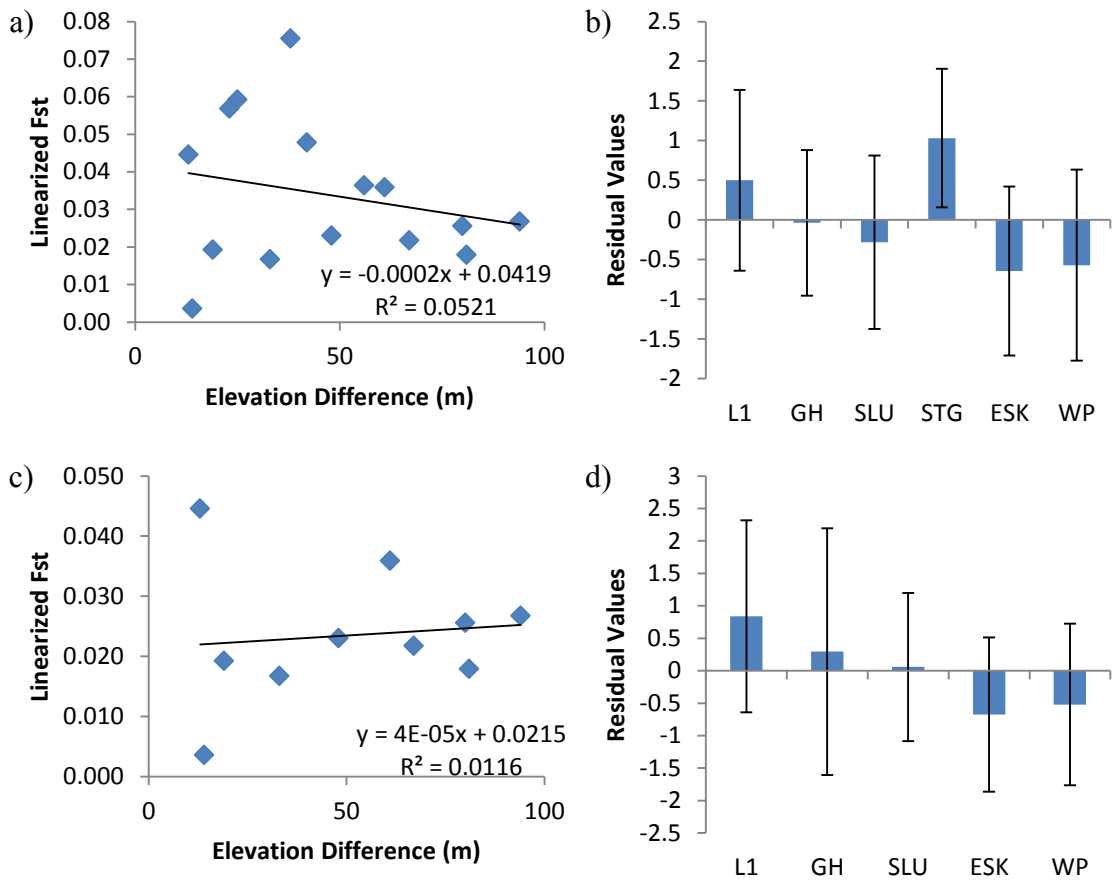

Figure L2 Correlation between pairwise linearized $\hat{F}_{ST}$ values and elevation (m) between samples of longnose suckers (*Catostomus catostomus*) collected from the northern six lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) and d) a plot of the residuals revealed no further outlier lakes.

APPENDIX M: PAIRWISE SLOPE MANTEL TESTS

Table M1 Results of Mantel test correlating slope and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from eight lakes within the Kogaluk River.

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	0.04	0.341	-70.45	1.39
Strange	<0.01	0.437	-71.84	0.00

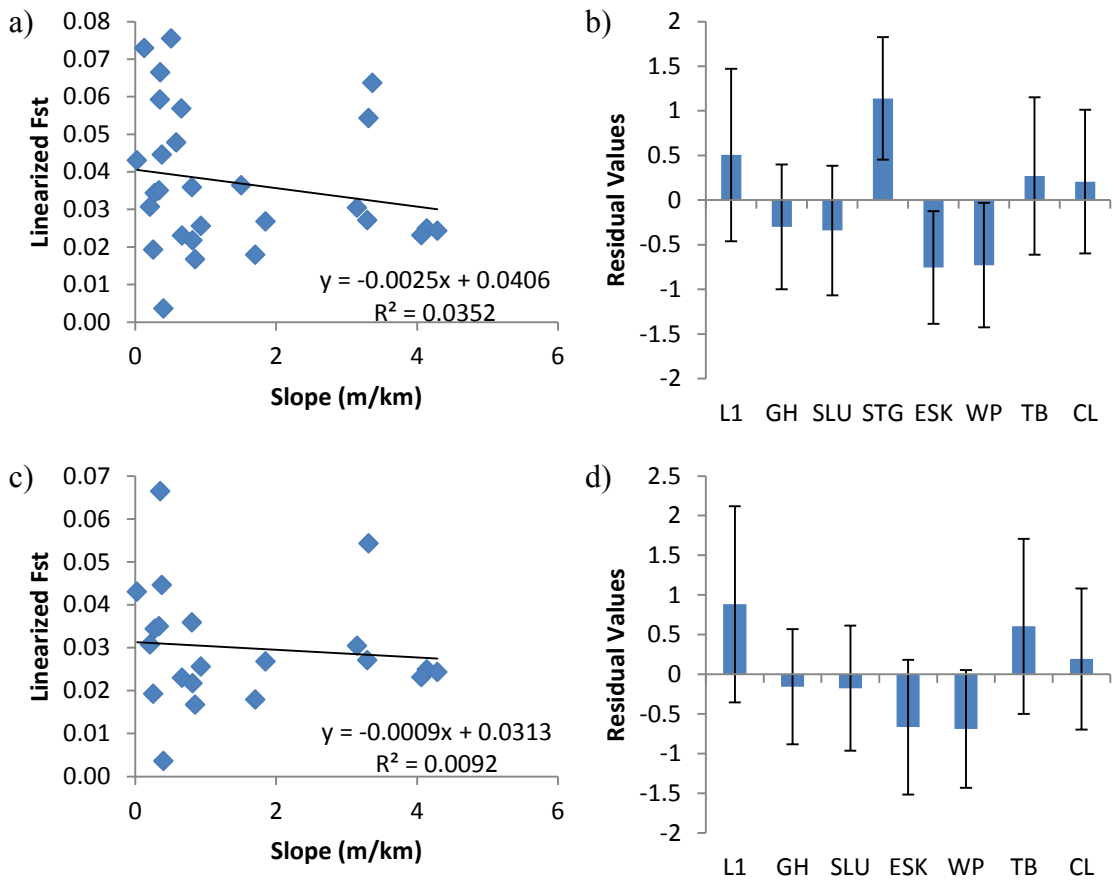

Figure M1 Correlation between pairwise linearized $\hat{F}_{ST}$ values and slope (m/km) between samples of longnose suckers (*Catostomus catostomus*) collected from eight lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) and d) a plot of the residuals revealed no further outlier lakes.

