MOLECULAR TOOLS REVEAL HIERARCHICAL STRUCTURE AND PATTERNS OF MIGRATION AND GENE FLOW IN BULL TROUT (*SALVELINUS CONFLUENTUS*) POPULATIONS OF SOUTH-WESTERN ALBERTA

WILL G. WARNOCK
BSc., University of Lethbridge, 2005

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

[MASTERS OF SCIENCE]

Biological Sciences Department
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

Abstract

Bull trout are a species of fish native to the coldwater mountain streams of Alberta. Because this species is of special conservation concern and displays finely dissected population structure, it is well suited as a model species to test the utility of versatile conservation genetics tools. One such tool, a genetic clustering method, was used to discern the hierarchical population structure of bull trout in the core of their range in South-West Alberta. The method also revealed patterns of gene flow by way of assignment tests. Populations defined by this method were then used as reference populations for mixed-migrant assignment tests, revealing that clustering method-defined populations may be more suitable for such tests rather than traditional approaches that define reference populations by sampling location. Combined with spatial data *a posteriori*, assignment tests had additional utility of discerning spatial scale of movement for juvenile and adult salmonids. This technique provided further evidence that assignment tests may be powerful indirect tools for evaluating migration, and that long-range inter-stream dispersal in juvenile salmonid fish may be more common than previously assumed.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CA: Castle River
Cb: Carbondale River
Cb_p: Carbondale population
COSEWIC: Committee on the status of endangered wildlife in Canada
Du: Dutch Creek
ESA: Endangered species act
ESU: Evolutionary significant unit
FL: Fork length
Ga: Gardiner Creek
Hi: Hidden Creek
Hi_p: Hidden population
HWE: Hardy-Weinberg equilibrium
IBD: Isolation-by-distance
LD: Linkage disequilibrium
LE: Linkage equilibrium
Lli: Lower Livingstone River
Lli_p: Lower Livingstone population
Lo: Lost Creek
LWD: Large woody debris
Mi: Mill Creek
Mi_p: Mill population
MU: Management unit
Nra: North Racehorse Creek
OMR: Oldman River
ORR: Oldman River Reservoir
Ra: Racehorse Creek
Ra_p: Racehorse population
RMP: Restricted movement paradigm
SARA: Species at risk act
Sca: South Castle River
SMM: Step-wise mutation model
Sra: South Racehorse Creek
TPM: Two-phase mutation model
TW: Oldman River tailwater
Uli: Upper Livingstone River
Uli_p: Upper Livingstone population
Wca: West Castle River
Wca_p: West Castle population

Chapter 1

Introduction: Bull trout: a model species for conservation using genetic analysis tools.

Species introduction

Bull trout (*Salvelinus confluentus*) are one of four trout species native to Alberta. Throughout most of the species' range in western North America, this fish has faced moderate to severe declines mainly attributable to overharvest, habitat degradation, blockage of migratory routes and competition/hybridization with invasive species. Declines tend to be most pronounced in areas on the southern and eastern periphery of the native range (McPhail and Baxter, 1996). Alberta represents nearly the entire north-eastern periphery of this range, and indeed in the Oldman River basin of South-West Alberta, the species now occupies only 31% of its former range (Fitch, 1997) (Figure 1-1).

Overharvest is generally regarded as the main contributor for historical decline of most Alberta bull trout populations (Groft, 1997). The aggressive nature of the fish makes it especially susceptible to angling. Because larger fish were generally targeted for harvest, this further decimated populations by selective removal of migratory, mature adults (Berry, 1997). In 1995, Alberta Fish and Wildlife implemented a complete province-wide moratorium of harvest as part of the new bull trout management and recovery plan (Berry, 1997), a regulation that is still in effect to this day. Other jurisdictions have implemented similar restrictive management plans. In British Columbia, the base harvest regulations are determined by a management-zone wide scale,

1

with further restrictions on specific waterbodies warranting higher conservation priority. Throughout the United States, all populations of the species have been listed as "threatened" under the federal Endangered Species Act (ESA) (USFWS, 1999). This listing has resulted in an integrated recovery plan that includes habitat protection as well at restrictive angling regulations. Bull trout are not currently protected under the Canadian federal Species at Risk Act (SARA), nor are they on the list of species to be reviewed by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) – the organization which provides recommendations to the federal government on species to be listed under the SARA.

Another risk factor in the decline of bull trout has been habitat degradation. Spawning takes place in the fall, generally in cold, fast 2nd or 3rd order mountain streams >2m width with suitable gravels, groundwater upwelling and cover (McPhail and Baxter, 1996; Dunham and Rieman, 1999, Baxter and Hauer, 2000). Anthropogenic disturbances known to increase siltation into the interstitial spaces of gravels and reduce aeration are devastating to egg development (Fraley and Shepard, 1989). Siltation may also reduce aquatic invertebrate productivity which serves as food sources for fry and juvenile fish (Nakano, 1992). Causes of increased siltation include logging, oil exploration and road and pipeline construction. Because eggs and juveniles must survive the winter, base flow levels are very important in these habitats. Activities (e.g. water withdrawal for artificial snow making) which decrease base flows may therefore result in reduced recruitment of bull trout (Post and Johnston, 2002). Suitable cover and macrohabitat complexity are especially important for rearing of fry and juveniles. Stable undercut banks, broken riffle substrate and large woody debris (LWD) are used as refuges for the secretive fry and

juveniles to avoid predation and competition (Dambacher and Jones, 1997; Goetz, 1997; Earle and McKenzie, 2001). Anthropogenic factors which may reduce habitat complexity are many. Channelization and culverts represent an extreme example of reduction in all or most critical habitat. Increased siltation from aforementioned causes may "cement" gravels and cobbles into the streambed and reduce fry habitat (McPhail and Baxter, 1996). Activities which result in reduced bank stability, such as logging or removal of riparian vegetation may eliminate undercut bank habitat. Finally, activities that reduce LWD presence, such as logging, may result in extreme changes to stream microhabitat characteristics which characterize many bull trout spawning streams (Hauer *et al.*, 1999).

Because many bull trout populations are composed primarily of migratory individuals as the effective spawning adult group, open migratory routes are critical. Fragmentation of riverine systems caused by dams and improperly designed culverts can eliminate an entire population if they block passage to spawning or over-wintering habitat. The effects of migratory impediments on genetics and life history strategy will be discussed further in this introduction and subsequent chapters.

Invasive species represent the final substantial threat to native bull trout populations. Alberta has a prolific history of stocking of non-native fishes into its lotic ecosystems. Introductions of exotic salmonines including brook trout (*Salvelinus fontinalis*), Northern Dolly Varden (*Salvelinus malma*), brown trout (*Salmo trutta*) and non-Athabasca strain rainbow trout (*Onchorynchus mykiss*) have taken place in the bull trout native range. Introductions of species native to Alberta into other parts of the province into areas where bull trout were the only salmonine species present has also taken place with westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and lake trout

(*Salvelinus namaycush*). Effects of introductions on bull trout range from minimal to complete extirpation (Post and Johnston, 2002). In lotic ecosystems, the most severe effects are from introduced brook trout, which tend to outcompete bull trout in resource-poor headwater tributaries (Gunckel *et al.*, 2002), freely hybridize and introgress with bull trout (Kanda *et al.*, 2002) or result in illegal harvest of bull trout through mistaken identity. Brook trout invasion may be exacerbated by habitat degradation, as "ideal" bull trout habitats with high complexity are more resistant to invasion from this species (Rich *et al.*, 2003).

Bull trout have been shown to exhibit three life history strategies entirely in freshwater: stream resident, fluvial and adfluvial. Stream residents are non-migratory, small forms (rarely exceeding 300mm) that rear and complete their lifecycle in headwater mountain streams (McPhail and Baxter, 1996). Fluvial trout rear in headwater mountain streams and migrate out to large tributaries or main-stem rivers at a certain age and size (usually 2+ years and 200-300mm) where they become piscivorous; they return to spawn in headwater streams at approximately 5 years of age, usually in alternate years (McPhail and Baxter, 1996; Fraley and Shepard, 1989). Adfluvial morphs rear in headwater mountain streams and migrate out to lakes where they become piscivorous; they return to spawn in headwater streams at approximately 5 years of age, usually every year if the habitat is productive enough (Stelfox, 1997). The two migratory forms may undergo long annual migrations of over 200 km to spawn (McPhail and Baxter, 1996; Shepard *et al.* 1984) and generally display precise homing to their natal stream (McPhail and Baxter, 1996; Swanberg, 1997; Bahr and Shrimpton, 2004). Outmigration of subadults from tributaries to rivers or lakes generally occurs in the high flow spring months, while

spawning migration of adults from lakes or rivers to tributaries usually takes place in late summer at lower flows (McPhail and Baxter, 1996). Sympatry of resident and migrant bull trout appears to be variable. In some areas, residents are only found above natural migratory barriers (Mcphail and Baxter, 1996), while in other areas, adults of the two forms are naturally sympatric (Nelson *et al.*, 2002). Bull trout have evolved variable life history and site specific spawning strategies to deal with the stochastic nature of mountain streams. Resident forms are highly adapted to their individual streams. Migrants spend the majority of their lives away from their natal streams, and therefore are able to avoid local extinction events, repopulating the area when the stream recovers sufficiently enough to spawn and rear juveniles in. Because these large migrants exhibit high longevity, (often alternate-year) iteroparity and high spawning site fidelity, they buffer populations from short-term disturbances which would otherwise extirpate residents (Dunham and Rieman, 1999; Neraas and Spruell, 2001).

An interesting question which has receive little study with respect to bull trout biology, and indeed most salmonid biology, is what the extreme long-range migratory tendencies are of the juvenile stage. Generally, juvenile salmonids are considered to be sedentary, rearing close to the area in which they are born (Gerking, 1959); however, recent evidence suggests that salmonid fish previously thought to be sedentary might in fact be quite mobile (Gowan *et al.*, 1994), even at the juvenile stage (Kennedy *et al.*, 2002; Homel and Budy, 2008; Rasmussen *et al.*, in press). While juvenile and subadult fish of migratory species must, at some point, emigrate to reach areas of higher productivity as adults, dispersal of younger juveniles into alternate rearing streams is

poorly studied and understood. Therefore, patterns of long-range dispersal of juvenile

bull trout between tributaries remains a mystery.

Bull trout are an excellent indicator species for pristine environments, are valued

in recreational fisheries and occupy an important ecological niche as top predator. The

species is also an excellent model species for many basic and applied ecological studies

and deserves special conservation efforts due to its susceptibility to overharvest and

habitat destruction.

Study area

The Oldman River basin is located in South-Western Alberta and Northern

Montana. It is one of 4 Alberta river basins that represent the mountain headwaters of the

Saskatchewan River drainage, which eventually drains into the Hudson Bay in Manitoba.

The headwaters are made up of 6 major sub-basin drainages which arise on the East

slopes of the continental divide. These sub-basins, from North to South are: upper

Oldman, Crowsnest, Castle, Waterton, Belly and St. Mary's (Figure 1-2).

The upper Oldman, Crowsnest and Castle river sub-basins are now all above a

large impassable hypolimnetic-draw dam which was constructed in 1990. All three rivers

now flow directly into the Oldman River Reservoir (ORR) which has been created by this

dam. Because the Crowsnest River sub-basin is considered devoid of a self-sustaining

bull trout population (Fitch, 1997; Warnock, personal observations), only the other two

will be considered in the study area. These two remaining sub-basins represent the core

of the bull trout's range in the entire Oldman River basin (Fitch, 1997).

The upper Oldman River sub-basin represents all waters flowing into the Oldman River above the ORR. There are 2 sets of falls on the main-stem of the river, one a seasonally passable set (Gap falls) where the river leaves the mountains, the other an impassable set (Upper Oldman falls) located near the headwaters. Bull trout occupy the entire range of this river from below the Upper Oldman falls to the reservoir (Fitch, 1997). The creeks Callum, Bob and Camp, which are foothill streams flowing into the river below the Gap falls are now considered to be devoid of bull trout populations. Above the Gap falls, the 6 creeks North and South Racehorse, Daisy, Vicary, Dutch, Hidden and the Livingstone River are all thought to have retained their bull trout, though the presence of resident and fluvial morphs in each is largely unknown (Fitch, 1997). Redd surveys provide some evidence for which tributaries appear to have the largest effective population sizes. In a study conducted by the Alberta Conservation Association, through 4 years of sampling, the highest number of redds were seen in the Livingstone River and Hidden Creek (Gerrand and Watmough, 1998).

The Castle River sub-basin represents all waters flowing into the Castle River above the ORR. This sub-basin is likely the least affected system in the entire Oldman River Basin with respect to population declines, as bull trout are thought to occupy 95% of their historical range (Fitch, 1997). Like the Oldman River, the Castle River has a seasonal set of falls (Castle falls) which may impede spawning migrations during the low-flow period when bull trout are thought to migrate; however, unlike the Oldman River, most of the major spawning tributaries for bull trout in this system are located below this set of falls (Gerrand and Watmough, 1998). Above Castle falls, the Castle River splits off into two branches: the West and South Castle Rivers, both containing bull

trout (Fitch, 1997). Below the falls, there are four tributaries containing bull trout: the Carbondale River and Gardiner, Lost and Mill Creeks (Fitch, 1997). Life history strategies for each stream are unknown, though it may be likely that the South and West Castle Rivers contain more resident bull trout due to their presence above a seasonal barrier. Redd surveys provide some indication for which tributaries appear to have the largest effective population sizes. Through 4 years of sampling, the highest number of redds were seen in Mill and Lost Creeks (Gerrand and Watmough, 1998).

Bull trout are presently found below the Oldman River dam in the main-stem of the tailwater section, albeit in low numbers (Fernet and O'Neil, 1997). These fish are presumed to represent upstream source populations, from one of the three possible sub-basins. It is unknown at this point what the specific source populations of these fish are.

Population genetics, basic concepts and methods

The field of molecular ecology is a recent and rapidly expanding field which uses quantifiable genetic data to address ecological questions. Phylogeography, population dynamics, mating systems, wildlife forensics, speciation and general evolutionary trends are just a partial list of the topics which may be addressed with molecular data. This study will be concerned with characterizing population genetic trends of the bull trout in the spatial network of the study area.

Population genetics is concerned with the partitioning of genetic variation within an individual species. Populations are groups of organisms that can freely interbreed within their own group, but have reduced breeding rates with other groups. As these

8

populations become more reproductively isolated from each other, they tend to undergo genetic divergence. Classical examples of genetic divergence are the result of allopatric divergence, when two populations diverge because they are geographically isolated from one another, and thus cannot freely interbreed with each other. Such examples of divergence have likely been responsible for the formation of most unique populations and sub-species of salmonid fishes of the genus *Onchorynchus* (Behnke, 2002). Over the course of Pleistocene glaciation events of the past several hundred thousand years, the rocky terrain of Western North America has had a dynamic history of freshwater migration routes for any fishes inhabiting its waters. Many sub-species and unique populations of cutthroat (*Onchorynchus clarki*) and rainbow (*Onchorynchus mykiss*) trout have been formed in this time by drainage transfers of common ancestor trout. This occurs either via headwater stream exchange or from proglacial lakes temporarily spanning drainages (Behnke, 2002). After these dispersal routes became impassable, usually in interglacial periods, the founding population in the virgin drainage and the original population underwent allopatric divergence. This may ultimately culminate, if there is enough time and evolutionary pressures, in sub-speciation or speciation.

It is no surprise that freshwater ecosystems have been important model systems for the study of population genetics, especially those with recent history of glaciations. Because most freshwater organisms are dependent on contiguous freshwater connections for dispersal, and with the recent expansion of knowledge of post-glacial migration routes, population genetics of inland fishes in particular has received much attention recently. In addition to providing answers to basic questions on population divergence, speciation and phylogeography, such studies also have practical applications in

conservation. Traditional management practices attempted to conserve organisms at the species level. It soon became apparent, however, that such practices led to dismal failures. Nowhere is this more apparent than in hatchery-raised pacific salmon stocking of much of the 20th century, which rarely strengthened and sometimes worsened local stocks (Behnke, 2002). Knowing that populations may rapidly diverge, and that these may represent adaptive differences to the local environment, it would be prudent to conserve the "local stock" or that which is most adapted to surviving in a specific ecosystem. This becomes especially important when considering that freshwater ecosystems, particularly coldwater lotic ones, are notoriously stochastic in nature and may impart significant selective pressures on any animals inhabiting them (McPhail, 2007). As such, several designations of conservation priority have been adopted. The evolutionary significant unit (ESU) is a population or complex of populations which contain a significant portion of evolutionary history (Moritz, 1994). This definition is obviously rather vague, but is useful for defining the most divergent population groupings below the sub-species or species taxonomic level. The management unit (MU) is a population which can be measured as genetically distinct due to low gene flow with neighboring populations (Mortiz, 1994). In salmonid fish literature, this term is usually interchangeable with the term "stock" (Allendorf and Luikart, 2007). As alluded to previously, conservation at the stock or MU level is often important in conservation of salmonid fishes due to the high degree of habitat specialization. This hierarchical approach to conservation reflects the hierarchical population structure that may be present across a species' range. This is especially important in salmonid fishes, which tend to display levels of hierarchical population structure that are organized according to

the nested hierarchy of the stream systems they occupy (Whiteley *et al.*, 2006a).

Detection of this hierarchical population structure may therefore be important in determining a guided management strategy for all levels of population structure in these species (Whiteley *et al.*, 2006a). With looming climate change, water quality and supply crises, coupled with their cultural and economic significance, salmonid fishes have become somewhat of a poster child for conservation in coldwater North American ecosystems. It is therefore a combination of both model system suitability and high conservation concern which has drawn such a large amount of population genetic research on these organisms.