Table M2 Results of Mantel test correlating slope and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from the northern six lakes within the Kogaluk River.

Lakes excluded	R^2 value	P	AIC_c	ΔAIC_c
None	0.07	0.239	-61.94	6.19
Strange	<0.01	0.558	-68.13	0.00

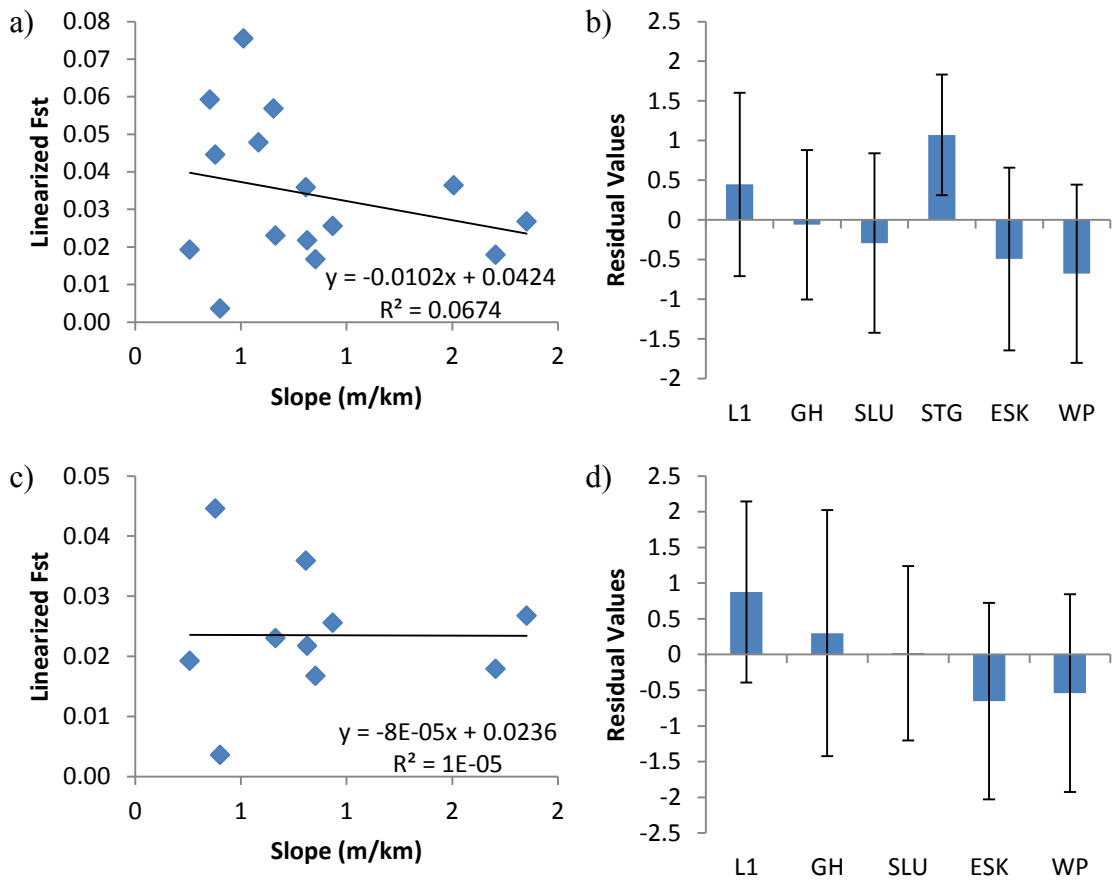

Figure M2 Correlation between pairwise linearized $\hat{F}_{ST}$ values and the slope (m/km) between samples of longnose suckers (*Catostomus catostomus*) collected from the northern six lakes within the Kogaluk River. a) Initially all lakes were included in the correlation and b) a barplot of the residuals for each lake indicates Strange lake as a potential outlier lake due to the lack of overlap between this lake's 95% confidence interval (brackets) with 0. This lake was subsequently removed and the correlation repeated c) and d) a plot of the residuals revealed no further outlier lakes.

APPENDIX N: INTERMEDIATE WATERFALLS MANTEL TESTS

Table N1 Results of Mantel test correlating number of intermediate waterfalls and pairwise linearized $\hat{F}_{ST}$s for longnose sucker (*Catostomus catostomus*) samples from the northern six lakes within the Kogaluk River. Significant *P*-values are in bold (α = 0.05).

Lakes excluded	R^2 value	*P*	AIC_c	ΔAIC_c
None	0.07	0.199	-70.81	14.72
Strange	0.30	0.061	-75.19	10.34
Strange, Lake 1	0.70	**0.027**	-85.52	0.00
Strange, Lake 1, Slushy	0.73	0.050	-82.23	3.29

APPENDIX O: ALLELIC RICHNESS AND $\widehat{F}_{ST}$ CORRELATIONS

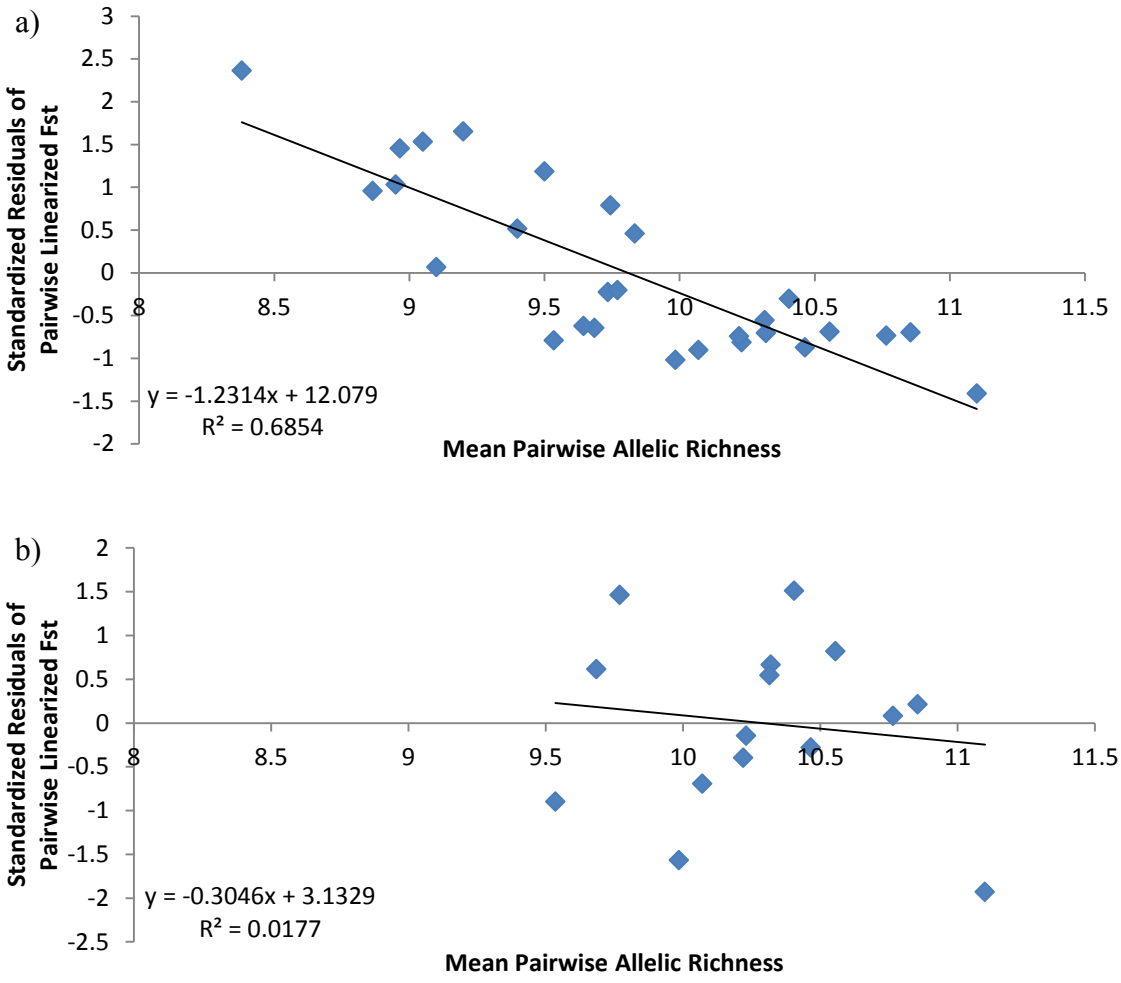

Figure O1 Correlation of the standardized residuals of pairwise linearized $\widehat{F}_{ST}$s after correlation between linearized $\widehat{F}_{ST}$ and geographic distance with allelic richness for a) all lakes, b) all lakes except Strange and Lake 1.

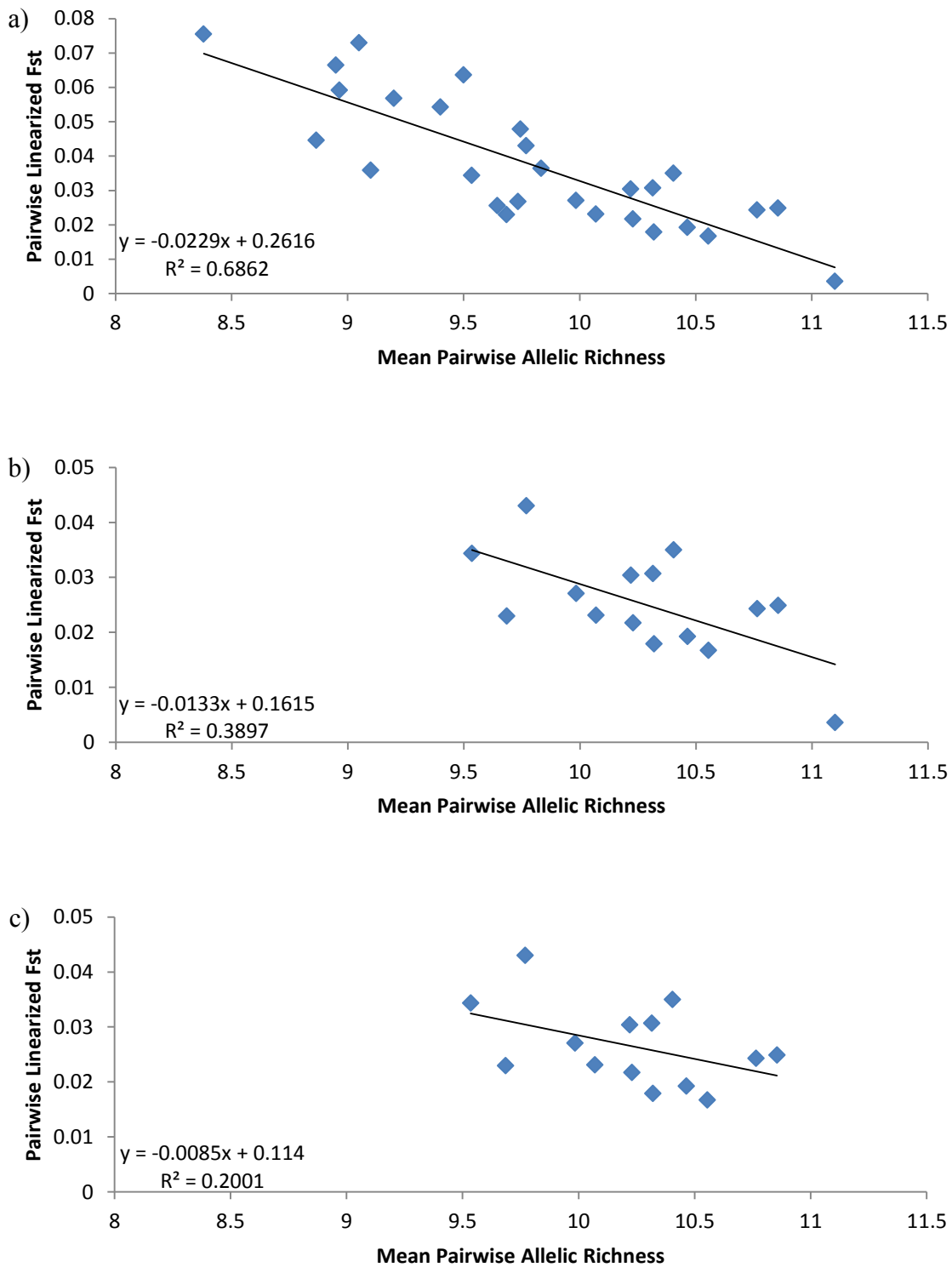

Figure O2 Correlation of pairwise linearized $\hat{F}_{ST}$s with allelic richness for a) all lakes, b) all lakes except Strange and Lake 1, c) all lakes except Strange and Lake 1 with the exclusion of the data point for Esker and WP152.