Intra and interpopulation variation is measurable, and it is important in determining overall population structure; however, we must understand the underlying causes that drive this variation and what trends can we expect in population structure. Assuming we are only examining neutral markers, some factors which will affect variation in populations are gene flow, genetic drift, mutation rate, social structure, mating system and effective population size. For the purpose of this study, we will assume the major forces driving population structure in bull trout populations are gene flow, genetic drift, mutation and effective population size.

High N_e
High N_{em}
Low N_e
Low N_{em}

Gene flow   Genetic Drift

Population Homogenization   Population Divergence
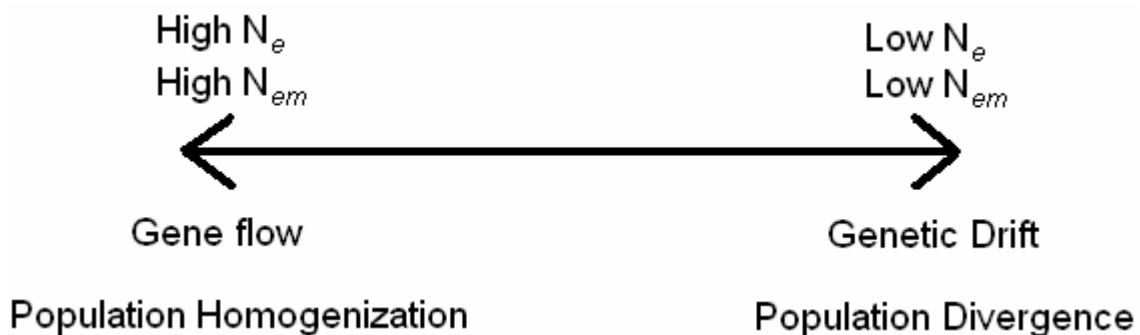
Figure 1-3: Forces driving population divergence or homogenization

Figure 1-3 shows how most of these forces affect interpopulation diversity. On the left side of the spectrum high gene flow between populations and large effective population size will convey weak population structure. On the right end of the spectrum, extremely reduced or no gene flow, high effects of genetic drift and low effective population size will convey strong population structure. The force of gene flow is determined by the number of effective migrants (N_em) which disperse and breed with populations other than the one in which they were born. Gene flow will therefore reduce interpopulation diversity, but increase intrapopulation diversity. Genetic drift is a phenomenon associated with random changes in allele frequencies from one generation to the next. As populations diverge, allele frequencies will randomly either fixate or be driven to extinction, a phenomenon known as coalescence. This will result in increased population divergence and a loss of heterozygosity, decreasing intrapopulation diversity. The rate of coalescence, and therefore drift, is especially pronounced in populations with small effective size (N_e). In extreme examples of drift, it may take a single or very few generations to drastically change allele frequencies. These acute examples are the result of founder or bottleneck events, where genotypes of the few founding individuals will have profound effects on the allele frequencies of the resulting population. Mutation is another force which will tend to increase both intrapopulation diversity and interpopulation diversity. This is a force which is also intimately linked to N_e, as *de novo* mutations will have larger effects on allele frequencies in small populations than large populations.

It is imperative to have an operational definition of what a population is. Unfortunately, the term population itself is loosely defined in population genetics. Different conceptual frameworks of populations can require different methods of quantification, and lead to inconsistency in quantifying diversity therein for conservation applications. Waples and Gaggiotti (2006) define two major theoretical frameworks for populations, the ecological and evolutionary paradigms. The ecological paradigm is defined as "A group of individuals of the same species that co-occur in space and time and have an opportunity to interact with each other." The evolutionary paradigm is defined as "A group of individuals of the same species living in close enough proximity that any member of the group can potentially mate with any other member." It is therefore apparent that the ecological paradigm is more focused on ecological interactions of organisms inhabiting the same environment; whereas the evolutionary paradigm is more focused on the reproductive interactions of the individuals inhabiting the same environment. Fish populations are generally considered unique under the evolutionary paradigm if they show significant divergence at neutral markers from neighbouring populations. Because neutral markers are not necessarily indicative of adaptive differences, I favour the term "genetic" rather than evolutionary paradigm as more appropriate. At this point, we must ask: "how much is divergence is deemed significant?" Again, this will vary from study to study based on the criteria for differentiation, the model by which it is estimated or the criteria for conservation. From an ecological point of view, we may consider that all fish occupying any specific rearing stream are a distinct population. Alternately, we may use a genetic point of view to define a population based on a grouping of individuals showing genotypic similarity, using no *a*

priori knowledge of sampling location – defining populations based on genotypic clustering methods (see glossary of genetic techniques). These genotypic clustering methods may also have additional advantages of detecting any hierarchical manner to the population structure that does exist, and providing assignment tests for use in determining admixture or dispersal rates between populations.

Classical assignment tests attempt to assign an individual organism to their population of origin. If the genotype of an individual is known, one can calculate the probability of that genotype occurring in any potential source population (reference population). Such tests are extremely valuable to assessing stock composition of mixed-migrant groups (Koljonen *et al.*, 2005) spatio-temporal tendencies of individual migrant populations (Potvin and Bernatchez, 2001), sex-biased dispersal (Hansen *et al.*, 2001), wildlife forensics (Primmer *et al.*, 2000), hybridization (Pritchard *et al.*, 2007a) or other specialized population questions. One important consideration is the definition of the reference population. Generally, reference populations used in assignment tests are defined *a priori* as geographic regions of origin (e.g., Neraas and Spruell, 2001), but as previously discussed, such a definition would be based on an ecological framework (Waples and Gaggiotti, 2006). Since reference populations are supposedly genetically recognisable units, a definition based on a genetic framework should be more appropriate as it should theoretically maximise assignment success.

Taxonomy of species of the genus *Salvelinus*, or the chars, has a convoluted history. This is largely due to phenotypic similarities between species and large phenotypic variation within species. Bull trout were only classified as a separate species from the Dolly Varden in 1978 (Cavender, 1978).

Previous studies have shown that bull trout tend to display high interpopulation but low intrapopulation genetic diversity (Spruell *et al.*, 1999; Neraas and Spruell, 2001; Costello *et al.*, 2003; Spruell *et al.*, 2003; Whiteley *et al.*, 2004, Whiteley *et al.*, 2006b). These trends are likely observed because of two ecological features of the fish: philopatry and low effective population size. Although bull trout may migrate long distances, they display high precision in homing to their natal habitat for spawning events (McPhail and Baxter, 1996; Spruell *et al.*, 1999). Such philopatric behaviour reduces the potential for gene flow between geographically distant populations, facilitating genetic divergence. Bull trout effective population sizes are low in comparison to most other salmonid species (Dunham and Rieman, 1999; Whiteley *et al.*, 2004). With low population size, the number of effective migrants (N_em) overall is reduced, leading to reduced probability of straying events that may lead to gene flow between populations (Whiteley *et al.*, 2004). In addition, small populations tend to have markedly different genetic features from large populations due to the large effect of genetic drift, especially acute drift effects such as bottlenecks or founder events.

Genetic patterns are variable throughout the species' range. Bull trout utilized two separate refugia during the most recent glaciation and hence have recolonized the North-

west part of the continent as two clades: the coastal and interior (Taylor *et al.*, 1999).

Although morphological differences seem limited (but detectable), the two clades are

highly distinguishable at both codominant and uniparental markers (Taylor *et al.*, 1999;

Spruell *et al.*, 2003). The interior clade has recolonized far more area than the coastal

clade following the last glaciation (Taylor *et al.*, 1999) and has colonized a more

extensive range than most other inland salmonids. This is likely attributable to the

migratory nature of the species and its ability to occupy the highest headwater tributaries

of drainages, facilitating headwater transfers. During the recolonization process, bull

trout faced harsh, stochastic environmental conditions; as a result, populations were

likely founded by few individuals or may have become bottlenecked (Costello *et al.*,

2003). Large-scale patterns of intrapopulation diversity are therefore observed throughout

the range of the interior clade, with populations central to the bull trout range displaying

the highest heterozygosities and the most recently founded populations on the periphery

of the range displaying the lowest heterozygosities (Costello *et al.*, 2003). This is

consistent with expected trends for a species which colonizes in a stepping-stone manner,

with repeated founder events (Costello *et al.*, 2003). At fine spatial scales, however,

heterozygosity may be affected greatly by landscape features, such as the presence of

migratory barriers, indicating the importance of contemporary factors that affect genetic

diversity (Castric *et al.*, 2001; Costello *et al.*, 2003). Interpopulation diversity seems to be

largely dominated by local features of the environment, and not as dependent on

historical colonization events (Costello *et al.*, 2003). At microsatellite markers, pairwise

population differentiation appears strong at fine spatial scales near the core of the bull

trout range (Neraas and Spruell, 2001) and slightly lower, but detectable near the

periphery (Whiteley *et al.*, 2006b). Where populations are isolated above barriers, resolution is higher (Costello *et al.*, 2003).

Until recently, it was not known what genetic differences might exist in populations with sympatric resident and migrant life-history forms. Homel *et al.* (2008) found that there is little to no genetic divergence between the two forms at neutral microsatellite markers. The authors do acknowledge, however, that quantitative phenotypic variation may be heritable and not detectable by indirect methods using neutral markers. Indeed, evidence does suggest neutral genetic variation (measured by F_{ST}) may underestimate phenotypic variation (measured by Q_{ST}) of some heritable traits between resident and migrant salmonids (Perry *et al.*, 2004; 2005).

Given that Alberta populations represent the extreme north-eastern periphery of the bull trout range, previous studies have shown that Alberta populations predictably show very low intrapopulation diversity and reasonably high interpopulation diversity (Groft, 1997; Taylor *et al.*, 1999; Costello *et al.*, 2003; Spruell *et al.*, 2003). It is not apparent how fine-scale structure in Alberta compares to other areas throughout the bull trout range, as all previous studies have been part of large-scale examinations of bull trout population genetics. It is also not apparent how bull trout populations in general are organized when using genetic clustering methods. It may be that such methods provide interpretable results of hierarchical population structure for this species, even at fine spatial scale.

Objectives of study

This thesis will be separated into three experimental chapters which will: 1) examine fine-scale population structure of bull trout in the species' core range of South-West Alberta by genetic clustering methods, 2) the utility of assignment tests to assign migratory bull trout to population of origin, and 3) the utility of assignment tests to quantify scale of migration in this species.

All potential spawning and rearing streams in the study area will be sampled for presence of bull trout. In those streams where bull trout are found, fish will be subject to a multilocus microsatellite DNA analysis. Basic population data will be calculated for each stream (H_e, H_o, A, F_{ST} etc.). Secondarily, populations will be defined based on genotypic clustering methods using no *a priori* knowledge of sampling location. Basic population data will be re-run on those clusters defined. To this end, I will examine fine-scale population structure of bull trout based on both ecological and genetic frameworks described by Waples and Gaggiotti (2006). Genotypic clustering methods will also be used to determine any hierarchical arrangement to the population structure in the spatial network of the drainage. Conservation implications to these differing layers of population structure will be discussed.

Migratory bull trout are common throughout the study area. If fine-scale genetic population differentiation is observed, then one could assign the genotypes of those migrant fish back to a population of origin. Populations defined by both stream-of-origin and genetic clustering from the previous chapter will be used as reference populations for genetic assignment. The relative success and spatial resolution of either technique will be

examined in order to determine if the genetic framework for defining populations is more analytically robust. This will aid researchers in future studies when defining reference populations used in mixed-migrant genetic assignment tests.

Genetic assignment tests have become extremely important and common tools for answering a variety of ecological questions. Two considerable questions in salmonid biology which have not largely been answered by these techniques remain. These are whether spatial scale of movement can be quantified by these tools, and what the extreme long-range inter-stream dispersal tendencies of juvenile fish are. Hierarchical populations defined by genetic clustering methods will be examined for the spatial scale of both adult and juvenile bull trout movement. This will provide valuable data for population-level conservation, as migratory tendencies of each locally adapted population will be ascertained. It will also potentially provide evidence for extreme long-range inter-stream movement of juvenile salmonid fishes.

Glossary of genetic techniques

Population genetics is concerned with natural genetic variation within and between populations in a species. Molecular techniques are therefore designed to quantify this genetic variation by way of allele frequencies. An allele is a character state of a gene found at a specific identifiable region on a chromosome, known as a locus. If there are 2 or more alleles commonly found at a locus in a population, it is said that this locus is polymorphic. Thus allele frequencies are quantified by measuring proportions of alleles found at polymorphic loci. Where populations show differing allele frequencies, this may be interpreted as genetic divergence; the greater the magnitude in differences, presumably the more divergent the populations.

Historically, variation was measured by looking at phenotypic characteristics which divergent populations tended to show (e.g. in fishes: gill raker number and length, body depth to length ratio etc.). Assuming these traits are heritable, this only indirectly measured genetic divergence, and genetic material that was under selective pressure. Selective pressure in undesirable when quantifying genetics because it need not follow predictable models of allelic variance based on Hardy-Weinberg equilibrium (HWE). Current robust statistical techniques which are used to measure genetic variability assume the 5 rules of HWE: no net mutation, no migration between populations, random mating, large population and no selection. Therefore when we measure intrapopulation or interpopulation diversity, we are assuming that there have been no recent major violations of these 5 rules. Additional confounds of measuring those genetic traits under selective pressure is that they do not always correspond to the average genetic distinctiveness of populations, and selective pressures are not constant in time and space.

Populations may therefore show marked phenotypic divergence, but display little genome-wide genetic divergence, or vice-versa. This will be explained further in the proceeding paragraph.

Current molecular techniques are concerned with directly measuring variation in genetic material which is neutral, i.e. under no selective pressure. This may be counter-intuitive as adaptive divergence is very important to population divergence and hence speciation ultimately; but to measure genome-wide divergence in all those genes with adaptive significance is impossible with current molecular techniques and knowledge. Neutral markers therefore give us a "benchmark" way of measuring average genome-wide divergence in the absence of such selective pressures. This is not to say that divergence at neutral markers is the only and most valuable measure of population distinctiveness for management practices. For example, populations of steelhead and resident rainbow trout show little (if measurable at all) genetic distinctiveness at neutral markers, but the phenotypic differences between the two sympatric pairs are obvious and must be at least partially heritable (Narum *et al.*, 2004). Indeed, this appears to be the case for sympatric resident and migrant bull trout (Homel *et al.*, 2008). In practice, if one population were eliminated, it may be impossible or take many generations for the alternate population to evolve once again from the remaining gene pool. Alternately, populations showing no overt phenotypic differences may show large genetic differences, what I will refer to as cryptic genetic divergence. Bull trout populations tend to show this cryptic genetic divergence, even at fine spatial scales (Whiteley *et al.*, 2006b). This divergence may even manifest itself at higher taxonomic units, recent genetic evidence led to designation of a new cryptic species of bat which displayed no overt phenotypic

difference but large genetic divergence (Mayer and Von Helversen, 2001). It is important, then, to draw from multiple indicators of population variation and take a conservative approach when reaching management decisions. Such practices maximize the persistence of natural genetic variation within and between populations of a species, allowing resiliency of wild populations in the face of an ever changing and increasingly impacted world.

There are two ways to approach measuring genetic variation in population genetics, depending on the degree of resolution the study requires. Large-scale resolution is concerned with measuring those deepest splits in population structure throughout a species range. Fine-scale resolution is concerned with more recent divergence of populations and thus more local population structure. Markers that display large-scale resolution are those that are uniparentally inherited. In animals, this is usually measured in mitochondrial (maternal) or non-recombinant regions of uniparentally inherited sex chromosome (paternal in most animal taxa, including fishes and mammals) markers. These markers do not undergo genetic recombination, so an historical archive of haplotype diversity is retained and may be used to infer deep population structure. Codominant markers are those which are inherited by both parents in diploid organisms, at (generally) non sex-linked nuclear chromosomes. Because of genetic recombination, these markers are generally more suited to infer more recent population structure. Resolution may also be affected by the physical type of marker used. Markers which are on regions of highly conserved regions of a genome (e.g., cytochrome b gene of the mitochondrial genome) undergo very low substitution rates. Such markers are appropriate for studies requiring extremely low resolution, where locus polymorphism may manifest

itself at higher taxonomic units. The opposite end of the spectrum is occupied by microsatellite markers. These are short tandem repeats of what is usually a di, tri or tetranuclotide sequence, alleles of which can be identified on a chromosome and characterized by overall base pair length (Li *et al.*, 2002). Because microsatellites are often highly polymorphic, mainly due to high mutation rate, they are very useful in displaying population structure (Li *et al.*, 2002). This study will use codominant microsatellite markers as it is concerned with fine-scale genetic partitioning – that within and among populations which are in close geographical proximity and share recent common ancestry.

Perhaps the most worrisome problem in molecular studies is the phenomenon of homoplasy, or convergent molecular evolution. Especially in markers such as microsatellites, alleles that appear identical between populations may not be identical by descent, but rather the result of both populations evolving the same alleles independently (Angers and Bernatchez, 1997). Such results obviously lead to under-estimation of molecular divergence. It is therefore important to use caution in interpreting results drawn from raw genetic data, and use multiple loci displaying moderate polymorphism so that robustness of such studies is maximized.

The Hardy-Weinberg Principle

The most basic concept in population genetics, and that which most quantitative population genetics is based off of is Hardy-Weinberg Equilibrium (HWE). Genotype frequencies of sexually mating organisms follow predictable modes of inheritance if several biological conditions are met under HWE. If a population is large, mating

randomly and undergoing no net migration, allele frequencies will stay the same between generations; providing the alleles are under no selective pressure or mutation process. Because alleles are inherited in a mathematically predictable fashion under these assumptions, it is important to determine whether a sampled population shows genotype frequencies conforming to those predicted under HWE. For simple tests of one population, one locus and 2 alleles, a simple chi square test can be used, whereby observed genotype frequencies are used to determine expected frequencies according to the equation:

$$p^2 + 2pq + q^2 = 1$$

Where p and q are the allele frequencies derived from observed genotype frequencies. The equation becomes notably more complicated when multiple alleles are present. Once genotype frequencies are found by multiplying population allele frequencies by the observed frequencies, the observed and expected values may be compared to arrive at a chi-square value:

$$x^2 = \sum \frac{(O-E)^2}{E}$$

Where O and E are observed and expected genotype frequencies, respectively. This chi-square value is than compared to the degrees of freedom on a chi-square table generated for a pre-determined alpha value, usually 0.05. Degrees of freedom are calculated by:

$$DF = \frac{Na(Na-1)}{2} \ or = Ng - Na$$

Where Na is the number of alleles and Ng are the total number of possible genotypes in the population. A p-value may now be compared to the pre-determined alpha to determine if the population is out of HWE.