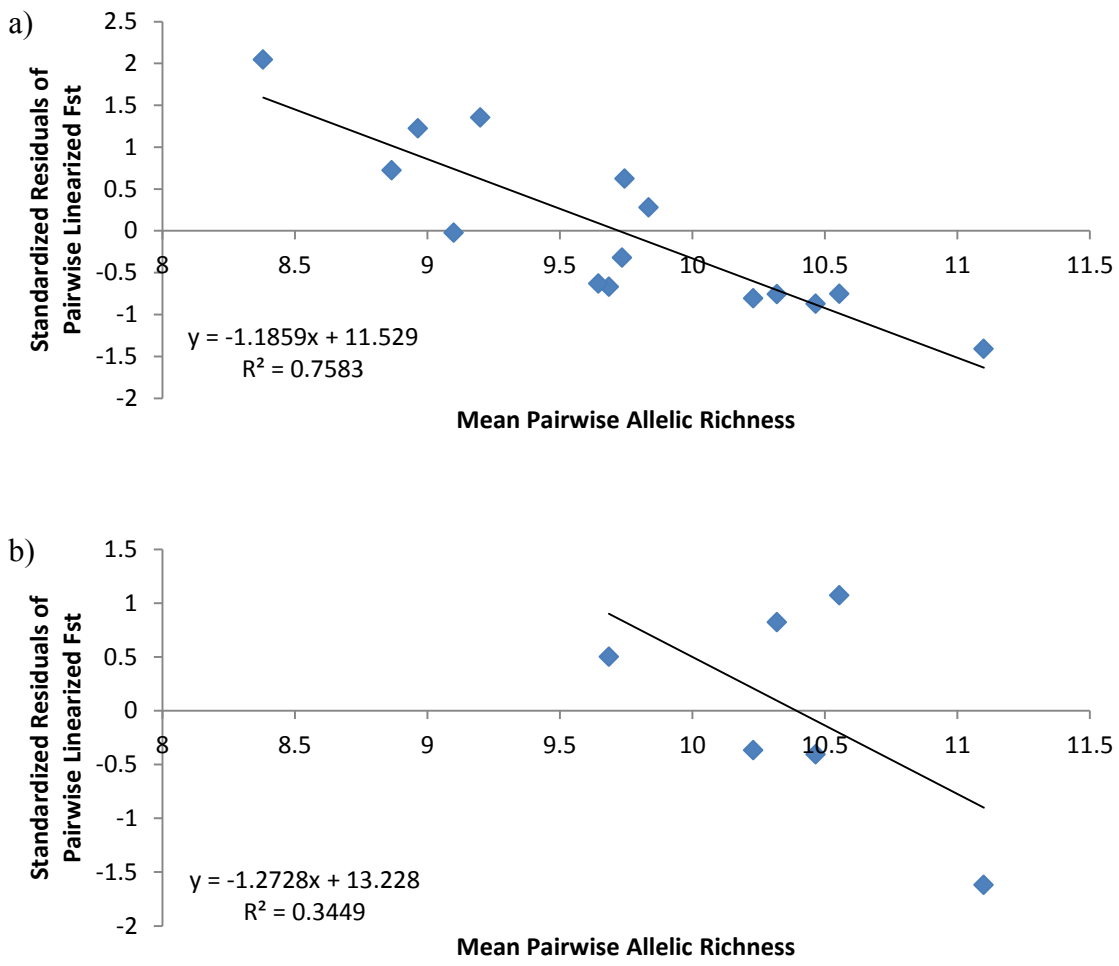

Figure O3 Correlation of the standardized residuals of pairwise linearized $\hat{F}_{ST}$s after correlation between linearized $\hat{F}_{ST}$ and geographic distance with allelic richness for a) all of the northern six lakes, b) all of the northern six lakes except Strange and Lake 1.

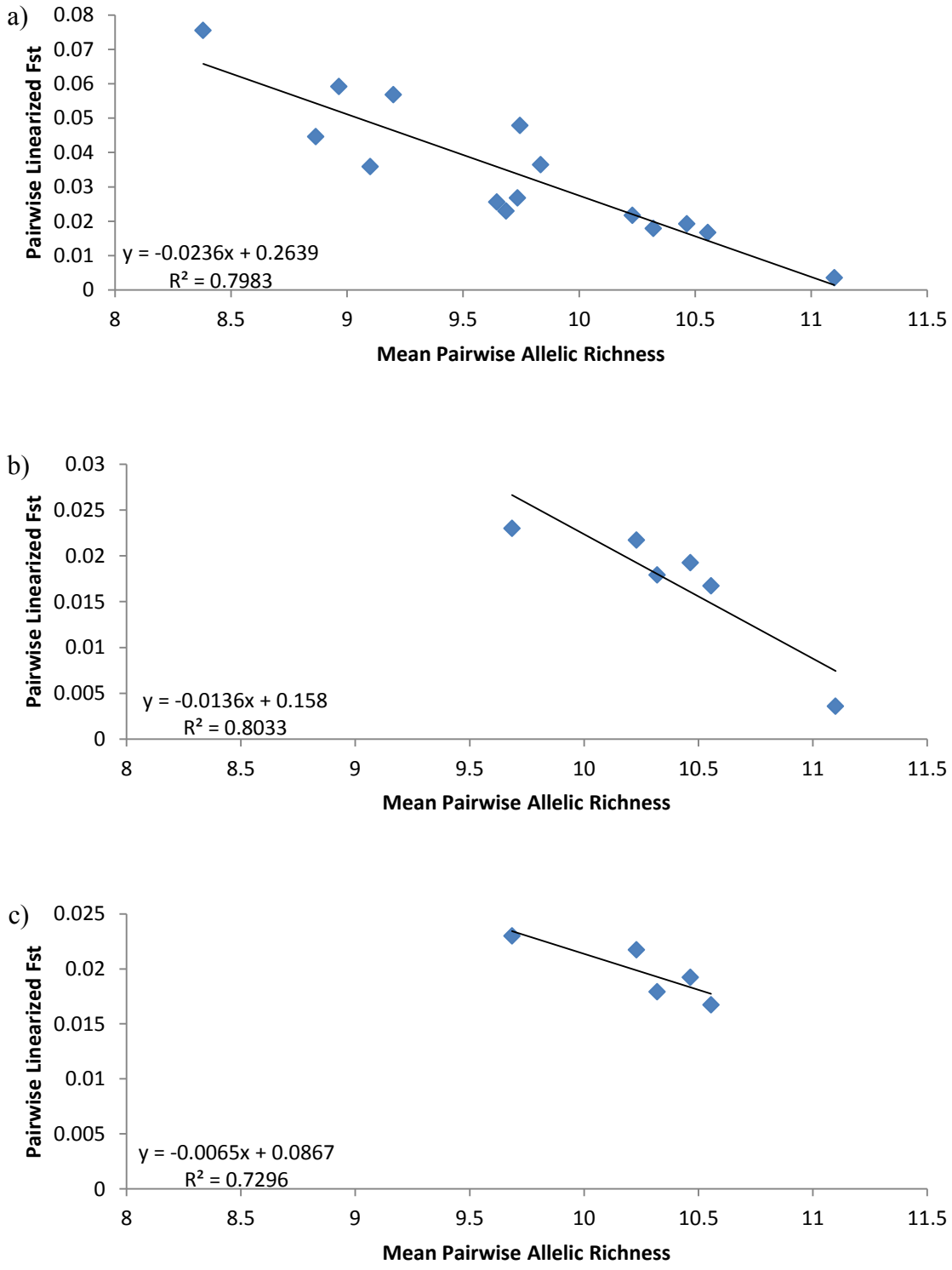

Figure O4 Correlation of pairwise linearized $\hat{F}_{ST}$s with allelic richness for a) all of the northern six lakes, b) all of the northern six lakes except Strange and Lake 1, c) all of the northern six lakes except Strange and Lake 1 with the exclusion of the data point for Esker and WP152.

APPENDIX P: PUTATIVE ORIGINS OF MIGRANTS

Table P1 Putative origin of longnose sucker samples collected in eight lakes according to GENECLASS2.

	Putative Origin							
Sampling Location	**Lake 1**	**Genetics H**	**Slushy**	**Strange**	**Esker**	**WP152**	**T-Bone**	**Cabot**
Lake 1	55				2	2		
Genetics H		199				1	1	
Slushy		1	98			2		2
Strange	1		1	117	1	1		1
Esker					138			
WP152	1				2	71		
T-Bone					1		114	
Cabot					2		1	54

Table Q1 Migration rates between lakes within the Kogaluk River calculated using BayesAss for the purpose of estimating the effective metapopulation size of the entire system using the Tufto-Hindar method (Tufto and Hindar 2003). Esker and WP152 samples were pooled together. Standard deviations are in parentheses.

		Emigration From						
		Lake 1	**Genetics H**	**Slushy**	**Strange**	**Esker/WP152**	**T-Bone**	**Cabot**
Immigration Into	**Lake 1**	0.9558 (0.0159)	0.0061 (0.006)	0.0067 (0.0065)	0.0072 (0.0067)	0.009 (0.0086)	0.01 (0.0085)	0.0052 (0.0052)
	Genetics H	0.0051 (0.0041)	0.9141 (0.0185)	0.012 (0.0071)	0.0071 (0.0048)	0.0461 (0.0161)	0.0138 (0.0084)	0.0019 (0.0019)
	Slushy	0.0082 (0.0066)	0.0079 (0.0074)	0.8751 (0.0229)	0.0076 (0.0065)	0.0848 (0.0219)	0.0126 (0.0107)	0.0038 (0.0038)
	Strange	0.0108 (0.0066)	0.0032 (0.0032)	0.0045 (0.0041)	0.9447 (0.0137)	0.0287 (0.0128)	0.0052 (0.0048)	0.0028 (0.0028)
	Esker/WP152	0.0136 (0.0089)	0.0161 (0.0097)	0.0188 (0.0103)	0.0111 (0.0069)	0.9315 (0.0177)	0.0058 (0.0053)	0.003 (0.0028)
	T-Bone	0.0056 (0.0052)	0.0057 (0.0054)	0.0049 (0.0044	0.0034 (0.0033)	0.0094 (0.0079)	0.9679 (0.0119)	0.003 (0.003)
	Cabot	0.0068 (0.0066)	0.0077 (0.0074)	0.0072 (0.007)	0.0092 (0.0083)	0.2432 (0.032)	0.0421 (0.0272)	0.6837 (0.0127)

Table Q2 Migration rates between lakes within the Kogaluk River calculated using BayesAss for the purpose of estimating the effective metapopulation size of the entire system (excluding T-Bone) using the Tufto-Hindar method (Tufto and Hindar 2003). Esker and WP152 samples were pooled together. Standard deviations are in parentheses.

		Emigration From					
		Lake 1	**Genetics H**	**Slushy**	**Strange**	**Esker/WP152**	**Cabot**
Immigration Into	**Lake 1**	0.9635 (0.0149)	0.0063 (0.0062)	0.0074 (0.0069)	0.0074 (0.0069)	0.01 (0.0092)	0.0054 (0.0053)
	Genetics H	0.0056 (0.0043)	0.925 (0.0156)	0.0145 (0.0078)	0.0077 (0.0051)	0.045 (0.0146)	0.0021 (0.002)
	Slushy	0.0087 (0.0069)	0.0089 (0.0079)	0.8885 (0.0221)	0.0092 (0.0075)	0.0807 (0.0219)	0.004 (0.0039)
	Strange	0.0113 (0.0066)	0.0033 (0.0032)	0.0046 (0.0041)	0.9504 (0.0132)	0.0274 (0.0123)	0.003 (0.0029)
	Esker/WP152	0.015 (0.0101)	0.0146 (0.0088)	0.0228 (0.0119)	0.0126 (0.0073)	0.932 (0.0183)	0.0031 (0.0029)
	Cabot	0.0074 (0.0072)	0.009 (0.0087)	0.0076 (0.0074)	0.0108 (0.0093)	0.2766 (0.0227)	0.6886 (0.0156)

Table Q3 Migration rates between the northern six lakes in the Kogaluk River calculated using BayesAss for the purpose of estimating the effective metapopulation size of the northern six lakes using the Tufto-Hindar method (Tufto and Hindar 2003). Standard deviations are in parentheses.