For a more exact test of probabilities, the software package GENEPOP 3.4 may be used (Raymond and Rousset, 2003). For 2 or 3 alleles, the p-value obtained for each locus is calculated by the equations outlined by Louis and Dempster (1987). These equations are either a Fisher's test (2 alleles) or a computational algorithm (3 alleles) which take into account sample size. For 4 or more alleles, a Markov Chain/Monte Carlo (MCMC) simulation is used, with default numbers of 100 batches and 1000 iterations/batch for the chain. The algorithms used in these simulations were developed by Guo and Thompson (1992) and are more compact, to save in computational time which was tedious for multiple alleles using the Louis and Dempster (1987) method. To test loci across all populations and populations across all loci, a Fisher's exact test is used.

Significant deviations between observed and expected genotypes within a population may reflect a number of difficulties in interpreting data or recent biological processes. If a locus shows significant differences among most or all populations tested, it may indicate a problem with the locus used, such as high presence of null alleles, linkage disequilibrium or high mutation rate at that locus. If significant differences are detected among all or most loci within a population, it could reflect an ecological event, such as non-random mating, migration or very recent evolution attributed to drift; likewise, artefacts of sampling may result in deviations, such as sampling two populations and classifying them putatively as one (Wahlund effect) or sampling only highly related individuals.

Intra and interpopulation diversity

Intrapopulation variation is very important in population genetics. Oft-reported measurements of intrapopulation diversity are expected heterozygosity (H_e) and mean allelic diversity (A). H_e is a value between 0 and 1 calculated using observed allele frequencies in the population according to the formula (for one locus):

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

Where p_i is the frequency of the ith of k alleles. We can see from this formula that as a locus becomes more polymorphic and as alleles are more equally represented in the population, heterozygosity will increase; hence, large heterozygosities are seen in populations showing high diversity. A is a simple mean of the alleles found at loci within a population. Because both measures are highly dependent on locus choice, cross literature comparisons of these values are subject to extreme caution.

Within populations, it is also important to test for inbreeding, as such results can lead to deviations from HWE. The most common statistic to measure inbreeding is the inbreeding coefficient, F_{IS}. This statistic is based on the probability of two alleles being identical by descent, and can be calculated from frequency data according to the formula:

$$F_{IS} = \frac{H_e - H_o}{H_e}$$

Where H_e is the expected heterozygosity within a randomly mating subpopulation (described above) and H_o is the observed heterozygosity per individual within the subpopulation. Values closer to 1 represent populations with an heterozygote deficiency, which may indicate inbreeding, while values closer to -1 represent populations with an heterozygote excess.

Fundamental to studies in population genetics is some kind of numerical representation of interpopulation diversity. The most common statistic to measure population divergence is F_{ST}. F_{ST} was originally developed by Wright (1951) as a measurement based on F_{IS}. Today the value is estimated by many different methods such as G_{ST} (Nei, 1972), θ (theta) (Weir and Cockerham, 1984) and R_{ST} (Slatkin, 1995) which all use allele frequencies as raw data. The values obtained are often used to describe interpopulation structure in a species, but may also be used to infer gene flow between populations (but see Neigel, 2002 for limitations of the latter).

Wright's original definition of F_{ST} can be derived from allele frequencies by the formula:

$$F_{ST} = \frac{Var(p)}{\bar{p}(1 - \bar{p})}$$

Where $Var(p)$ is the variance of frequencies among populations and $\bar{p}$ is the mean allele frequency.

Because heterozygosity can be viewed as a measure of genetic diversity, global F_{ST} can also be calculated with simple frequency data according to the formula:

$$F_{ST} = \frac{H_T - \overline{H_e}}{H_T}$$

where H_T is a measure of the total expected heterozygosity among all populations being tested and $\overline{H_e}$ is the average expected heterozygosity among all populations being tested. Generally in fine-scale population studies, the interest lies in pairwise F_{ST} values, where parameters from only 2 populations need be considered. To obtain an F_{ST} value across multiple loci, the F_{ST} values can simply be averaged. Unfortunately, this method makes

certain simplistic assumptions unlikely to be found in real datasets. Weir and Cockerham (1984) have outlined a series of statistical approaches which do not make assumptions of sample size, numbers of populations and multiple alleles; as a result, their estimator, θ (theta) has caught on as the most common F_{ST} analogue. Their tests can also be used to find values of F_{IS}. The program FSTAT (Goudet, 1995) calculates θ based on raw allelic data and performs sequential Bonferroni adjustment to F_{ST} values to test for significance (Rice, 1989).

F_{ST} suffers from many limitations, particularly when making inferences of gene flow (Neigel, 2002) or cross-literature comparisons (Hedrick, 2005). Partly, this may be due to traditional estimators of F_{ST} (including θ) being based on an infinite alleles model, which assumes that every mutation event creates a new allele (Selkoe and Toonen, 2006). Several derived estimators take into account that microsatellite alleles tend to evolve according to a stepwise mutation model (e.g.: Slatkin's (1995) R_{ST}). This is because microsatellite alleles separated by a single repeat difference are usually more closely related than those which are two or more repeats different. Using this knowledge, these estimators may provide a more realistic value of population divergence (but see Selkoe and Toonen, 2006 for limitations) than F_{ST}. The other major limitation of traditional F_{ST} – based methods is that F_{ST} is not a standardized value and will vary depending on the mutation rate of the locus (Neigel, 1997), intrapopulation diversity (Charlesworth, 1998; Hedrick, 1999, 2005) or minor sampling errors. It is therefore important to use caution in comparing F_{ST} values between different studies, particularly those which use different markers or examine different taxa. Nevertheless, θ is often the only value reported in

population genetic research and thus remains the most important value for cross-literature comparisons and the standard value for population differentiation.

Population structure defined by genetic clustering methods

        The program STRUCTURE 2.2 (Pritchard, 2000) uses a model-based clustering approach to define populations with no *a priori* knowledge of sampling location. The model assumes that that there are K populations within the sample (if $K > 2$, it is called an admixed sample), each containing its own set of allele frequencies at each locus. Each population is assumed to be under HWE and linkage equilibrium. The program uses a Bayesian estimator to probabilistically assign individuals to one of the K populations based on minimization of deviations from HWE and linkage equilibrium. An overall value of the log probability of the data conforming to the given K is calculated as $\ln \Pr(X|K)$. By running the program at different values of K, the true K may correspond to the highest $\ln \Pr(X|K)$ value which captures the major structure in the data (Pritchard *et al.*, 2007b). The current version of the program also allows inferences of structure based on a linkage model, which takes into account locus correlations due to historical admixture (Falush *et al.*, 2003).

        Recently, Evanno *et al.* (2005) have run simulated datasets which assess the relative performance of STRUCTURE on inferring clusters given different modes of gene flow between populations. They also tested resolution of the program based on an *ad hoc* statistic which assists estimation of K. The 3 models of gene flow tested are: island, contact and hierarchical island. The *ad hoc* statistic ΔK is calculated according to the formula:

$$\Delta K = \frac{\overline{L''(K)}}{s.d.[L(K)]}$$

Where $\overline{L''(K)}$ is the mean second order rate of change of $\ln\Pr(X|K)$ relative to K over

20 runs, and $s.d.[L(K)]$ is the standard deviation of $\ln\Pr(X|K)$.

ΔK only captures the most major structure in the data, which will correspond to

the highest level of differentiation in non-island models of dispersal (Evanno *et al.*,

2005). Therefore, an hierarchical STRUCTURE analysis is warranted where there may

be further population subdivision (Vaha *et al.*, 2007). This involves finding the most

major clusters in data by pooling all samples into STRUCTURE, and estimating K.

Samples within each K are then subjected to a second round of STRUCTURE analysis,

and so forth until all the subpopulations are identified.

Genetic assignment tests

Assignment tests initially became robust following the methods of Paetkau *et al*.

(1995). The method is based on comparing allele frequencies of individuals to the overall

frequency profile of a reference population, and calculating a log-likelihood value

thereof. This will hereby be referred to as the frequency method. For individuals this is

calculated by first obtaining the expected allele frequencies at each locus for each

reference population. The individual's alleles are then used as a basis in each population

for multiplying across loci and log transforming according to the formula (for 2

alleles/locus):

$$Log(\prod p^2 x \prod 2pq)$$

Note that for missing alleles, a value of 0.01 is used to reduce the likelihood of incorrect assignment, although as long as this value is kept small, its absolute value does not greatly affect results (Paetkau *et al.*, 2004). The value with the highest log likelihood (least negative) is then considered the most likely population of origin for an individual genotype.

Note that the frequency method assumes HWE, random mating and that all reference populations have been sampled. Obviously, obtaining samples from every possible reference population is impossible or difficult in most studies. To deal with this confound, an alternate Bayesian approach was derived by Rannala and Mountain (1997) to calculate assignment probabilities without the assumption that all reference populations have been sampled (Manel *et al.*, 2005). This method will be referred to as the Bayesian method.

For either method, two main analyses may be conducted: a self or mixed-migrant test. Self assignment uses individuals which have pre-defined membership in a putative population and attempts to assign them to populations within the entire dataset. This can be useful for ascertaining the reliability of reference populations in mixed-migrant assignment tests or useful in and of itself for inferring dispersal between populations (Berry *et al.*, 2004). Note that mis-assignment is not necessarily an indicator of first-generation migration between populations. It can be an indicator of past gene flow, homoplasy or incomplete baseline population data. For these self-assignment tests, both methods use a leave-one-out test, which removes each individual from their baseline population and tests it as an unknown against all possible populations (Efron, 1983).

Mixed-migrant assignment takes individuals of unknown origin and attempts to assign them to a given reference population.

A problem with aforementioned assignment tests is that there are no confidence values given for individual assignments. An individual will always be assigned to a given reference population, even if the ultimate likelihood of assignment to any of the reference populations is very small (Hansen *et al.*, 2001). Certain exclusion-based methods have been developed to address this problem (Cornuet *et al.*, 1999; Paetkau *et al.*, 2004). These tests use a Monte Carlo re-sampling procedure which simulates a number of potential individuals based on the genotypes of the given reference populations. A frequency-distribution of likelihood values for each potential genotype is then generated. If the genotype of the individual being tested falls into the tails of the distribution (a pre-defined alpha of 0.01 or 0.001 usually), the individual is excluded from that population of origin. Such tests may then be used in tandem with classical assignment tests to increase robustness of analysis (Manel *et al.*, 2005).

The program GENECLASS 2 (Piry *et al.*, 2004) calculates self and mixed-assignment probabilities by either the frequency or Bayesian methods, and can compute probabilities of exclusion according to the methods of Cornuet *et al.* (1999) and Paetkau *et al.* (2004).

Detecting family structure, linkage disequilibrium, null alleles and genetic bottlenecks

Family structure may be very important in the genetic structuring of salmonid populations (Bentzen *et al.*, 2001). Because the mating system of salmonids can occur between few individuals, it is important that the estimates of allele frequencies in

populations are derived from more than a representative sample of one or few family groups. If a family group with a small number of effective breeders gives rise to highly related offspring, a slight observed heterozygote excess relative to that predicted under HWE will be observed (Luikart and Cornuet, 1999). If the population, however, consists of several of these familial groups, an heterozygote deficiency may be observed ("family Wahlund effect," Castric *et al.*, 2002). To identify exact fullsibs, the program KINSHIP may be used (Goodnight and Queller, 1999). Once these individuals are recognized, they may be removed from subsequent analyses so that they do not bias the overall population data (DeHaan and Arden, 2007).

Linkage disequilibrium (LD) may result when one measured marker is sufficiently close on the chromosome to another marker which is measured or close to a gene under selective pressure. This results in unequal assortment of genotypes as predicted by Mendelian inheritance. Such processes generally occur when a new population is founded, or a population is recently admixed (Mueller, 2004). During meiosis, genetic recombination may occur when aligned chromosomes exchange genetic material. As the frequency of recombination increases and in subsequent generations, the loci approach linkage equilibrium again, thus LD degrades over time. It is therefore important to test for LD in loci measured as presence may be the result of important recent biological events such as assortative mating, admixture or population bottlenecks.

Null alleles are referred to by population genetics as alleles which fail to amplify during the PCR (Chapuis and Estoup, 2007). This often results in scoring of loci as homozygous rather than heterozygous due to laboratory errors. Indirect evidence of null alleles is therefore provided by departures from HWE due to a heterozygote deficiency. It

is important to note though, that detection of null alleles can be the result of perfectly valid biological processes and have nothing to do with scoring errors. Sampling of two distinct populations and classifying them putatively as one can lead to such an heterozygote deficiency, what is referred to as the Wahlund effect (Dakin and Avise, 2004). A high amount of inbreeding may also lead to an heterozygote deficiency (Wright, 1921). Finally, as mentioned earlier, sampling few familial groups in the same population may be detected as an heterozygote deficiency (Castric *et al.*, 2002).

Genetic bottlenecks are the result of an acute reduction in effective population size. Tests can be directly performed to detect recent bottlenecks, or may be inferred from other data. Direct tests, such as done by the program BOTTLENECK, (Piry *et al.*, 1999) look for heterozygosity excess observed relative to heterozygosity expected under mutation-drift equilibrium (Luikart and Cornuet, 1998). This phenomenon occurs because allelic diversity is lost quicker than heterozygosity when a population goes through a bottleneck. Bottlenecks may also leave signatures by LD or HWE departures.

References

Allendorf, F. W., & Luikart, G. (2007). *Conservation and the genetics of populations*. Malden, MA: Blackwell Publishing.

Angers, B., & Bernatchez, L. (1997). Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution, 14*(3), 230-238.

Bahr, M. A., & Shrimpton, J. M. (2004). Spatial and quantitative patterns of movement in large bull trout (*Salvelinus confluentus*) from a watershed in north-western British Columbia, Canada, are due to habitat selection and not differences in life history. *Ecology of Freshwater Fish, 13*(4), 294-304.

Baxter, C. V., & Hauer, F. R. (2000). Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (*Salvelinus confluentus*). *Canadian Journal of Fisheries and Aquatic Sciences, 57*(7), 1470-1481.

Behnke. (2002). *Trout and salmon of North America*. New York, NY: The Free Press.

Bentzen, P., Olsen, J. B., McLean, J. E., Seamons, T. R., & Quinn, T. P. (2001). Kinship analysis of Pacific salmon: Insights into mating, homing, and timing of reproduction. *Journal of Heredity, 92*(2), 127-136.

Berry, D.K. (1997). Alberta's bull trout management and recovery plan. Pp.89-98. *In:* W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology, 13*(3), 551-561.

Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus Fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity, 89*, 27-35.

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

Cavender, T. M. (1978). Taxonomy and distribution of the bull trout, *Salvelinus confluentus* (Suckley) from the American Northwest. *California Fish and Game, 64*, 139-174.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*(3), 621-631.

Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution, 15*(5), 538-543.

Cornuet, J. M., Piry, S., Luikart, G., Estoup, A., & Solignac, M. (1999). New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics, 153*(4), 1989-2000.

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.

Dakin, E. E., & Avise, J. C. (2004). Microsatellite null alleles in parentage analysis. *Heredity, 93*(5), 504-509.

Dembacher, J.M. & K. Jones. (1997). Stream habitat of juvenile bull trout populations in Oregon and benchmarks for habitat quality. Pp 353-360. *In*: M.K. Brewin, A.J. Paul & M. Montina (ed.) Bull Trout II Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

DeHaan, P., & Arden, W. (2007). Genetic analysis of bull trout in the Jarbidge River watershed, Nevada/Idaho. U.S. Fish and Wildlife Service. Reno, NV. 22 pp.

Dunham, J. B., & Rieman, B. E. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications, 9*(2), 642-655.

Earle, J.E & J.S. McKenzie. (2001). Habitat use by juvenile bull trout in mountain streams in the Copton Creek drainage, Alberta, and its relation to mining activity. Pp. 121-128. *In*: M.K. Brewin, A.J. Paul & M. Montina (ed.) Bull Trout II Conference Proceedings, Trout Unlimited Canada, Calgary, AB

Efron, B. (1983). Estimating the error rate of a prediction rule - improvement on cross-validation. *Journal of the American Statistical Association, 78*(382), 316-331.

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

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics, 164*(4), 1567-1587.

Fernet, D.A. & J.O. O'Neil. (1997). Use of radio telemetry to document seasonal movements, and spawning locations for bull trout in relation to a newly created reservoir. Pp. 427-434. *In:* W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Fitch L. (1997). Bull trout in southwestern Alberta: notes on historical and current distribution. Pp. 147-160. *In*: W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Fraley, J. J., & Shepard, B. B. (1989). Life-history, ecology and population status of migratory bull trout (*Salvelinus confluentus*) in the Flathead lake and river system, Montana\. *Northwest Science, 63*(4), 133-143.

Gerking, S. D. (1959). The restricted movement of fish populations. *Biological Reviews of the Cambridge Philosophical Society, 34*(2), 221-242.

Gerrand, M., & Watmough, M. (1998). 1998 Spawning survey of bull trout in the Oldman River drainage. Prepared for Alberta Fish and Wildlife. Alberta Conservation Association. Blairmore, AB. 43 pp.

Goetz, F.A. (1997). Habitat use of juvenile bull trout in Cascade mountain streams of Oregon and Washington. Pp. 339-352. *In:* W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Goodnight, K. F., & Queller, D. C. (1999). Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology, 8*(7), 1231-1234.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity, 86*(6), 485-486.