		Emigration From				
		Lake 1	**Genetics H**	**Slushy**	**Strange**	**Esker/WP152**
Immigration Into	**Lake 1**	0.9686 (0.0142)	0.0064 (0.0063)	0.008 0.0074)	0.0075 (0.007)	0.0095 (0.0088)
	Genetics H	0.0055 (0.0042)	0.9253 (0.0168)	0.0152 (0.0082)	0.0074 (0.0049)	0.0467 (0.0158)
	Slushy	0.009 (0.0071)	0.01 (0.0087)	0.8969 (0.0212)	0.0076 (0.0068)	0.0765 (0.021)
	Strange	0.0122 (0.0071)	0.0033 (0.0033)	0.0051 (0.0044)	0.9526 (0.0132)	0.0268 (0.0127)
	Esker/WP152	0.015 (0.0099)	0.0209 (0.0111)	0.0306 (0.0144)	0.01 (0.0065)	0.9234 (0.0219)

APPENDIX R: LAKE AREA AND $\widehat{N}_e$ CORRELATIONS

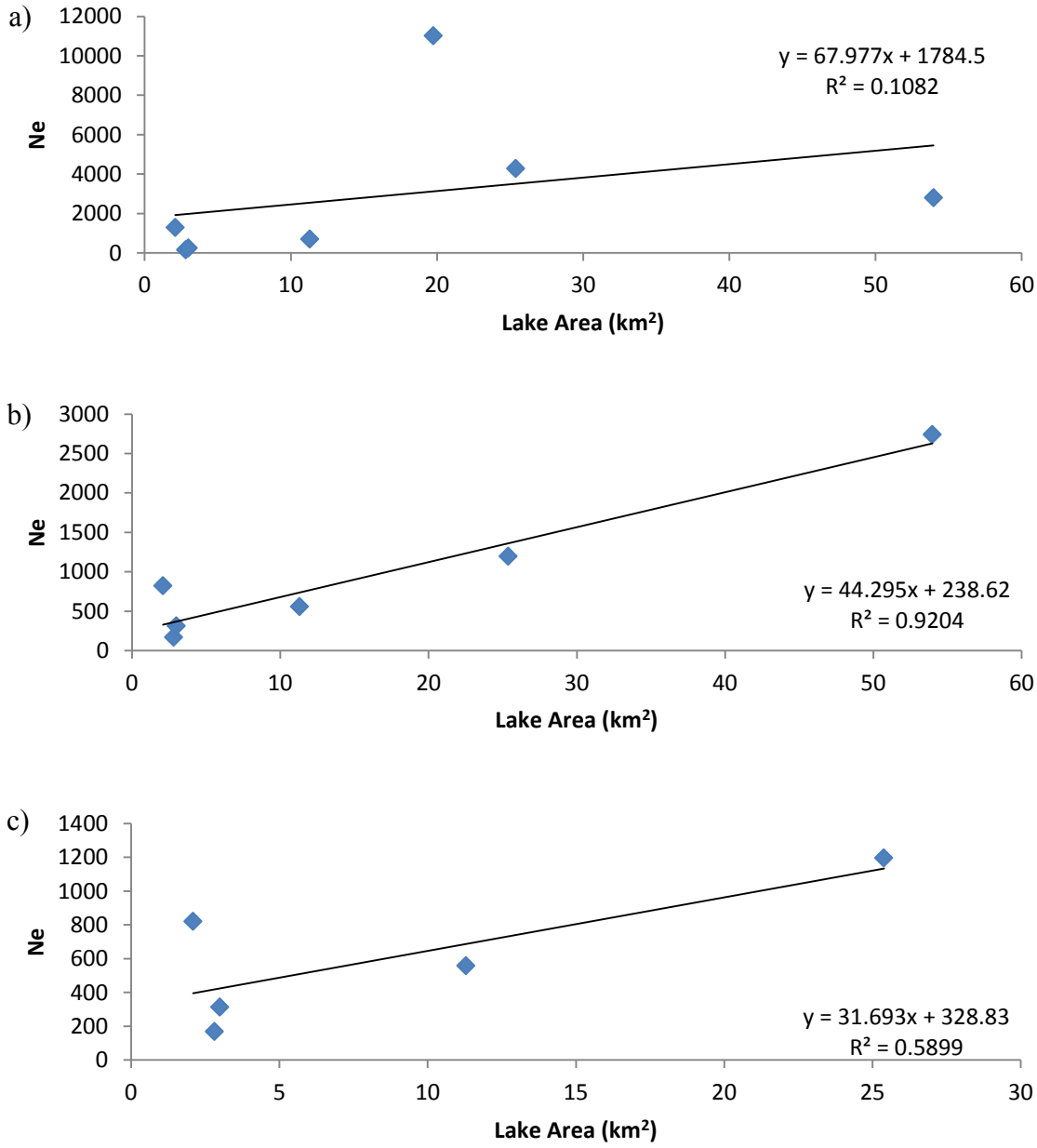

Figure R1 Correlation between $\widehat{N}_e$ values calculated with and without migrants identified in GENECLASS2 with lake area (km^2). a) $\widehat{N}_{e(LDNE)}$ values calculated without migrants correlated with lake area. b) $\widehat{N}_{e(LDNE)}$ values calculated with the inclusion of migrants correlated with lake area after the negative $\widehat{N}_{e(LDNE)}$ value for T-Bone was removed from the correlation analysis. c) Same as b) but with the $\widehat{N}_{e(LDNE)}$ for Esker/WP152 removed.