Gowan, C., Young, M. K., Fausch, K. D., & Riley, S. C. (1994). Restricted movement in stream salmonids - a paradigm lost. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(11), 2626-2637.

Groft, D.G. (1997). DNA fingerprinting of Alberta bull trout (*Salvelinus confluentus*) populations. M.Sc. Thesis. University of Lethbridge, Lethbridge, AB.

Gunckel, S. L., Hemmingsen, A. R., & Li, J. L. (2002). Effect of bull trout and brook trout interactions on foraging habitat, feeding behavior, and growth. *Transactions of the American Fisheries Society, 131*(6), 1119-1130.

Guo, S. W., & Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics, 48*(2), 361-372.

Hansen, M. M., Kenchington, E., & Nielsen, E. E. (2001). Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries, 2*(2), 93-112.

Hauer, F. R., Poole, G. C., Gangemi, J. T., & Baxter, C. V. (1999). Large woody debris in bull trout (*Salvelinus confluentus*) spawning streams of logged and wilderness watersheds in northwest Montana. *Canadian Journal of Fisheries and Aquatic Sciences, 56*(6), 915-924.

Hedrick, P. W. (1999). Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution, 53*(2), 313-318.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution, 59*(8), 1633-1638.

Homel, K., & Budy, P. (2008). Temporal and spatial variability in the migration patterns of juvenile and subadult bull trout in northeastern Oregon. *Transactions of the American Fisheries Society, 137*(3), 869-880.

Homel, K., Budy, P., Pfrender, M. E., Whitesel, T. A., & Mock, K. (2008). Evaluating genetic structure among resident and migratory forms of bull trout (*Salvelinus confluentus*) in Northeast Oregon. *Ecology of Freshwater Fish*(In Press).

Kanda, N., Leary, R. F., & Allendorf, F. W. (2002). Evidence of introgressive hybridization between bull trout and brook trout. *Transactions of the American Fisheries Society, 131*(4), 772-782.

Kennedy, B. P., Klaue, A., Blum, J. D., Folt, C. L., & Nislow, K. H. (2002). Reconstructing the lives of fish using Sr isotopes in otoliths. *Canadian Journal of Fisheries and Aquatic Sciences, 59*(6), 925-929.

Koljonen, M. L., Pella, J. J., & Masuda, M. (2005). Classical individual assignments versus mixture modeling to estimate stock proportions in Atlantic salmon (*Salmo salar*) catches from DNA microsatellite data. *Canadian Journal of Fisheries and Aquatic Sciences, 62*(9), 2143-2158.

Li, Y. C., Korol, A. B., Fahima, T., Beiles, A., & Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology, 11*, 2453-2465.

Louis, E. J., & Dempster, E. R. (1987). An exact test for Hardy-Weinberg and multiple alleles. *Biometrics, 43*(4), 805-811.

Luikart, G., Allendorf, F. W., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity, 89*(3), 238-247.

Luikart, G., & Cornuet, J. M. (1998). Empirical Evaluation of a Test for Identifying Recently Bottlenecked Populations from Allele Frequency Data. *Conservation Biology, 12*(1), 228-237

Luikart, G., & Cornuet, J. M. (1999). Estimating the effective number of breeders from heterozygote excess in progeny. *Genetics, 151*(3), 1211-1216.

Manel, S., Gaggiotti, O. E., & Waples, R. S. (2005). Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution, 20*(3), 136-142.

Mayer, F., & Von Helversen, O. (2001). Sympatric distribution of two cryptic bat species across Europe. *Biological Journal of the Linnean Society, 74*(3), 365-374.

McPhail, J. D. (2007). *The freshwater fishes of British Columbia*. Edmonton, AB: The University of Alberta Press.

McPhail, J. D., & Baxter, J. S. (1996). A review of bull trout (*Salvelinus confluentus*) life-history and habitat use in relation to compensation and improvement

opportunities. University of British Columbia. Fisheris Management Report No. 104. Vancouver, BC. 31 pp.

Moritz, C. (1994). Defining evolutionarily-significant-units for conservation. *Trends in Ecology & Evolution, 9*(10), 373-375.

Mueller, J. C. (2004). Linkage disequilibrium for different scales and applications. *Briefings in Bioinformatics, 5*(4), 355-364.

Nakano, S., Fausch, K. D., Furukawatanaka, T., Maekawa, K., & Kawanabe, H. (1992). Resource utilization by bull char and cutthroat trout in a mountain stream in Montana, USA. *Japanese Journal of Ichthyology, 39*(3), 211-217.

Narum, S. R., Contor, C., Talbot, A., & Powell, M. S. (2004). Genetic divergence of sympatric resident and anadromous forms of *Oncorhynchus mykiss* in the Walla Walla River, USA. *Journal of Fish Biology, 65*(2), 471-488.

Nei, M. (1972). Genetic distance between populations. *American Naturalist, 106*(949), 283-&.

Neigel, J. E. (1997). A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology and Systematics, 28*, 105-128.

Neigel, J. E. (2002). Is F-ST obsolete? *Conservation Genetics, 3*(2), 167-173.

Nelson, M. L., McMahon, T. E., & Thurow, R. F. (2002). Decline of the migratory form in bull charr, *Salvelinus confluentus*, and implications for conservation. *Environmental Biology of Fishes, 64*(1-3), 321-332.

Neraas, L. P., & Spruell, P. (2001). Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology, 10*(5), 1153-1164.

Paetkau, D., Calvert, W., Stirling, I., & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology, 4*(3), 347-354.

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.

Perry, G. M. L., Audet, C., & Bernatchez, L. (2005). Maternal genetic effects on adaptive divergence between anadromous and resident brook charr during early life history. *Journal of Evolutionary Biology, 18*(5), 1348-1361.

Perry, G. M. L., Audet, C., Laplatte, B., & Bernatchez, L. (2004). Shifting patterns in genetic control at the embryo-alevin boundary in brook charr. *Evolution, 58*(9), 2002-2012.

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.

Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity, 90*(4), 502-503.

Post, J.R. & Johnston, F.D. (2002). Status of the Bull Trout (*Salvelinus confluentus*) in Alberta. Alberta Sustainable Resource Development, Fish and Wildlife Division, and Alberta Conservation Association, Wildlife Status Report No. 39. Edmonton, AB. 40 pp. + 4 app.

Potvin, C., & Bernatchez, L. (2001). Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed-stock analyses. *Molecular Ecology, 10*(10), 2375-2388.

Primmer, C. R., Koskinen, M. T., & Piironen, J. (2000). The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. *Proceedings of the Royal Society of London Series B-Biological Sciences, 267*(1453), 1699-1704.

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

Pritchard, J. K., Wen, X., & Falush, D. (2007). Documentation for STRUCTURE software: Version 2.2. Available from http://pritch.bsd.uchicago.edu/software.

Pritchard, V. L., Jones, K., & Cowley, D. E. (2007). Estimation of introgression in cutthroat trout populations using microsatellites. *Conservation Genetics, 8*(6), 1311-1329.

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

Rasmussen, J.B., Trudeau, V., & Morinville, G. (IN PRESS). Estimating the scale of fish feeding movements in rivers using $\delta^{13}C$ signature gradients. *Journal of animal ecology.*

Raymond, M., & Rousset, F. (1995). GENEPOP (version-1.2) - Population-genetics software for exact tests and ecumenicism. *Journal of Heredity, 86*(3), 248-249.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution, 43*(1), 223-225.

Rich, C. F., McMahon, T. E., Rieman, B. E., & Thompson, W. L. (2003). Local-habitat, watershed, and biotic features associated with bull trout occurrence in Montana streams. *Transactions of the American Fisheries Society, 132*(6), 1053-1064.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters, 9*(5), 615-629.

Shepard, B., S. Leathe, T. Weaver, and M. Enk. (1984). Monitoring levels of fine sediment within tributaries to Flathead Lake and impacts of fine sediment on bull trout recruitment. Proc. Wild Trout Symposium Ill, Mammoth Hot Springs, Wyoming.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics, 139*(1), 457-462.

Spruell, P., Hemmingsen, A. R., Howell, P. J., Kanda, N., & Allendorf, F. W. (2003). Conservation genetics of bull trout: Geographic distribution of variation at microsatellite loci. *Conservation Genetics, 4*(1), 17-29.

Spruell, P., Rieman, B. E., Knudsen, K. L., Utter, F. M., & Allendorf, F. W. (1999). Genetic population structure within streams: microsatellite analysis of bull trout populations. *Ecology of Freshwater Fish, 8*(3), 114-121.

Stelfox JD. (1997). Seasonal movements, growth, survival and population status of the adfluvial bull trout population in Lower Kananaskis Lake, Alberta. Pp 309-316. *In:* W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Swanberg, T. R. (1997). Movements of and habitat use by fluvial bull trout in the Blackfoot River, Montana. *Transactions of the American Fisheries Society, 126*(5), 735-746.

Taylor, E. B., Pollard, S., & Louie, D. (1999). Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Molecular Ecology, 8*(7), 1155-1170.

U.S. Fish and Wildlife Service. (1999). Endangered and threatened wildlife and plants. Determination of threatened status for bull trout in the conterminous United States, final rule. FR 64, 58909-58933

Vaha, J. P., Erkinaro, J., Niemela, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology, 16*(13), 2638-2654.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology, 15*(6), 1419-1439.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution, 38*(6), 1358-1370.

Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2004). Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Molecular Ecology, 13*(12), 3675-3688.

Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2006). Can common species provide valuable information for conservation? *Molecular Ecology, 15*(10), 2767-2786.

Whiteley, A. R., Spruell, P., Rieman, B. E., & Allendorf, F. W. (2006). Fine-scale genetic structure of bull trout at the southern limit of their distribution. *Transactions of the American Fisheries Society, 135*(5), 1238-1253.

Wright, S. (1921). Systems of mating. II. The effects of inbreeding on the genetic composition of a population. *Genetics, 6*(2), 124-143.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics, 15*, 323-354.

Appendix
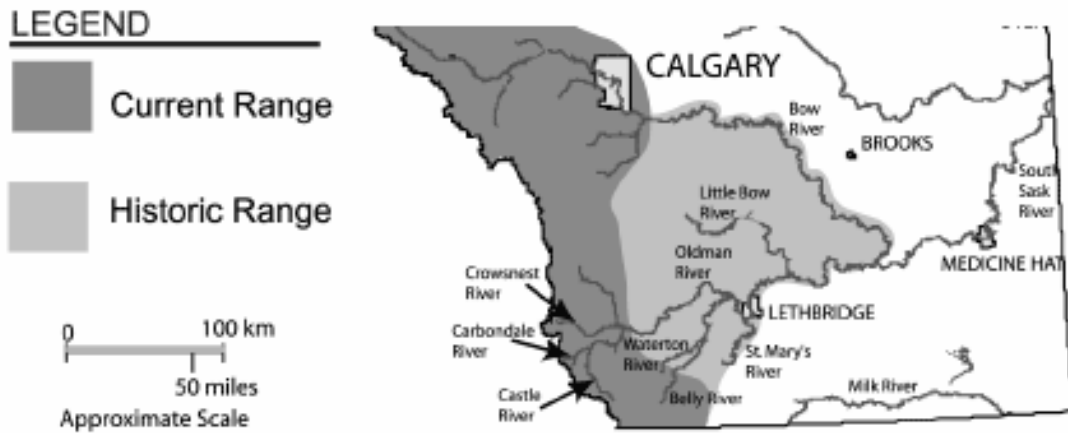
Figure 1-1: Historic and contemporary range of bull trout in Southern Alberta. Image modified from Post and Johnston, 2002.
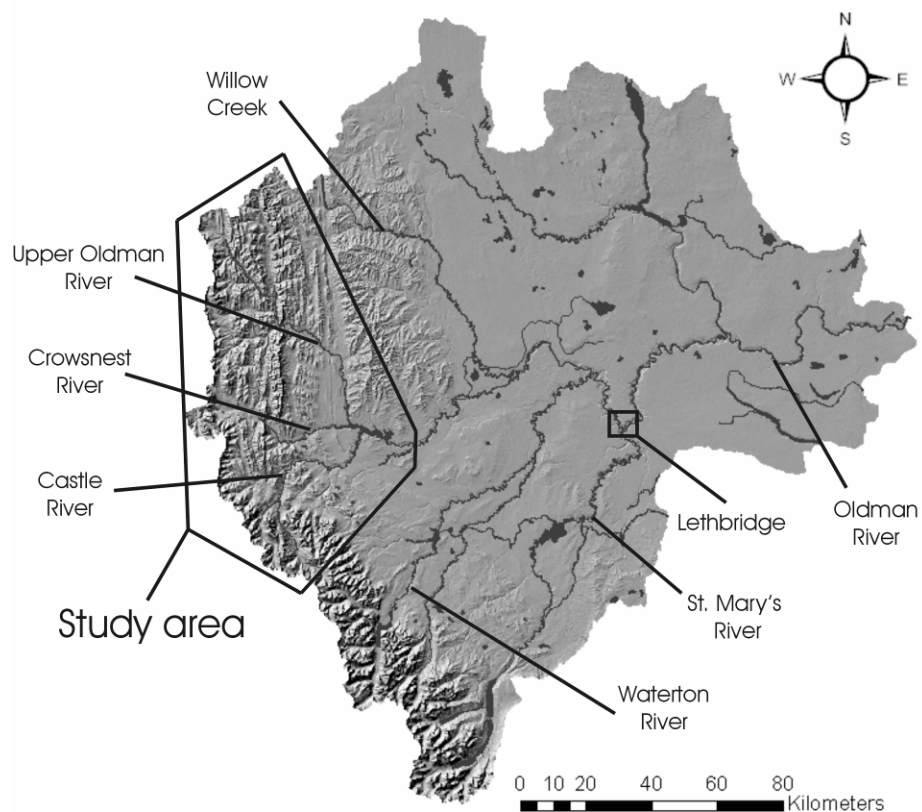
Figure 1-2: Oldman River drainage, with study area including the two sub-basins holding the core bull trout range (Upper Oldman and Castle).

Chapter 2

Genetic clustering methods reveal bull trout (*Salvelinus confluentus*) fine-scale

population structure as a spatially nested hierarchy.

Introduction

Stream-spawning salmonid fishes have long been recognized to display finely

dissected population structure in stream networks, largely due to the strong homing

tendencies of these organisms. In its most extreme manifestation, this population

structure organization would be expected as genetically distinguishable clusters of

individuals found at or near to specific spawning sites. If such population structure

existed, it would be entirely warranted to define populations *a priori*, by site-of-origin.

This has been a classic paradigm of defining populations in conservation genetics studies

of salmonid fishes; such a paradigm may be unrealistic because of two key assumptions:

that each site-of-origin represents a genetically distinguishable unit, and that each

individual sampled has not dispersed from other sites.

Departures from the site-of-origin model using neutral genetic markers are likely

common in real stream ecosystems. Firstly, if organisms commonly exchange genetic

material between sites or if genetic drift has insufficient time to lead to genetic

divergence, these spatially discrete sites should be genetically indistinguishable from one

another. Secondly, site-of-origin does not necessarily reflect population-of-origin. It is

common for stream salmonids to disperse into sites from which they did not originate

(Gowan *et al*., 1994); therefore when an investigator samples a site, it may be composed of immigrants that are inappropriate representatives of the defined population.

An alternate approach to population identification which is not as affected by these confounds involves the use of genetic clustering methods (Pritchard *et al*., 2000; Dawson and Belkhir, 2001; Corander *et al*., 2004; Chen *et al*., 2007). Such methods are based on a genetic or evolutionary framework, as they define populations on the basis of individual genotypes, regardless of capture location (Waples and Gaggiotti, 2006). STRUCTURE (Pritchard *et al*., 2000) is the most common program used for genetic clustering. This clustering algorithm assigns individuals to populations in such a way as to minimize linkage disequilibrium (LD) and departures from Hardy-Weinberg equilibrium (HWE) in a pre-defined cluster value of K. Log-likelihood values for each K value tested are provided, allowing the investigator to ascertain the true number of clusters within the dataset. This approach overcomes the effects of the two confounding factors inherent to a site-of-origin approach in that: 1) collections of individuals from sites with insufficient genetic divergence will be grouped together, and 2) immigrants from other sites are assigned automatically to genetic clusters from whence they originated, not capture location.

Assignment test results may be combined *a posteriori* with spatial data to provide information about movements and gene flow (Manel *et al*., 2005). Gene flow may be indirectly determined by the prevalence of admixture between identified populations. Assignment tests may therefore be powerful tools in answering a wide range of conservation and ecological questions.

An additional advantage of the clustering method used in STRUCTURE is the ability to detect hierarchical population structure. This is because populations may be structured according to an hierarchical island model of gene flow, where genetic "archipelagos" are differentiated from one another, and further population sub-structuring occurs within each archipelago (Giles *et al*., 1998; Waples and Gaggiotti, 2006). The rates of gene flow between populations within each archipelago are higher than rates between archipelagos, but all rates occur in a symmetrical manner (Slatkin and Voelm, 1991) (Figure 2-1). When STRUCTURE is fully utilized in this manner, one can reveal all levels of population structure within the spatial network of the study area; combined with contemporary dispersal and admixture rates, this approach may provide an intuitive guide for a meaningful conservation approach.

In theoretical and applied bull trout conservation genetics literature, populations are always defined by stream-of-origin (Spruell *et al*., 1999; Neraas and Spruell, 2001; Costello *et al*., 2003; Spruell *et al*., 2003; Whiteley *et al*., 2006a; DeHaan *et al*., 2007; Kassler and Mendel, 2007); however, the two aforementioned assumptions of this approach may have occluded interpretations of these studies' results. Firstly, the majority of representative samples used in these studies come from juvenile fish (0-3 yrs age), which are assumed rear in the stream in which they were born. The assumption that juvenile fish remain in their stream-of-origin is consistent with a restricted movement paradigm (Gerking, 1959); however, recent studies have cast doubt on the ubiquity of this model (Gowan *et al*., 1994; Gowan and Fausch, 1996) even in juvenile salmonids (Kennedy *et al*., 2002; Rasmussen *et al*., in press). Secondly, previous studies of bull trout have shown that pairwise F_{ST} values between streams are generally high and

significant, even at fine spatial scales (Whiteley *et al*., 2006a; Dehaan and Arden, 2007; Kassler and Mendel, 2007). It should be noted, however, that F_{ST} is a relative value, and appears high in bull trout likely because of low measured intrapopulation heterozygosity (Hedrick, 1999). Where tributary populations are only composed of few breeders, family structure is also likely to result in an overestimate of significant population divergence (Hansen *et al*., 1997). This is because putative populations may appear as specific family clusters, which may display high pairwise F_{ST} values, and go undetected by HWE departures. These features may have led to an overestimate of meaningful population divergence in bull trout at the stream-of-origin level.

The goals of this chapter are:

1)   to provide an hierarchical description of the population structure of bull trout in Southwest Alberta using a genetic clustering method

2)   to examine contemporary gene flow within the spatial network of the drainage. Such results will provide valuable information for a guided management strategy and provide support for the genetic clustering method as a valuable tool for use in salmonid conservation genetics.

Materials and methods

Study area and sample collection

The Oldman River complex, as described in the introduction, represents the largest complex in the Oldman River drainage. To sample all possible bull trout populations, or sites-of-origin, I backpack electrofished all streams that could possibly

support substantial bull trout populations. Such streams were selected on the basis of accessibility for bull trout migrants, previous sampling literature, and physical stream characteristics such as size and gradient (Dunham and Rieman, 1999). On the basis of those criteria, 21 streams were sampled, with 13 yielding bull trout numbers large enough to possibly support a self-sustaining population (Figure 2-2). The Castle River sub-basin contained 6 streams: the Carbondale River (Cb), Gardiner Creek (Ga), Lost Creek (Lo), the West (Wca) and South (Sca) Castle Rivers, and Mill Creek (Mi). The upper Oldman River sub-basin contained 7 streams: South (Sra) and North (Nra) and the main-stem Racehorse (Ra) Creeks, Dutch Creek (Du), Hidden Creek (Hi), and the Lower (Lli) and Upper (Uli) Livingstone Rivers. The latter two streams were separated on the basis of a seasonally passable set of falls, which may reduce gene flow between and lead to divergence of the two sites.

In 2006 and 2007, streams were sampled in the summer months, and a target of 30 fish per stream was caught. Each fish was weighed, measured and an adipose fin clip was taken and stored in 99% ethanol. In addition, each fish >20mm FL had scales taken for aging to determine if it was an adult resident or migrant. Because streams of migrant bull trout are primarily used for spawning by adults and as nurseries for juveniles (McPhail and Baxter, 1996), juvenile fish make up the bulk of the sample size for the populations. This may lead to 2 possible confounds: high chance of sampling family groups (Hansen *et al*., 1997) and possibly sampling individuals of different origin, as juvenile salmonid movement between tributaries may be a common phenomenon (Kennedy *et al*., 2002). In order to minimize the former, multiple sites within the same

stream were chosen to avoid sampling sibling groups that may occur in clusters of close

spatial proximity (Bentzen, 2001).

DNA extraction, amplification and microsatellite genotyping

DNA was extracted from adipose fins using a QIAGEN DNeasy^{TM} tissue kit

following the recommended procedure in the manual.

All samples were screened for 9 microsatellite loci, chosen based on clarity of

resolution, degree of polymorphism and ability to distinguish bull X brook trout hybrids:

Sco102, Sco216, Smm22, Sfo18, Sco105, Sco106, Sco215, Sco220 and *Omm1128* (See

Costello *et al.*, 2003; DeHaan and Arden, 2005 for details). PCRs were conducted with

fluorescently labeled primers (for PCR procedure, see Costello *et al.*, 2003). PCR

products were then assayed on a Beckman-Coulter CEQ 8000 automated genotyper. Raw

scores were compiled into the program MICROCHECKER to screen for genotyping

errors (Van Oosterhout *et al.*, 2004). Any individuals displaying missing data for more

than 4 alleles were discarded from analysis.

Genetic data analysis

For stream-of-origin

Departures from HWE across loci and streams were tested in the program

GENEPOP 4.0 (Raymond and Rousset, 1995) using Markov chain parameters of 10,000

dememorizations, 100 batches and 5,000 iterations per batch (Guo and Thompson, 1992).

P-values of tests were adjusted using sequential Bonferroni adjustment (Rice, 1989) to

test for significance. Where significance was encountered, GENEPOP was used to identify whether departures were due to heterozygote excess or deficiency.

Basic statistical data on allele frequencies, observed (H_o) and expected (H_e) heterozygosity were calculated in GENALEX 6 (Peakall and Smouse, 2006) across loci and streams. Mean allelic richness (A) per stream was calculated in FSTAT v 2.9.3 (Goudet, 1995) because the program corrects for differences in sample size between given populations. F_{IS} and F_{ST} (θ) were also calculated in FSTAT, which uses the methods outlined in Weir and Cockerham (1984). Additionally, FSTAT was used to test for genotypic (linkage) disequilibrium at all locus pairs within and across sites. To test for significance, p-values were adjusted using the sequential Bonferroni technique described above.

Kassler and Mendel (2007) argue that excluding highly related individuals from samples is not warranted in bull trout studies because bull trout populations are often composed of fullsibs that contribute to the overall reproductive success of the population. In another study, DeHaan (2007) removed fullsibs from analysis because family structure biased populations out of HWE. This is undesirable when conducting further tests which assume HWE within populations (eg, F_{ST}, assignment tests etc.). Because both interpretations may be justified, it is difficult to decide on an appropriate way to analyze the data. Caution must be taken in avoiding results such as the family Wahlund effect (Castric *et al.*, 2002), when family groupings are disproportionately represented in the population due to sampling artifacts and thus may bias the genotypic makeup of the entire population. For this reason, I have chosen to evaluate fullsib presence *a posteriori* to determine whether data was biased due to family structure. Fullsibs were identified in the

program KINSHIP v1.3.1 (Goodnight and Queller, 1999) using 1000 pairwise

simulations and assuming that fullsibs have a relatedness value of 0.5. Family groups are

a large part of salmonid populations (Bentzen, 2001) so it is inevitable that sites will yield

many fullsibs, especially from streams with smaller spawning aggregations.

Finally, population bottlenecks were evaluated by the program BOTTLENECK

(Piry *et al.*, 1999). Although microsatellites largely evolve via a step-wise mutation

model (SMM), occasional mutations result in multi-repeat differences that violate the

stepwise model (Balloux and Lugon-Moulin, 2002). This may be the case for some loci

screened, such as Sfo18, Sco106 and Sco220, which show multi-repeat differences in

allelic patterns across populations (Figure 2-3). To account for this, the program allows

for a two-phase mutation model (TPM), which allows the user to specify the proportion

of stepwise and multi-step mutations within the model. The recommended value of 90%

SMM within the TPM was chosen, which is realistic for microsatellite data (Luikart *et

al.*, 1998). To test for significance, the 1-way Wilcoxon sign-rank test was used, which is

the most powerful in detecting significance with limited sample sizes and/or loci (Piry *et

al.*, 1999).

For Genetic Clustering

The program STRUCTURE 2.2 (Pritchard *et al.*, 2000) was used to infer

population structure by genetic clustering methods. The model parameters of admixture

and correlated allele frequencies were used. These account for recent gene flow between

populations and some flexibility in linkage disequilibrium within populations. Such

settings are most flexible for dealing with real biological phenomena (Pritchard *et al*., 2007).

Recently, Evanno *et al*. (2005) assessed the ability of STRUCTURE to detect population structure according to an hierarchical model. The authors found that the program only captured the major structure in the data, at the archipelago level; however, subsequent STRUCTURE analysis may be performed on each of the identified clusters to find further population structure within each archipelago (e.g., Vaha *et al.*, 2007). I used the "hierarchical STRUCTURE analysis" approach outlined by Vaha *et al*. (2007) to account for varying levels of population structuring within the drainage. A first round of STRUCTURE was conducted with the entire pooled dataset to find the true value of *K*. Values of $\ln \Pr(X|K)$ were estimated for *K*=1 to *K*=13. Because STRUCTURE only captures the major population structure within the sample (Pritchard *et al*., 2007), the true value of *K* was assumed to correspond to major archipelagos within the drainage. The major archipelagos found were then extracted from the dataset and separately analysed in a second round of STRUCTURE analysis.

True values of *K* were estimated by the ΔK method of Evanno *et al*. (2005). This method is an *ad hoc* statistic which aids the researcher in finding the optimal value of *K* across multiple iterations. Where ΔK failed to converge on the true value of *K*, I used the highest mean value of $\ln \Pr(X|K)$ (Pritchard *et al*., 2007; Vaha *et al*., 2007).

A burn-in length of 100,000 with 100,000 MCMC repeats was sufficient to capture major structure in the data, but for subsequent rounds a burn-in length of 300,000-500,000 with 300,000-500,000 MCMC repeats was used due to higher success

in estimating ΔK. All analyses were conducted with 10 iterations per K estimate. To visualize data, the run of highest value for $\ln \Pr(X|K)$ was selected for the true K found.

Values of membership (q) were assessed for all individuals within each genetic cluster found by the program. Individuals were assigned to a population based on their highest q-value. The numbers of mis-assigned individuals were tabulated to determine the rate of mis-assignment in the dataset. These mis-assigned individuals may represent dispersed individuals which were caught in a geographic region different from the "home range" of the assigned population (Berry et al., 2004). To avoid assigning individuals to multiple populations, any individual found in the second round to show membership (q>0.75) to a particular group that had displayed q<0.5 in the previous round was discarded from analysis (Vaha et al., 2007). Individuals showing q<0.75 to any particular population were assumed to be of admixed ancestry, with the highest two population rankings as the ancestral populations of the admixed individual. The threshold value of 0.75 was selected because this is the expected assignment value of an F2 hybrid between two genetically distinguishable populations. Admixture rates were calculated within and between each population as the number of admixed individuals over the number of individuals within the population pair (Manel et al., 2005).

For both rounds of STRUCTURE analysis, clusters found were subject to the same population genetic analyses conducted for the site-of-origin technique. Basic allele frequency data, H_o, H_e, A, F_{IS}, private alleles, HWE, genotypic disequilibrium, null alleles, population bottlenecks and pairwise F_{ST} were calculated for all clusters found.

Other considerations

Mill Creek (Mi) is the only stream in the complex that contains a sympatric population of bull and exotic brook trout. All fish throughout the drainage (including this creek) were screened for brook-trout specific alleles that are identifiable in 5 of the 9 loci used for the study. Any hybrids identified were removed from subsequent analyses.

The migratory tendencies of bull trout in the Oldman River complex are poorly understood. In order to determine whether streams had resident or migrant populations, all potential adult (>20mm FL) bull trout caught were aged by scales, and weighed. Size-at-age for these bull trout was used to determine if the population consisted of migrant, resident or both life histories. If fish of both life histories were found sympatrically, they were separately analysed to determine if genetic differences could be found.

It is unknown what the long-term effects are of impassable dams on bull trout genetics. Yamamoto *et al*. (2004) demonstrated quick change (~30 yrs) in genetic population structuring in white-spotted char following habitat fragmentation due to damming. Their results show loss of genetic diversity and increased population divergence in headwater populations above dams; however, this is likely attributable to the extreme fragmentation of headwaters and drastic change of headwater population life-history from a migratory to a non-migratory form. The effects of the recently constructed dam in this study's watershed are not as extreme, and indeed, in a study of similar scope to mine, there was no evidence of effects from an old (~100 yrs) dam on bull trout genetics (Whiteley *et al*., 2006a). As such, genetic effects of the dam will not be addressed directly in this chapter, but results of this study may be used as a baseline for future studies addressing the genetic effects of the dam.

Results

Intra and inter-stream variation

Basic statistical data on allele frequencies, observed (H_o) and expected (H_e) heterozygosity across loci and streams can be found in Tables 2-1 and 2-2. A summary of results from all pertinent tests can be found in Table 2-3.

Intra-stream variation ranged from $A = 2.6$ and $H_e = 0.409$ in the Upper Livingstone River (Uli) to $A = 5.1$ and $H_e = 0.593$ in the South Castle River (Sca) (Table 2-3). For inter-stream variation, out of 78 pairwise examinations of F_{ST} between streams, all but three were found to be significant (Table 2-3). Absolute F_{ST} values ranged from -0.0077 to 0.3174. Global F_{ST} over all streams was found to be 0.158.

Tests for deviations from HWE yielded a single stream (Wca) falling out of HWE (Table 2-4). Further tests to attribute this to heterozygote deficiency or excess indicated that *Sco216* and *Omm1128* showed high F_{IS} values (Table 2-2), and presence of null alleles (Table 2-3), indications of heterozygote deficiency; however, exact tests for heterozygote deficiency were not significant at any loci (Table 2-3). Only one other stream (Dutch Creek) showed significant heterozygote deficiency at a locus, however, the stream was not significantly out of HWE, as 6 out of 9 loci in fact showed an opposite trend towards heterozygote excess (Tables 2-2 and 2-3).

Evidence for LD was found in a single locus pair in the Upper Livingstone River (Table 2-5). Evidence for population bottlenecks was found in 9 out of the 13 streams at the 0.05 significance level and 5 at the 0.01 significance level (Table 2-3).

Hierarchical population structure

A first round of STRUCTURE analysis revealed $\Delta K = 3$ as the strongest value of number of clusters in the drainage (Figure 2-4). These corresponded geographically to the Castle, Upper Oldman and Livingstone archipelagos (Figure 2-5). A second round of analysis revealed $\Delta K = 2$ within each of these archipelagos, again generally grouping individuals from streams with the highest geographical proximity (Figure 2-5 and Table 2-6). Populations found at this level were the West and Mill (Mi_p) group within the Castle area, the Racehorse (Ra_p) and Hidden (Hi_p) groups within the Upper Oldman area, and the Lower (Lli_p) and Upper (Uli_p) Livingstone groups within the Livingstone area. A third round of analysis only found further structuring within the West Castle group, where $\Delta K = 2$ was found, representing the Carbondale (Cb_p) and West Castle (Wca_p) groups (Table 2-6). The finest population structure found by STRUCTURE therefore consisted of 7 populations, arranged hierarchically within three archipelagos (Figure 2-5). For a visual representation of clustering analysis, see Figure 2-6

Coarsest level of structure: the three archipelagos

Basic statistical data on allele frequencies, observed (H_o) and expected (H_e) heterozygosity across loci and archipelagos can be found in Tables 2-7 and 2-8. A summary of results from all pertinent tests can be found in Table 2-9.

Intra-archipelago variation ranged from $A = 3.6$ and $H_e = 0.437$ in the Livingstone to $A = 7.2$ and $H_e = 0.597$ in the Castle (Table 2-9). For inter-archipelago variation, out of 3 pairwise examinations of F_{ST} between archipelagos, all were found to be significant

(Table 2-9). Absolute F_{ST} values ranged from 0.121 to 0.218. Global F_{ST} over all populations was found to be 0.146.

As expected in subdivided populations, tests for deviations from HWE revealed all archipelagos were out of HWE (Table 2-10). Further tests to attribute this to heterozygote deficiency or excess indicated that 4 loci in the Castle and 1 locus in the Livingstone showed significance for heterozygote deficiency (Table 2-10) This was corroborated in all archipelagos by high F_{IS} values (Table 2-8), and presence of null alleles (Table 2-9).

Evidence for LD was found in all archipelagos, but only in 1-2 locus pairs in each (Table 2-11). Evidence for a population bottleneck was only found in the Oldman archipelago at both the 0.05 and 0.01 significance levels.

Finest level of structure: the 7 populations

Basic statistical data on allele frequencies, observed (H_o) and expected (H_e) heterozygosity across loci and populations can be found in Tables and 2-12 and 2-13. A summary of results from all pertinent tests can be found in Table 2-14.

Intra-population variation ranged from $A = 2.3$ and $H_e = 0.372$ in Uli$_p$ to $A = 6.0$ and $H_e = 0.592$ in Mi$_p$ (Table 2-14). For inter-population variation, out of 21 pairwise examinations of F_{ST} between populations, all were found to be significant (Table 2-14). Absolute F_{ST} values ranged from 0.0782 to 0.3555. Global F_{ST} over all populations was found to be 0.178.

Tests for deviations from HWE yielded a single population (Mi $_p$) falling out of HWE (Table 2-15). Further tests to attribute this to heterozygote deficiency or excess

indicated that 3 loci showed heterozygote deficiency (Table 2-15). This was corroborated by high F_{IS} values (Table 2-13), and presence of null alleles (Table 2-14) at this population.

Evidence for LD was found in a single locus pair in Uli $_p$ and Mi $_p$ (Table 2-16). Evidence for population bottlenecks were found in Ra $_p$, Hi $_p$ and Uli $_p$ at a significance level of 0.05, but only Ra $_p$ and Hi $_p$ at a significance level of 0.01 (Table 2-14).

Family structure was determined by presence of fullsib pairs within each population. The Wca$_p$, Ra$_p$, Lli$_p$ and Uli$_p$ populations displayed strong family structure, while the Cb$_p$ and Hi$_p$ populations showed weak family structure. Family structure in the Mi$_p$ population appeared "patchy," as several unrelated familial groups were found (Figure 2-7)

Admixture within the hierarchically structured drainage

Individual assignments revealed asymmetrical gene flow by admixture. Inter-archipelago admixture rates varied from 2.4% between the Livingstone and Castle archipelagos to 4.5% between the Castle and Oldman archipelagos (Figure 2-8). The Livingstone archipelago had the highest intra-archipelago admixture rate at 35%, while the Oldman had the lowest at 19% (Figure 2-8).