APPENDIX S: ELEVATION AND $\widehat{N}_e$ CORRELATIONS

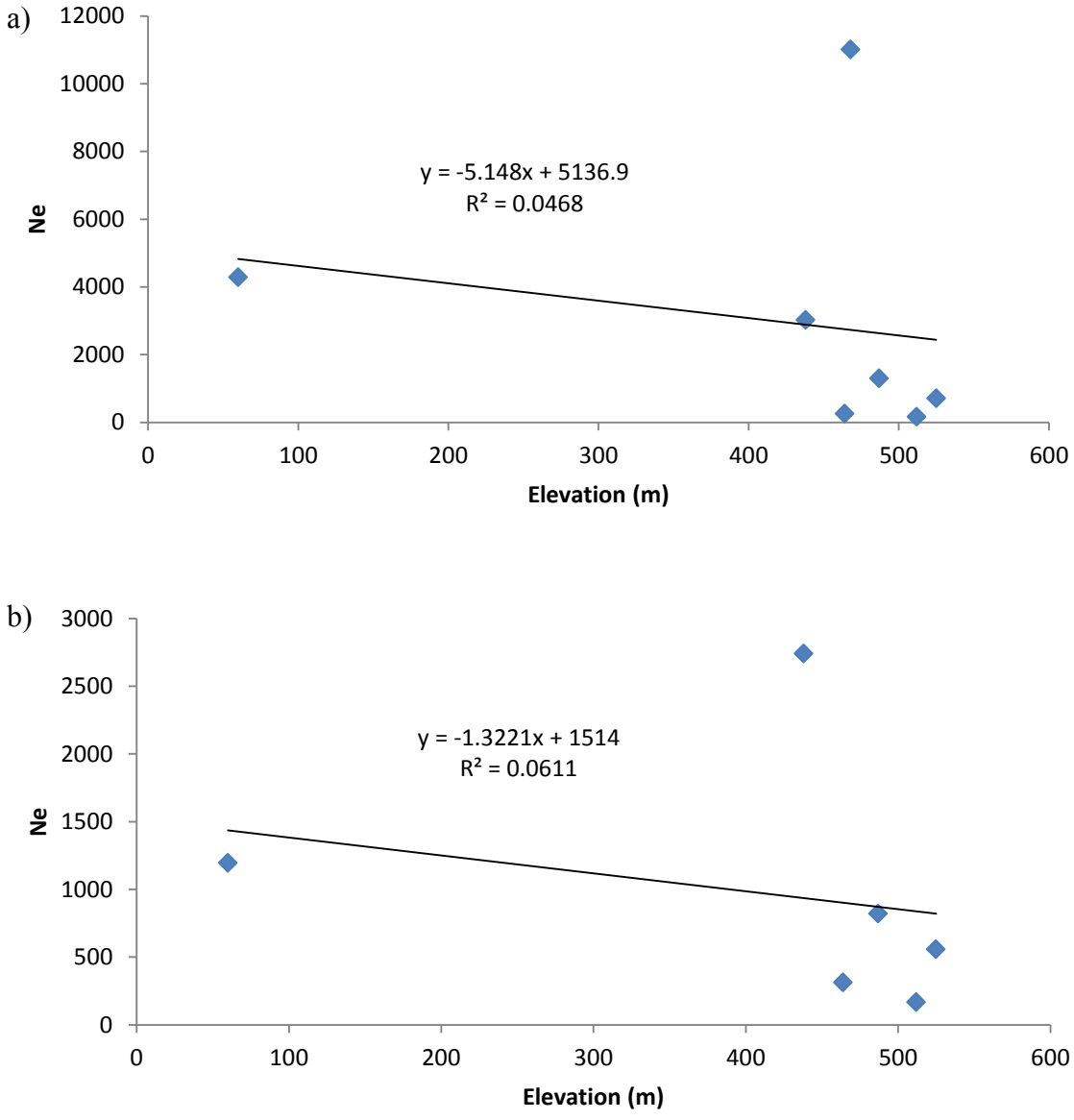

Figure S1 Correlation between $\widehat{N}_{e(LDNE)}$ values calculated with and without migrants identified in GENECLASS2 with lake elevation above sea level (m). a) $\widehat{N}_{e(LDNE)}$ values calculated without migrants. b) $\widehat{N}_{e(LDNE)}$ values calculated with the inclusion of migrants correlated with elevation after the negative $\widehat{N}_{e(LDNE)}$ value for T-Bone was removed from the correlation analysis.

APPENDIX T: DIYABC PRE-EVALUATION SUMMARY STATISTICS

Table T1 Pre-evaluation of three colonization scenarios of longnose suckers (*Catostomus catostomus*) in the Kogaluk River based on comparison of summary statistics of the observed data set with those for the simulated data sets. Values for each scenario denote the proportion of data sets that have a summary statistic value less than that of the observed value: * denotes a value >0.95 or <0.05. NAL is the mean number of alleles (one-sample), HET is the mean genic diversity (one-sample), VAR is the mean size variance (one-sample), H2P is the mean genic diversity (two-sample), V2P is the mean size variance (two-sample), DM2 is the $(d\mu)^2$ distance (two sample). Within the summary statistics column 1 is Lake 1, 2 is Genetics H, 3 is Slushy, 4 is Strange, 5 is Esker/WP152, 6 is T-Bone, 7 is Cabot.

Summary Statistics	Observed Value	Scenario 1	Significance	Scenario 2	Significance	Scenario 3	Significance
NAL_1_1	8.5294	0.1886		0.1843		0.1856	
NAL_1_2	11.4706	0.309		0.3007		0.3058	
NAL_1_3	11.2941	0.3323		0.3233		0.3261	
NAL_1_4	9.8824	0.235		0.2292		0.2315	
NAL_1_5	14.1765	0.4626		0.4223		0.461	
NAL_1_6	11.0588	0.2984		0.3017		0.2781	
NAL_1_7	10.7059	0.3018		0.3206		0.3143	
HET_1_1	0.6517	0.0679		0.0674		0.0665	
HET_1_2	0.642	0.063		0.0627		0.0614	
HET_1_3	0.645	0.0647		0.0641		0.0634	
HET_1_4	0.6109	0.0498	*	0.0497	*	0.0495	*
HET_1_5	0.6573	0.0702		0.0696		0.0687	
HET_1_6	0.6356	0.06		0.0597		0.06	
HET_1_7	0.6475	0.065		0.0657		0.0605	
VAR_1_1	11.0212	0.3218		0.3194		0.3204	
VAR_1_2	14.8178	0.4188		0.4153		0.4172	
VAR_1_3	12.9595	0.3724		0.3696		0.3704	
VAR_1_4	10.827	0.3163		0.3141		0.3148	
VAR_1_5	14.7701	0.4164		0.406		0.415	
VAR_1_6	13.3952	0.3791		0.3803		0.3721	
VAR_1_7	14.4487	0.3991		0.4032		0.403	
H2P_1_1&2	0.6547	0.0627		0.0611		0.0621	
H2P_1_1&3	0.6584	0.0618		0.0595		0.0581	
H2P_1_1&4	0.6453	0.0559		0.054		0.0528	
H2P_1_1&5	0.6617	0.0678		0.0651		0.0671	
H2P_1_1&6	0.6603	0.0591		0.061		0.0599	
H2P_1_1&7	0.6669	0.0615		0.0628		0.0615	
H2P_1_2&3	0.6498	0.0581		0.0562		0.0541	
H2P_1_2&4	0.6479	0.0566		0.0545		0.0527	
H2P_1_2&5	0.6563	0.0621		0.0587		0.0618	
H2P_1_2&6	0.6498	0.0534		0.0556		0.0541	