Locus tests

No loci showed significant HWE deviations consistently across populations (Table 2-17). When testing across populations, the locus pair Sco105/Sco2220 showed significance in only 2 out of 7 populations.

Other considerations

A single F1 bull X brook trout hybrid from Mi was identified by being heterozygous for bull and brook trout specific alleles at all 5 hybrid-detecting loci. This individual was excluded from subsequent analyses.

No evidence of sympatric resident and migrant bull trout were found anywhere in the drainage. The Upper Livingstone River had no fish >320mm FL, and many of the fish found were 5 years of age or older (data not shown). In addition, many of the fish between 250 and 320mm FL had secondary sex characteristics typical of sexually mature trout (vibrant coloring, kyped jaws etc.); on the basis of these characteristics, I interpret this population to be a resident one. This life history was not found in any other streams in the study area.

Discussion

Intra and inter-stream variation

By strict definition, nearly all streams-of-origin can be thought of as different management units, or stocks. This is because most show significant genetic divergence at neutral markers (Moritz, 1994). These patterns of population structure at the stream-specific level are similar to those found in previous fine-scale studies of bull trout (Whiteley *et al*., 2006a; Dehaan and Arden, 2007; Kassler and Mendel, 2007). It is unlikely, however, that all these streams contain truly separate populations. Large, significant pairwise F_{ST} is likely inflated by an effect of family structure between streams

(Hansen *et al*., 1997) and low intra-stream heterozygosity (Hedrick, 1999). Furthermore, hierarchical STRUCTURE analysis assignment tests revealed significant mis-assignment rates in the dataset. Because mis-assignment is likely to reflect true dispersal between populations (Berry *et al*., 2004), this introduces a further confound of including immigrant individuals into a population sample. These confounds provide evidence for questioning the stream-of-origin approach, as they violate the two assumptions outlined in the introduction of this chapter. Results of hierarchical STRUCTURE analysis therefore may be used as a more parsimonious alternative approach to determining the population structure in the study area.

Hierarchical population structure

Hierarchical STRUCTURE analysis revealed population structure in an hierarchical manner down to a relatively fine spatial scale. At the coarsest level of population structure, the archipelago level, divergence was relatively high, as shown in the range of pairwise F_{ST} values. The Oldman and Livingstone archipelagos were the most differentiated, which is counter-intuitive due to their geographical proximity. The highest intra-archipelago variation was seen in the Castle, which is likely due to this being the largest of the three populations found at this level of structure. All three archipelagos showed significant HWE departures due to heterozygote deficit, which was expected due to further population structuring (Wahlund effect). This was confirmed by subsequent STURCTURE analysis, which successfully determined population structure within each archipelago.

Within the Oldman and Livingstone archipelagos, 2 populations were found. In the Castle archipelago, 3 populations were found with an additional round of STRUCTURE analysis. All of these populations were significantly different from one another, though pairwise F_{ST} reveals differences between populations which would not be expected under a simple model of isolation-by-distance (IBD). This may be due to relatively recent dramatic drift events or presence of barriers, which may lead to variable patterns of IBD (Costello *et al*., 2003; Taylor *et al*., 2003). All populations conformed to HWE with the exception of Mi_p. This may indicate the existence of further sub-structuring which was not detectable by the analysis, or by the presence of patchy family structure (Figure 2-7) creating a family Wahlund effect in this population; however, given the large geographical distances between the areas, the more likely explanation is further true sub-structuring between the Mill and South Castle groups. The samples appear different in $K = 2$ output from STRUCTURE, though neither ΔK or the highest mean value of $\ln \Pr(X|K)$ suggested that this was the most likely solution (Figure 2-10).

STRUCTURE may have failed to subdivide these groups because of insufficient genetic differentiation between them (Waples and Gaggiotti, 2006). Indeed, when these groups are subdivided, both approach HWE (Table 2-18). It may be warranted in future studies to use multiple clustering programs (eg., Vaha *et al*., 2007; Fedy *et al*., 2008) to explore multiple avenues of solutions for correctly determining population structure.

Genetic bottlenecks were only found in the Hi_p and Ra_p populations using a threshold of α = 0.01. This is consistent with the strong trends of bottleneck detection seen in these areas at both the stream specific and archipelago levels of structure. There are two likely phenomena which may explain a true bottleneck occurring in these

populations. First, angling pressure may have historically reduced the numbers of migratory bull trout far more in the upper Oldman area than in the Castle area (Fitch, 1997). Because anglers would selectively target large-bodied migrants for harvest, this explains why no bottleneck is seen in the Livingstone archipelago, which is found in the same area but contains highly resident fish. Secondly, the Gap falls are a seasonal migration barrier for fish in the main-stem of the Oldman River found below all tributaries the Ra_p and Hi_p populations use for spawning. Because bull trout migrate in the later summer months when stream flows are reduced (McPhail and Baxter, 1996), this barrier may create a natural limit on numbers of returning spawners to these migrant populations. This may cause a perpetual population bottleneck, which may vary in severity depending on river flows during the upstream migration run of sexually mature adults. Note that these two reasons may not be mutually exclusive, and may have interacted in the past century to create the bottleneck which has been observed.

Linkage disequilibrium was observed at the same locus pair in both the Mi_p and Uli_p populations. Because no other locus pairs showed this trend, the result was interpreted as biologically irrelevant.

Admixture

Inter-archipelago admixture was highest between the Oldman and Castle groups. This is corroborated by the lowest pairwise F_{ST} value that is found between archipelagos. Conversely, admixture rates were lowest between the Castle and Livingstone groups, which did not show the highest pairwise F_{ST} value. Further discontinuities were found within archipelagos, as low or high values of F_{ST} could not be intuitively predicted by

admixture rates. I would therefore suggest caution in interpreting admixture values from STRUCTURE as quantifiable indicators of gene flow between populations. There are several reasons why this may be the case, as will be discussed below.

F_{ST} values represent historic divergence values between given populations, while admixture rates from assignment tests only indicate a contemporary snapshot of connectivity processes (Palstra *et al*., 2007). Additionally, admixture rates obtained by this paper are based on number of admixed individuals found relative to sample size between two populations. Because population sizes are not equal in the drainage (i.e. sample sizes do not reflect population sizes), admixture rates are not likely to affect each population in an equivalent manner. For example, although relatively little admixture occurs between the Livingstone and Castle archipelagos, it is possible that the effect on Livingstone genetics would be far greater than that in the Castle. This is because the overall population size of the Livingstone is much smaller and less diverse than the Castle, which is more resistant to gene flow and drift events (high N_e, heterozygosity and allelic richness).

If historic gene flow may not be determined based on contemporary admixture rates found by STRUCTURE, what use do such measurements have? In terms of genetics, the picture of gene flow shows us that contemporary processes deviate from Wright's (1959) original simple hierarchical island model (Figure 2-1). This is because connectivity, measured by admixture rate, occurs in an asymmetrical, and possibly stepping-stone manner within and between archipelagos (Figure 2-8). This contemporary picture may tell us the current gene flow dynamics in the system, and may allow an

opportunity for continued monitoring to explore how gene flow occurs as a dynamic

process in the nested hierarchy of the river system.

As previously mentioned, the assignment tests used not only have the potential to

measure admixture, but to measure dispersal rates as well. For the purposes of this study,

the presence of dispersal found due to mis-assignment indicates an advantage of the

clustering method over the site-of-origin technique; however, these assignment tests may

be used in and of themselves to describe migratory dispersal tendencies (Berry *et al.*,

2004). The utility of such techniques will be explored thoroughly in Chapter 4.

Conservation implications and recommendations

Hierarchical STRUCTURE analysis revealed levels of population structure within

the drainage system. At what level does one draw the line for management? Major

clusters defined by the first round of STRUCTURE analysis are likely a good starting

point, as it appears that gene flow is likely very low between the three major

archipelagos; however, significant subdivision within each of these archipelagos is

present and may encompass some adaptive genetic differences due to local adaptation

(Taylor 1991). Therefore, it could be argued that all levels of hierarchical population

structure detectable by this program reveal important information for determining

conservation strategies (Pearse and Crandall, 2004; Whiteley *et al.*, 2006b). The coarsest

level of population structure determined in this study identified major groupings with

unique evolutionary trajectories, between which contemporary gene flow is minimal. The

finest level of population structure revealed locally adapted populations that may warrant

special consideration as appropriate management units.

The large diversity of the Castle archipelago implies that this area is the least affected by drift, and may be more resistant to disturbances or extinction events than the other two (Dunham and Rieman, 1999). This may be attributable to the higher population size, and lower anthropogenic impacts on this system (Fitch, 1997). Within this archipelago, the Mi_p population is the most diverse, which is likely attributable to its robust spawning run of highly fluvial fish (Gerrand and Watmough, 1998; Golder, 1998). All attempts possible should be made to maintain the strong spawning runs of such large, important populations. The main spawning stream in this case is Mill Creek, which should be managed in order to minimize obstructions or fragmenting influences that can restrict upstream or downstream movements of either adults or juveniles. Although brook trout were introduced to Mill Creek over 60 years ago (Wig, personal communication), bull trout appear to remain the numerically dominant species (Warnock, unpublished data) and hybridization between the two species was found to be very rare within this area. Such resistance to the biological effects of brook trout invasion is in stark contrast to devastating effects commonly reported in other systems (Gunckel *et al*., 2002; Paul *et al*., 2003). Anthropogenic disturbances that reduce habitat complexity may increase the risk of brook trout invasiveness (Rich *et al*., 2003); therefore high habitat complexity within the stream must be maintained in order to maximize the resistance of this important bull trout population to brook trout invasion. The fall angling closure in this stream is warranted, and should be maintained to protect the spawning events that occur.

The Wca_p population has the lowest diversity within the archipelago and is the most genetically divergent, which may be partially explained by the location of the spawning stream above a seasonal migration barrier, with a subsequent seasonal

migration barrier located approximately ½ way up the stream itself (Costello *et al*., 2003; Whiteley *et al*., 2006a). Although the population is not considered resident, it is possible that these barriers serve to reduce population size by perpetually reducing the number of migrants on a yearly basis. Because of this low diversity, small effective population size and difficult access to spawning areas, this population is likely the most sensitive to extinction events in the archipelago (Hedrick and Kalinowski, 2000). Protection of critical spawning habitats in such small systems is essential. Critical spawning habitats occur in both the upper section (Gerrand and Watmough, 1998) and in the lower section of the river, immediately downstream from the subsurface flow area (Warnock, personal observations).

The Oldman archipelago is second to the Castle with respect to genetic diversity. Within the archipelago, the Hi_p population is the most diverse, which is likely attributable to its much larger spawning population size (Gerrand and Watmough, 1998; Blackburn, personal communication). It appears that Hidden Creek is the main source tributary for this population, and thus would be the region of highest conservation priority to maintain the central spawning aggregation. A fall closure to angling in this stream, similar to that of Mill Creek may be warranted to protect these spawning events. Despite the high census population size, genetic diversity is still lower than that observed in the Cb_p and Mi_p populations of the Castle archipelago. This may be due to the genetic bottleneck observed, as bottlenecks cause a loss of allelic diversity and - to a lesser extent - heterozygosity (Hedrick and Miller, 1992). Large current population size indicates that a rapid recovery of population size has taken place, but the loss of neutral genetic diversity may not necessarily be indicative of a sharp loss of fitness. Such a phenomenon has been

observed in the Northern elephant seal, which experienced a severe reduction in population size in the late 1800s -early 1990s (20-100 individuals) to a population size of 175,000 today (Weber *et al.*, 2000). Nevertheless, while numbers may have recovered sufficiently following the bottleneck event, the population should continue to be monitored and protected.

The Ra_p population appears to be highly divergent from other populations in the drainage and has low intrapopulation diversity. Two possible reasons for this are high degree of isolation of the spawning population and/or severe recent population bottlenecks, greatly increasing the effects of drift by reducing effective population size. The presence of a seasonal barrier on both the North and South Racehorse streams give evidence for the former, while the high significance value of a test for a population bottleneck gives evidence for the latter. As mentioned earlier, the cause for the bottleneck may have been caused or exacerbated by the presence of such seasonal barriers, perpetually reducing breeding numbers for the populations by frequently cutting off access to spawning grounds. Management practices should focus on a recovery plan for this population, as it may be sensitive to extinction events (Dunham and Rieman, 1999). A survey of the spawning locations should take place to determine how much spawning, if any, may occur below the seasonally passable migration barriers. If little spawning occurs, artificial augmentation of spawning grounds (e.g., Merz *et al*., 2004) may be warranted below these barriers to augment population recovery or if a long drought period elapses which may extirpate the local population.

The Livingstone archipelago is highly divergent, while displaying the least intra-archipelago diversity of the three, indicating that effects of drift are pronounced in this

area. The Uli_p is the least diverse and most divergent within this archipelago. This is attributable to the presence of a seasonal barrier and resident life-history. The barrier reduces immigration into the Uli_p region, while residency reduces outmigration of Uli_p fish. This represents the only area in the drainage where a non-migratory, non-piscivorous population of bull trout naturally co-exists with a pure (non-hybridized) Westslope cutthroat trout population (Robinson, 2007). Such an occurrence is likely rare in Rocky Mountain streams and may be useful as a model system for future studies investigating ecological interactions between these two species. The unique ecological features of such populations likely result in different selective pressures from migrant populations, which face different challenges (Taylor, 1991); therefore, neutral genetic divergence measured for this study may mirror a high degree of adaptive divergence within this population. Genetic challenges may also be unique to resident populations, which may have to deal with effects of inbreeding depression and low diversity (Hedrick and Kalinowski., 2000). Although our results do not provide evidence of high inbreeding rates via F_{IS} values, this population had the highest rate of fullsib occurrence, indicating high likelihood of inbreeding. Low diversity was observed, but there are many examples of small populations with low diversity that continue to persist, even in the face of novel challenges (Amos and Balmford, 2001). Conservation efforts should aim to maintain this population in its natural state.

The Lli_p is the most diverse population in the group, but still well below the diversity observed in most other populations in the drainage. Large-bodied fish occur frequently in the Livingstone river below the falls (Blackburn, personal communication) and a large number of redds were observed in the area in 1997 (Gerrand and Watmough,

1997), so it is assumed that the life-history of this population is indeed migratory. If this population is migratory and of moderate size, what is preventing it from genetically converging with adjacent populations in the Oldman archipelago, increasing intrapopulation diversity? Assuming assignment values from STRUCTURE are correct, this likely results from the high degree of admixture observed with the Uli_p population. The low diversity of the Uli_p may then be acting as a genetic "drain" which reduces the diversity within the Lli_p population. Because the Uli_p population is more divergent from other populations, the Lli_p population may act as a stepping-stone intermediate by which the entire archipelago is linked to the other two major genetic clusters in the drainage. It may be, then, that the Uli_p is not completely isolated from long-range gene flow, as it may only receive gene flow via a proxy population rather than directly from other sources. The importance of conserving of the Lli_p may then be additionally due to its value as a unique genetic intermediate between a resident and migrant population. Furthermore, this population may act as a genetic "rescue" population for the Uli_p population, providing relief from effects of inbreeding depression and/or low diversity discussed in the preceding paragraph. Characterization of the interaction between these two populations should continue in future studies.

Conclusions

When populations were defined by site-of-origin, results of this study were largely consistent with previous fine-scale population genetic studies on bull trout (Whiteley *et al*., 2006a; Dehaan and Arden, 2007; Kassler and Mendel, 2007). The logic

used in this approach, however, makes assumptions about the spatial and genetic patterns

of populations which are likely to be violated in real ecosystems. Defining populations by

a genetic clustering method, however, may represent a more parsimonious approach to

ascertaining population structure in the spatial network of a river system. The population

structure revealed by STRUCTURE adheres to a modified hierarchical island model in

which patterns of contemporary gene flow are complex. Population structure found by

this method is not as strong as may be interpreted by site-of-origin methods, which may

be biased by inflated F_{ST} values and dispersal between populations. It is, of course,

possible that STRUCTURE has failed to determine all the population structure present in

the drainage (Waples and Gagiotti, 2006) or has provided inappropriate assignment

values; however, because results are largely interpretable and intuitive given a posteriori

geographic and ecological information, it appears that STRUCTURE has merit in

identifying true population structure and proper units of conservation at fine spatial scales

for salmonid fishes.

Overall, our results appear generally consistent with findings of other authors that

distal headwater populations (Vaha *et al*., 2008) and populations isolated above barriers

(Costello *et al*., 2003; Whiteley *et al*., 2006a) display reduced intrapopulation diversity

and increased interpopulation divergence. I suggest added complexity to the latter,

however, by interpreting most barriers as not all-or-none migration impediments, but as

semi-permeable genetic filters which may operate in a temporally variable manner.

References

Amos, W., & Balmford, A. (2001). When does conservation genetics matter? *Heredity, 87*, 257-265.

Balloux, F., & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology, 11*(2), 155-165.

Bentzen, P., Olsen, J. B., McLean, J. E., Seamons, T. R., & Quinn, T. P. (2001). Kinship analysis of Pacific salmon: Insights into mating, homing, and timing of reproduction. *Journal of Heredity, 92*(2), 127-136.

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology, 13*(3), 551-561.

Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus Fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity, 89*, 27-35.

Chen, C., Durand, E., Forbes, F., & Francois, O. (2007). Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes, 7*(5), 747-756.

Corander, J., Waldmann, P., Marttinen, P., & Sillanpaa, M. J. (2004). BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics, 20*(15), 2363-2369.

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics, 144*(4), 2001-2014.

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.

Dawson, K. J., & Belkhir, K. (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetical Research, 78*(1), 59-77.

DeHaan, P., & Arden, W. (2007). Genetic analysis of bull trout in the Jarbidge River watershed, Nevada/Idaho. U.S. Fish and Wildlife Service. Reno, NV. 22 pp.

Dehaan, P. W., & Ardren, W. R. (2005). Characterization of 20 highly variable tetranucleotide microsatellite loci for bull trout (Salvelinus confluentus) and cross-amplification in other Salvelinus species. *Molecular Ecology Notes, 5*(3), 582-585.

Dunham, J. B., & Rieman, B. E. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications, 9*(2), 642-655.

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

Fedy, B. C., Martin, K., Ritland, C., & Young, J. (2008). Genetic and ecological data provide incongruent interpretations of population structure and dispersal in naturally subdivided populations of white-tailed ptarmigan (*Lagopus leucura*). *Molecular Ecology, 17*(8), 1905-1917.

Fitch L. (1997). Bull trout in southwestern Alberta: notes on historical and current distribution. Pp. 147-160. *In*: W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Gerrand, M., & Watmough, M. (1998). 1998 Spawning survey of bull trout in the Oldman River drainage. Prepared for Alberta Fish and Wildlife. Alberta Conservation Association. Blairmore, AB. 43 Pp.

Giles, B. E., Lundqvist, E., & Goudet, J. (1998). Restricted gene flow and subpopulation differentiation in *Silene dioica*. *Heredity, 80*, 715-723.

Golder, & Associates. (1998). The Oldman River bull trout radio telemetry study 1993-1996. Prepared by Golder Associates Limited. Calgary, AB. 31 pp. + 2 app.

Goodnight, K. F., & Queller, D. C. (1999). Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology, 8*(7), 1231-1234.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity, 86*(6), 485-486.

Gowan, C., Young, M. K., Fausch, K. D., & Riley, S. C. (1994). Restricted movement in stream salmonids - a paradigm lost. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(11), 2626-2637.

Gowan, C., & Fausch, K. D. (1996). Mobile brook trout in two high-elevation Colorado streams: Re-evaluating the concept of restricted movement. *Canadian Journal of Fisheries and Aquatic Sciences, 53*(6), 1370-1381.

Guo, S. W., & Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics, 48*(2), 361-372.

Hansen, M. M., Nielsen, E. E., & Mensberg, K. L. D. (1997). The problem of sampling families rather than populations: Relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. *Molecular Ecology, 6*(5), 469-474.

Hedrick, P. W. (1999). Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution, 53*(2), 313-318.

Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics, 31*, 139-162.

Hedrick, P. W., & Miller, P. S. (1992). Conservation genetics - techniques and fundamentals. *Ecological Applications, 2*(1), 30-46.

Kassler, T., & Mendel, G. (2007). Genetic characterization of bull trout from the Walla Walla River basin. Washington Dept. of Fish and Wildlife. Olympia, WA. 42 pp. + app.

Kennedy, B. P., Klaue, A., Blum, J. D., Folt, C. L., & Nislow, K. H. (2002). Reconstructing the lives of fish using Sr isotopes in otoliths. *Canadian Journal of Fisheries and Aquatic Sciences, 59*(6), 925-929.

Luikart, G., Allendorf, F. W., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity, 89*(3), 238-247.

Manel, S., Gaggiotti, O. E., & Waples, R. S. (2005). Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution, 20*(3), 136-142.

McPhail, J. D., & Baxter, J. S. (1996). A review of bull trout (*Salvelinus confluentus*) life-history and habitat use in relation to compensation and improvement opportunities. University of British Columbia. Fisheris Management Report No. 104. Vancouver, BC. 31 pp.

Merz, J. E., Setka, J. D., Pasternack, G. B., & Wheaton, J. M. (2004). Predicting benefits of spawning-habitat rehabilitation to salmonid (*Oncorhynchus* spp.) fry production in a regulated California river. *Canadian Journal of Fisheries and Aquatic Sciences, 61*(8), 1433-1446.

Moritz, C. (1994). Defining evolutionarily-significant-units for conservation. *Trends in Ecology & Evolution, 9*(10), 373-375.

Neraas, L. P., & Spruell, P. (2001). Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology, 10*(5), 1153-1164.

Palstra, F. P., O'Connell, M. F., & Ruzzante, D. E. (2007). Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. *Molecular Ecology, 16*(21), 4504-4522.

Paul, A. J., Post, J. R., & Stelfox, J. D. (2003). Can anglers influence the abundance of native and nonnative salmonids in a stream from the Canadian Rocky Mountains? *North American Journal of Fisheries Management, 23*(1), 109-119.

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

Pearse, D. E., & Crandall, K. A. (2004). Beyond F-ST: Analysis of population genetic data for conservation. *Conservation Genetics, 5*(5), 585-602.

Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity, 90*(4), 502-503.

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

Pritchard, J. K., Wen, X., & Falush, D. (2007). Documentation for STRUCTURE software: Version 2.2. Available from http://pritch.bsd.uchicago.edu/software.

Rasmussen, J.B., Trudeau, V., & Morinville, G. (in press). Estimating the scale of fish feeding movements in rivers using d13C signature gradients. *Journal of animal ecology*

Raymond, M., & Rousset, F. (1995). GENEPOP (version-1.2) - population-genetics software for exact tests and ecumenicism. *Journal of Heredity, 86*(3), 248-249.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution, 43*(1), 223-225.

Rich, C. F., McMahon, T. E., Rieman, B. E., & Thompson, W. L. (2003). Local-habitat, watershed, and biotic features associated with bull trout occurrence in Montana streams. *Transactions of the American Fisheries Society, 132*(6), 1053-1064.

Robinson, M. D. (2007). The ecological consequences of hybridization between native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and introduced rainbow

trout (*O. mykiss*) in South Western Alberta. M.Sc. Thesis. University of Lethbridge, Lethbridge, AB.

Slatkin, M., & Voelm, L. (1991). Fst in a hierarchical island model. *Genetics, 127*(3), 627-629.

Spruell, P., Hemmingsen, A. R., Howell, P. J., Kanda, N., & Allendorf, F. W. (2003). Conservation genetics of bull trout: Geographic distribution of variation at microsatellite loci. *Conservation Genetics, 4*(1), 17-29.

Spruell, P., Rieman, B. E., Knudsen, K. L., Utter, F. M., & Allendorf, F. W. (1999). Genetic population structure within streams: microsatellite analysis of bull trout populations. *Ecology of Freshwater Fish, 8*(3), 114-121.

Taylor, E. B. (1991). A review of local adaptation in salmonidae, with particular reference to pacific and atlantic salmon. *Aquaculture, 98*(1-3), 185-207.

Taylor, E. B., Stamford, M. D., & Baxter, J. S. (2003). Population subdivision in westslope cutthroat trout (*Oncorhynchus clarki lewisi*) at the northern periphery of its range: evolutionary inferences and conservation implications. *Molecular Ecology, 12*(10), 2609-2622.

Vaha, J. P., Erkinaro, J., Niemela, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology, 16*(13), 2638-2654.

Vaha, J. P., Erkinaro, J., Niemela, E., & Primmer, C. R. (2008). Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. *Evolutionary Applications, 1*(1), 137-154.

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

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology, 15*(6), 1419-1439.

Weber, D. S., Stewart, B. S., Garza, J. C., & Lehman, N. (2000). An empirical genetic assessment of the severity of the northern elephant seal population bottleneck. *Current Biology, 10*(20), 1287-1290.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution, 38*(6), 1358-1370.

Whiteley, A. R., Spruell, P., Rieman, B. E., & Allendorf, F. W. (2006a). Fine-scale genetic structure of bull trout at the southern limit of their distribution. *Transactions of the American Fisheries Society, 135*(5), 1238-1253.

Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2006b). Can common species provide valuable information for conservation? *Molecular Ecology, 15*(10), 2767-2786.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics, 15*, 323-354.

Yamamoto, S., Morita, K., Koizumi, I., & Maekawa, K. (2004). Genetic differentiation of white-spotted charr (*Salvelinus leucomaenis*) populations after habitat fragmentation: Spatial-temporal changes in gene frequencies. *Conservation Genetics, 5*(4), 529-538.
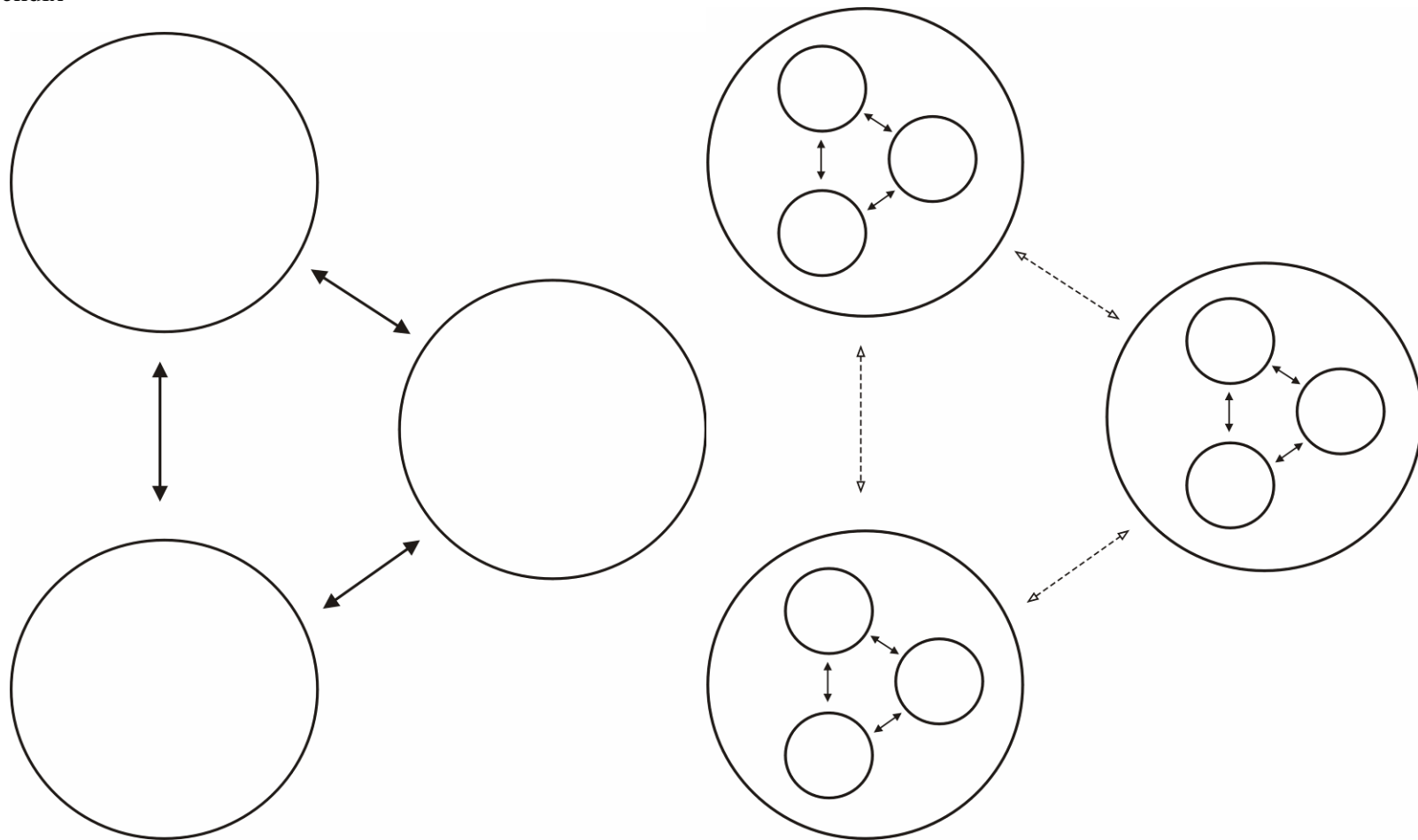
Figure 2-1: Gene flow according to Wright's original (left) and hierarchical (right) island model. In the hierarchical model, solid lines indicate higher rates of gene flow, while dashed lines indicate lower rates.
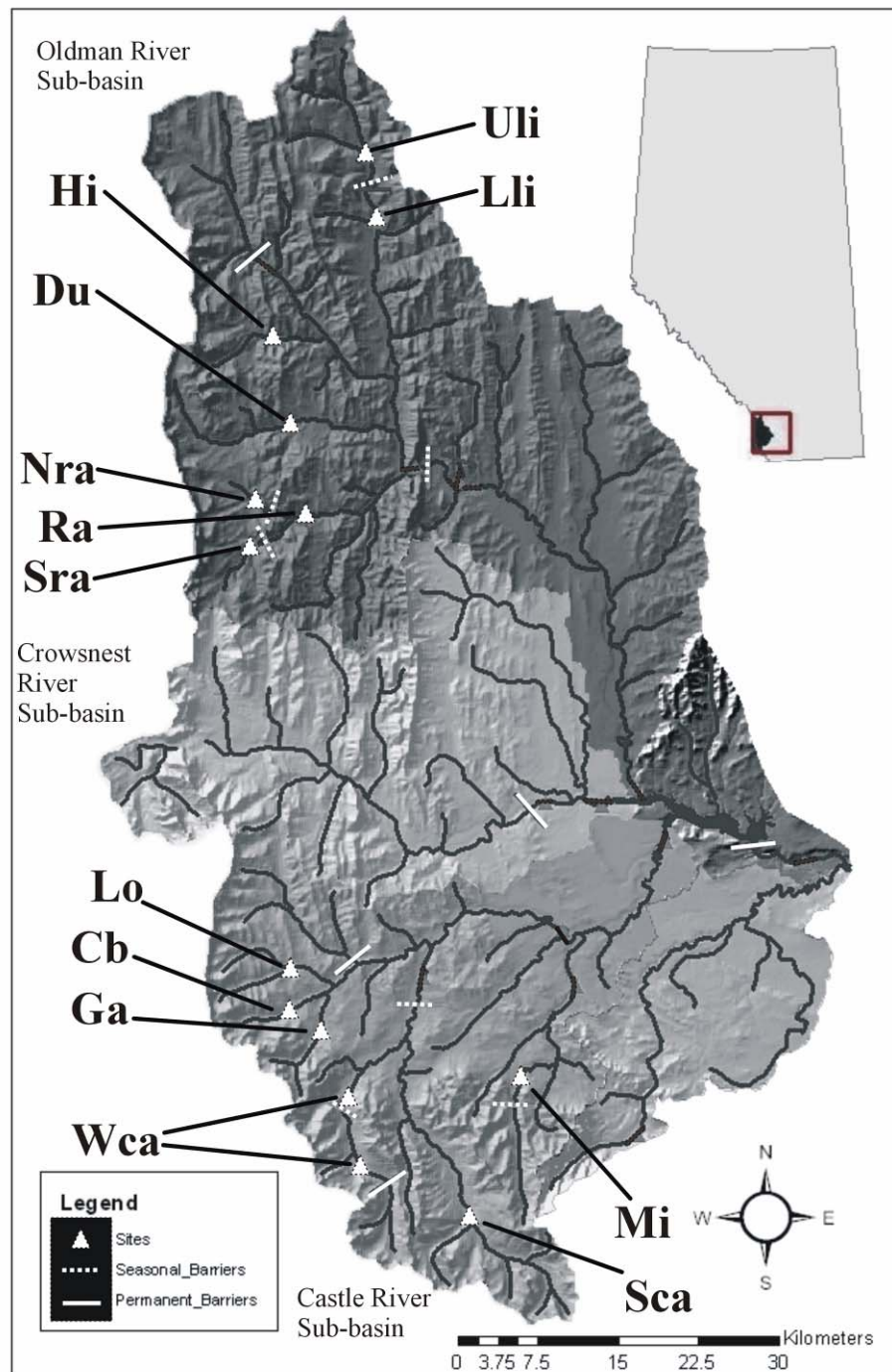
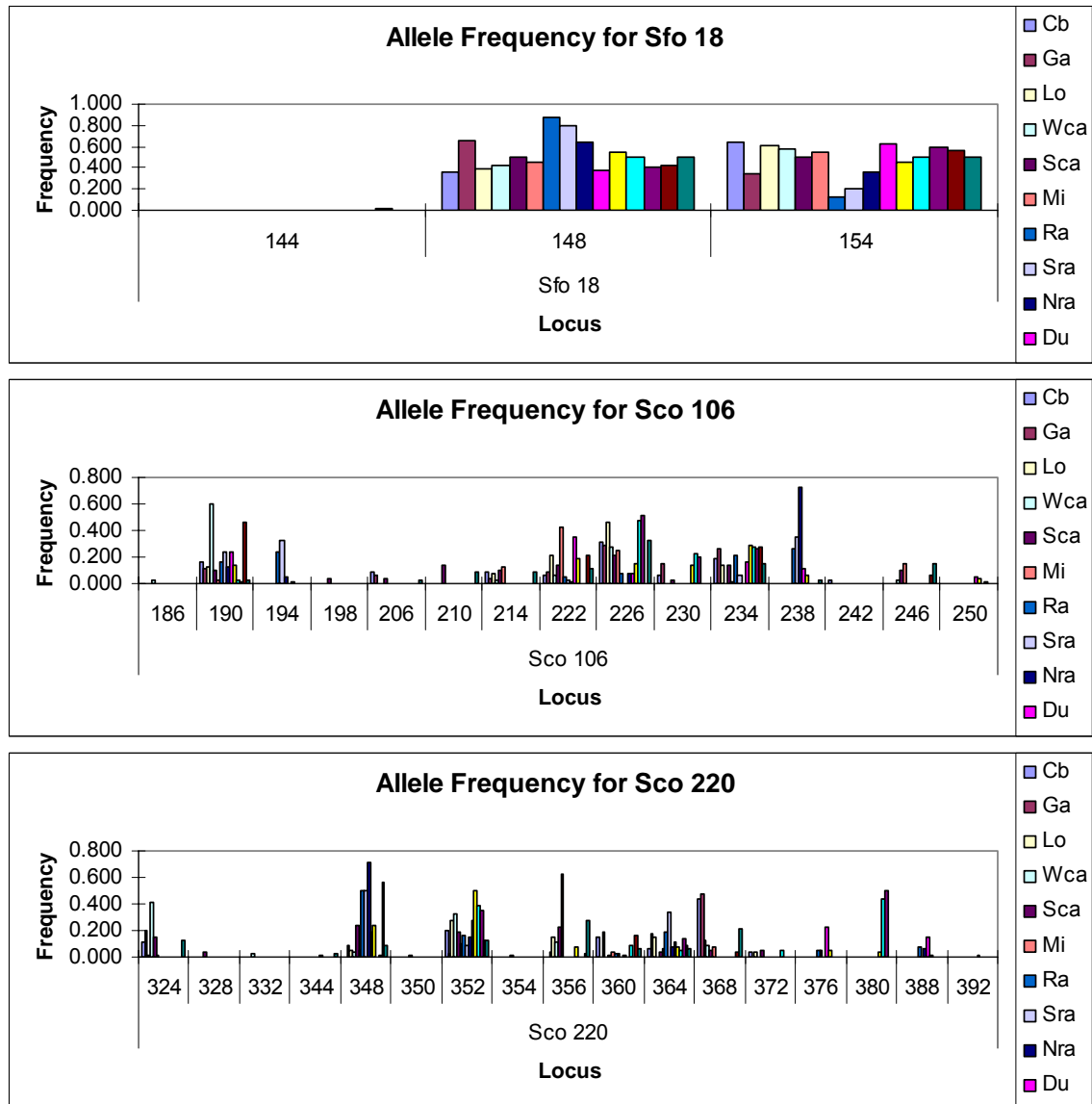Figure 2-2: Map of streams-of-origin in the Oldman River complex.