Summary Statistics	Observed Value	Scenario 1	Significance	Scenario 2	Significance	Scenario 3	Significance
H2P_1_2&7	0.6489	0.0579		0.0589		0.0577	
H2P_1_3&4	0.6448	0.054		0.0514		0.0485	*
H2P_1_3&5	0.6586	0.0645		0.0615		0.0588	
H2P_1_3&6	0.6541	0.0536		0.0556		0.0553	
H2P_1_3&7	0.6548	0.0575		0.0588		0.0561	
H2P_1_4&5	0.6519	0.0606		0.0579		0.0548	
H2P_1_4&6	0.6457	0.0497	*	0.0517		0.0514	
H2P_1_4&7	0.6394	0.0512		0.0523		0.0503	
H2P_1_5&6	0.6594	0.0585		0.0615		0.059	
H2P_1_5&7	0.6605	0.0637		0.0677		0.0642	
H2P_1_6&7	0.6456	0.0576		0.055		0.0563	
V2P_1_1&2	14.1647	0.3981		0.3939		0.3973	
V2P_1_1&3	12.4615	0.3535		0.3497		0.3501	
V2P_1_1&4	11.1994	0.3208		0.3172		0.3175	
V2P_1_1&5	14.0422	0.3956		0.3868		0.3946	
V2P_1_1&6	13.1125	0.3655		0.3663		0.3628	
V2P_1_1&7	12.8525	0.358		0.3586		0.3599	
V2P_1_2&3	14.529	0.4059		0.4012		0.4022	
V2P_1_2&4	13.6992	0.385		0.3806		0.3813	
V2P_1_2&5	14.9917	0.4178		0.4098		0.4168	
V2P_1_2&6	15.3552	0.4216		0.4209		0.4199	
V2P_1_2&7	15.0024	0.416		0.4144		0.4164	
V2P_1_3&4	12.0381	0.3423		0.3385		0.3373	
V2P_1_3&5	14.3403	0.4022		0.3939		0.3973	
V2P_1_3&6	13.5716	0.3768		0.3771		0.3745	
V2P_1_3&7	13.505	0.3759		0.3758		0.3757	
V2P_1_4&5	13.4659	0.3802		0.3723		0.3753	
V2P_1_4&6	12.6578	0.3537		0.3542		0.3518	
V2P_1_4&7	12.2184	0.3436		0.3438		0.3433	
V2P_1_5&6	14.9199	0.411		0.4078		0.4094	
V2P_1_5&7	14.7772	0.41		0.4054		0.4108	
V2P_1_6&7	13.8887	0.3855		0.3856		0.3824	
DM2_1_1&2	1.6435	0.8152		0.7914		0.8493	
DM2_1_1&3	1.2967	0.7362		0.7085		0.6672	
DM2_1_1&4	1.5674	0.7943		0.7689		0.7321	
DM2_1_1&5	0.8552	0.7114		0.6467		0.7496	
DM2_1_1&6	2.5939	0.8553		0.8773		0.9008	
DM2_1_1&7	1.1016	0.5958		0.6385		0.6533	
DM2_1_2&3	1.7107	0.8409		0.8185		0.7875	
DM2_1_2&4	1.8693	0.8613		0.84		0.8123	
DM2_1_2&5	0.7292	0.7329		0.6616		0.7759	

Summary Statistics	Observed Value	Scenario 1	Significance	Scenario 2	Significance	Scenario 3	Significance
DM2_1_2&6	4.7639	0.945		0.9498		0.9683	*
DM2_1_2&7	1.6887	0.8054		0.8199		0.8288	
DM2_1_3&4	1.1052	0.7223		0.6928		0.6331	
DM2_1_3&5	0.793	0.7279		0.6603		0.5707	
DM2_1_3&6	1.6825	0.763		0.8047		0.8367	
DM2_1_3&7	0.6889	0.4292		0.4938		0.4638	
DM2_1_4&5	0.7525	0.7223		0.6506		0.5607	
DM2_1_4&6	2.3546	0.8502		0.8734		0.9082	
DM2_1_4&7	1.7208	0.7996		0.8147		0.7838	
DM2_1_5&6	2.7431	0.892		0.9474		0.9318	
DM2_1_5&7	0.8668	0.5748		0.7931		0.67	
DM2_1_6&7	1.1901	0.8003		0.7015		0.7878	

APPENDIX U: DIYABC MODEL CHECKING SUMMARY STATISTICS

Table U1 Model checking of the best colonization scenario (colonization from the west) of longnose suckers (*Catostomus catostomus*) in the Kogaluk River based on comparison of summary statistics of the observed data set with those derived from the posterior distribution of parameters. Values for each scenario denote the proportion of data sets that have a summary statistic value less than that of the observed value: * denotes a value >0.95 or <0.05, ** denotes a value >0.99 or <0.01, * denotes a value >0.999 or <0.001. N2P is the number of alleles (two-sample), FST is the pairwise FST (two-sample). Within the summary statistics column 1 is Lake 1, 2 is Genetics H, 3 is Slushy, 4 is Strange, 5 is Esker/WP152, 6 is T-Bone, 7 is Cabot.

Summary Statistic	Observed Value	Proportion (simulated>observed)	Significance
N2P_1_1&2	12.2941	0.466	
N2P_1_1&3	12.0588	0.249	
N2P_1_1&4	11.2353	0.1915	
N2P_1_1&5	14.1765	0.5415	
N2P_1_1&6	12.3529	0.47	
N2P_1_1&7	11.7059	0.321	
N2P_1_2&3	12.9412	0.36	
N2P_1_2&4	12.5882	0.345	
N2P_1_2&5	14.4706	0.552	
N2P_1_2&6	13.2941	0.54	
N2P_1_2&7	13	0.492	
N2P_1_3&4	12.4706	0.2515	
N2P_1_3&5	14.5294	0.5355	
N2P_1_3&6	13.1176	0.412	
N2P_1_3&7	12.6471	0.3245	
N2P_1_4&5	14.4118	0.53	
N2P_1_4&6	12.6471	0.375	
N2P_1_4&7	12.2941	0.3075	
N2P_1_5&6	14.9412	0.6495	
N2P_1_5&7	14.5882	0.597	
N2P_1_6&7	13.1176	0.5565	
FST_1_1&2	0.0428	0.4325	
FST_1_1&3	0.0351	0.554	
FST_1_1&4	0.0702	0.7845	
FST_1_1&5	0.0248	0.456	
FST_1_1&6	0.063	0.72	
FST_1_1&7	0.0517	0.6915	
FST_1_2&3	0.0224	0.561	
FST_1_2&4	0.0562	0.882	
FST_1_2&5	0.0179	0.5145	
FST_1_2&6	0.0338	0.5695	

Summary Statistic	Observed Value	Proportion (simulated>observed)	Significance
FST_1_2&7	0.0265	0.6005	
FST_1_3&4	0.0543	0.997	**
FST_1_3&5	0.0166	1	***
FST_1_3&6	0.0422	1	***
FST_1_3&7	0.0297	1	***
FST_1_4&5	0.0375	0.9885	*
FST_1_4&6	0.0686	0.998	**
FST_1_4&7	0.06	0.9985	**
FST_1_5&6	0.032	0.994	**
FST_1_5&7	0.0232	0.9995	***
FST_1_6&7	0.0228	0.968	*