Figure 2-3: Allelic patterns for microsatellite loci suspected of deviating from stepwise model of mutation.

Locus	Allele	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli
Sco_102	168	0.630	0.357	0.483	0.260	0.283	0.357	0.740	0.794	0.600	0.741	0.833	0.696	0.667
	172	0.326	0.595	0.517	0.700	0.567	0.589	0.260	0.206	0.400	0.259	0.167	0.304	0.333
	176	0.043	0.048	0.000	0.040	0.150	0.054	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sco_216	222	0.286	0.091	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	226	0.286	0.523	0.346	0.750	0.466	0.339	0.000	0.028	0.000	0.204	0.081	0.679	0.414
	230	0.375	0.364	0.596	0.154	0.362	0.565	1.000	0.972	1.000	0.796	0.887	0.321	0.571
	234	0.054	0.023	0.058	0.058	0.155	0.065	0.000	0.000	0.000	0.000	0.032	0.000	0.014
	238	0.000	0.000	0.000	0.000	0.017	0.032	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Smm_22	213	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.000	0.000	0.000	0.000	0.000	0.000
	217	0.000	0.000	0.000	0.000	0.000	0.017	0.100	0.156	0.180	0.037	0.033	0.000	0.000
	221	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014
	225	0.000	0.000	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.014
	229	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.083	0.534	0.681
	233	0.104	0.022	0.172	0.000	0.034	0.121	0.000	0.000	0.000	0.000	0.100	0.121	0.264
	237	0.167	0.022	0.052	0.042	0.086	0.017	0.000	0.000	0.000	0.000	0.000	0.086	0.000
	241	0.021	0.000	0.172	0.000	0.052	0.138	0.000	0.000	0.000	0.019	0.000	0.000	0.000
	245	0.000	0.000	0.000	0.000	0.017	0.000	0.040	0.063	0.080	0.000	0.000	0.017	0.000
	249	0.063	0.022	0.121	0.000	0.017	0.017	0.080	0.125	0.020	0.111	0.017	0.190	0.000
	253	0.083	0.000	0.017	0.063	0.017	0.052	0.000	0.000	0.000	0.000	0.000	0.000	0.014
	257	0.104	0.152	0.103	0.104	0.103	0.103	0.260	0.219	0.280	0.204	0.100	0.034	0.000
	261	0.083	0.022	0.052	0.021	0.241	0.000	0.020	0.063	0.040	0.000	0.000	0.000	0.000
	265	0.042	0.043	0.034	0.021	0.172	0.172	0.040	0.000	0.040	0.148	0.000	0.000	0.000
	269	0.000	0.000	0.017	0.125	0.103	0.086	0.080	0.063	0.000	0.074	0.133	0.000	0.000
	273	0.000	0.000	0.000	0.000	0.017	0.052	0.320	0.250	0.320	0.148	0.133	0.000	0.000
	277	0.000	0.022	0.000	0.021	0.000	0.121	0.040	0.063	0.000	0.000	0.333	0.000	0.014
	281	0.021	0.087	0.000	0.083	0.000	0.052	0.000	0.000	0.020	0.056	0.000	0.000	0.000
	285	0.000	0.022	0.017	0.063	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	289	0.042	0.174	0.000	0.146	0.034	0.017	0.000	0.000	0.000	0.019	0.000	0.000	0.000
	293	0.000	0.043	0.017	0.167	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	297	0.000	0.022	0.000	0.042	0.017	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	301	0.000	0.043	0.000	0.021	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	305	0.000	0.000	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	309	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	313	0.042	0.022	0.052	0.000	0.017	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	317	0.021	0.043	0.052	0.000	0.000	0.000	0.000	0.000	0.000	0.037	0.000	0.000	0.000
	321	0.104	0.130	0.017	0.042	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	325	0.063	0.043	0.034	0.000	0.000	0.000	0.000	0.000	0.020	0.093	0.067	0.000	0.000
	329	0.042	0.043	0.017	0.042	0.000	0.000	0.000	0.000	0.000	0.019	0.000	0.000	0.000
	333	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.037	0.000	0.000	0.000
Sfo_18	148	0.365	0.652	0.397	0.423	0.431	0.450	0.882	0.794	0.648	0.370	0.548	0.648	0.409
	154	0.635	0.348	0.603	0.577	0.569	0.550	0.118	0.206	0.352	0.630	0.452	0.352	0.591
Sco_105	176	0.017	0.130	0.207	0.115	0.267	0.645	0.417	0.389	0.519	0.093	0.113	0.000	0.014
	180	0.879	0.870	0.534	0.827	0.717	0.339	0.500	0.556	0.463	0.537	0.581	0.393	0.541
	184	0.103	0.000	0.259	0.038	0.017	0.016	0.083	0.056	0.019	0.370	0.306	0.607	0.432
	188	0.000	0.000	0.000	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014
Sco_106	186	0.000	0.000	0.000	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	190	0.159	0.109	0.121	0.600	0.096	0.031	0.158	0.235	0.130	0.241	0.133	0.019	0.015

	194	0.000	0.000	0.000	0.000	0.000	0.000	0.237	0.324	0.056	0.000	0.017	0.000	0.000
	198	0.000	0.000	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	206	0.091	0.065	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	210	0.000	0.000	0.000	0.000	0.135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	214	0.091	0.043	0.069	0.020	0.096	0.125	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	222	0.068	0.087	0.207	0.060	0.135	0.422	0.053	0.029	0.019	0.352	0.183	0.000	0.000
	226	0.318	0.283	0.466	0.280	0.212	0.250	0.079	0.000	0.074	0.074	0.150	0.481	0.515
	230	0.068	0.152	0.000	0.000	0.019	0.000	0.000	0.000	0.000	0.000	0.133	0.231	0.197
	234	0.182	0.261	0.138	0.000	0.135	0.016	0.211	0.059	0.000	0.167	0.283	0.269	0.258
	238	0.000	0.000	0.000	0.000	0.000	0.000	0.263	0.353	0.722	0.111	0.067	0.000	0.000
	242	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	246	0.000	0.000	0.000	0.020	0.096	0.156	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	250	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.056	0.033	0.000	0.015
Sco_215	288	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Sco_220	324	0.109	0.196	0.019	0.413	0.148	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	328	0.000	0.000	0.000	0.000	0.037	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	332	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	344	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.000	0.000	0.000
	348	0.000	0.087	0.056	0.043	0.241	0.048	0.500	0.500	0.712	0.185	0.242	0.000	0.015
	352	0.196	0.022	0.278	0.326	0.185	0.113	0.167	0.083	0.154	0.278	0.500	0.386	0.353
	356	0.000	0.043	0.148	0.109	0.222	0.645	0.000	0.000	0.000	0.000	0.081	0.000	0.000
	360	0.152	0.000	0.185	0.000	0.019	0.032	0.024	0.028	0.000	0.019	0.000	0.091	0.000
	364	0.065	0.174	0.148	0.000	0.037	0.065	0.190	0.333	0.077	0.111	0.081	0.045	0.132
	368	0.435	0.478	0.130	0.087	0.056	0.081	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	372	0.043	0.000	0.037	0.000	0.056	0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.000
	376	0.000	0.000	0.000	0.000	0.000	0.000	0.048	0.056	0.000	0.222	0.048	0.000	0.000
	380	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.432	0.500
	388	0.000	0.000	0.000	0.000	0.000	0.000	0.071	0.000	0.058	0.148	0.016	0.000	0.000
	392	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.000	0.000	0.000
Omm_1128	340	0.038	0.056	0.000	0.000	0.000	0.121	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	344	0.308	0.028	0.196	0.080	0.155	0.172	0.000	0.059	0.000	0.038	0.000	0.000	0.000
	348	0.192	0.306	0.107	0.200	0.362	0.241	0.354	0.412	0.130	0.250	0.259	0.196	0.014
	352	0.404	0.611	0.661	0.720	0.448	0.466	0.375	0.353	0.593	0.692	0.655	0.804	0.986
	356	0.058	0.000	0.036	0.000	0.034	0.000	0.271	0.176	0.278	0.019	0.086	0.000	0.000

Table 2-1: Allele frequencies of all loci in 13 streams-of-origin. Calculated in GENALEX 6.

Pop	Locus	Na	Ho	He	F
Cb	Sco_102	3.000	0.609	0.494	-0.231
	Sco_216	4.000	0.643	0.693	0.073
	Smm_22	15.000	0.917	0.910	-0.008
	Sfo_18	2.000	0.500	0.464	-0.078
	Sco_105	3.000	0.241	0.216	-0.118
	Sco_106	8.000	0.727	0.814	0.107
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	6.000	0.739	0.732	-0.010
	Omm_1128	5.000	0.692	0.700	0.012
Ga	Sco_102	3.000	0.524	0.516	-0.015
	Sco_216	4.000	0.727	0.586	-0.242
	Smm_22	19.000	0.957	0.906	-0.055
	Sfo_18	2.000	0.435	0.454	0.042
	Sco_105	2.000	0.261	0.227	-0.150
	Sco_106	7.000	0.826	0.803	-0.028
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	6.000	0.696	0.693	-0.004
	Omm_1128	4.000	0.500	0.529	0.055
Lo	Sco_102	2.000	0.759	0.499	-0.519
	Sco_216	3.000	0.654	0.521	-0.254
	Smm_22	18.000	0.931	0.899	-0.036
	Sfo_18	2.000	0.379	0.479	0.207
	Sco_105	3.000	0.690	0.605	-0.141
	Sco_106	5.000	0.724	0.702	-0.031
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	8.000	0.778	0.823	0.055
	Omm_1128	4.000	0.500	0.512	0.024
Wca	Sco_102	3.000	0.360	0.441	0.183
	Sco_216	4.000	0.192	0.409	0.530
	Smm_22	15.000	0.958	0.901	-0.064
	Sfo_18	2.000	0.385	0.488	0.212
	Sco_105	4.000	0.346	0.301	-0.150
	Sco_106	6.000	0.640	0.557	-0.149
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	6.000	0.783	0.701	-0.116
	Omm_1128	3.000	0.280	0.435	0.357
Sca	Sco_102	3.000	0.467	0.576	0.190
	Sco_216	4.000	0.552	0.628	0.121
	Smm_22	17.000	0.793	0.875	0.093
	Sfo_18	2.000	0.586	0.490	-0.195
	Sco_105	3.000	0.367	0.415	0.116
	Sco_106	10.000	0.846	0.870	0.027
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	9.000	0.815	0.827	0.015
	Omm_1128	4.000	0.793	0.643	-0.234
Mi	Sco_102	3.000	0.607	0.522	-0.162
	Sco_216	4.000	0.548	0.561	0.023
	Smm_22	15.000	0.759	0.894	0.152
	Sfo_18	2.000	0.633	0.495	-0.279
	Sco_105	3.000	0.387	0.469	0.174
	Sco_106	6.000	0.625	0.718	0.130
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	7.000	0.516	0.557	0.073
	Omm_1128	4.000	0.724	0.681	-0.064
Ra	Sco_102	2.000	0.280	0.385	0.272
	Sco_216	1.000	0.000	0.000	#N/A
	Smm_22	10.000	0.960	0.802	-0.198
	Sfo_18	2.000	0.235	0.208	-0.133
	Sco_105	3.000	0.458	0.569	0.195
	Sco_106	6.000	0.789	0.796	0.009
	Sco_215	1.000	0.000	0.000	#N/A

	Sco_220	6.000	0.619	0.678	0.087
	Omm_1128	3.000	0.667	0.661	-0.009
Sra	Sco_102	2.000	0.412	0.327	-0.259
	Sco_216	2.000	0.056	0.054	-0.029
	Smm_22	8.000	0.938	0.834	-0.124
	Sfo_18	2.000	0.412	0.327	-0.259
	Sco_105	3.000	0.611	0.537	-0.138
	Sco_106	5.000	0.824	0.711	-0.158
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	5.000	0.833	0.628	-0.327
	Omm_1128	4.000	0.706	0.671	-0.052
Nra	Sco_102	2.000	0.400	0.480	0.167
	Sco_216	1.000	0.000	0.000	#N/A
	Smm_22	9.000	0.640	0.776	0.175
	Sfo_18	2.000	0.481	0.456	-0.056
	Sco_105	3.000	0.778	0.516	-0.506
	Sco_106	5.000	0.333	0.453	0.264
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	4.000	0.577	0.461	-0.252
	Omm_1128	3.000	0.704	0.555	-0.268
Du	Sco_102	2.000	0.519	0.384	-0.350
	Sco_216	2.000	0.407	0.324	-0.256
	Smm_22	13.000	0.926	0.880	-0.052
	Sfo_18	2.000	0.148	0.466	0.682
	Sco_105	3.000	0.481	0.566	0.149
	Sco_106	6.000	0.778	0.770	-0.011
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	8.000	0.815	0.804	-0.014
	Omm_1128	4.000	0.615	0.456	-0.348
Hi	Sco_102	2.000	0.333	0.278	-0.200
	Sco_216	3.000	0.226	0.206	-0.099
	Smm_22	9.000	0.867	0.821	-0.056
	Sfo_18	2.000	0.387	0.495	0.218
	Sco_105	3.000	0.581	0.556	-0.044
	Sco_106	8.000	0.900	0.822	-0.095
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	7.000	0.581	0.675	0.140
	Omm_1128	3.000	0.414	0.496	0.166
Lli	Sco_102	2.000	0.250	0.423	0.409
	Sco_216	2.000	0.429	0.436	0.018
	Smm_22	7.000	0.586	0.655	0.104
	Sfo_18	2.000	0.630	0.456	-0.380
	Sco_105	2.000	0.429	0.477	0.102
	Sco_106	4.000	0.846	0.643	-0.316
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	5.000	0.455	0.652	0.303
	Omm_1128	2.000	0.179	0.316	0.434
Uli	Sco_102	2.000	0.333	0.444	0.250
	Sco_216	3.000	0.343	0.502	0.317
	Smm_22	6.000	0.444	0.466	0.047
	Sfo_18	2.000	0.515	0.483	-0.066
	Sco_105	4.000	0.514	0.520	0.013
	Sco_106	5.000	0.697	0.629	-0.108
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	4.000	0.618	0.608	-0.016
	Omm_1128	2.000	0.029	0.028	-0.014

Table 2-2: Basic statistical data on number of alleles (*Na*), observed (*Ho*) and expected (*He*) heterozygosity and fixation index (*F*) for each locus in each stream-of-origin. Calculated in Genalex 6.

Population	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli	Global
N	29	23	29	27	30	32	25	18	27	27	31	29	37	364
pw *Fst*, Cb		*	*	*	*	*	*	*	*	*	*	*	*	N/A
pw *Fst*, Ga	0.0505		*	*	*	*	*	*	*	*	*	*	*	N/A
pw *Fst*, LO	0.0497	0.0742		*	*	*	*	*	*	*	*	*	*	N/A
pw *Fst*, Wca	0.114	0.0787	0.1051		*	*	*	*	*	*	*	*	*	N/A
pw *Fst*, Sca	0.0613	0.0484	0.0452	0.0672		*	*	*	*	*	*	*	*	N/A
pw *Fst*, Mi	0.1455	0.1327	0.0688	0.1686	0.0622		*	*	*	*	*	*	*	N/A
pw *Fst,* Ra	0.1962	0.1873	0.1627	0.2878	0.1646	0.1907		NS	NS	*	*	*	*	N/A
pw *Fst*, Sra	0.1869	0.1891	0.1645	0.2827	0.1612	0.193	-0.008		*	*	*	*	*	N/A
pw *Fst*, Nra	0.2398	0.2359	0.1883	0.3057	0.1907	0.2109	0.0559	0.0601		*	*	*	*	N/A
pw *Fst*, Du	0.0998	0.1357	0.0561	0.1729	0.1003	0.1349	0.1159	0.1146	0.1502		*	*	*	N/A
pw *Fst*, Hi	0.1232	0.1576	0.0842	0.2181	0.1304	0.1612	0.0909	0.1018	0.1527	0.0344		*	*	N/A
pw *Fst*, Lli	0.1674	0.1698	0.1145	0.2053	0.1685	0.2284	0.2625	0.275	0.3174	0.159	0.1547		NS	N/A
pw *Fst*, Uli	0.1775	0.1992	0.1142	0.232	0.1898	0.2396	0.2863	0.2968	0.3129	0.1577	0.1483	0.0419		N/A
Mean allelic richness, FSTAT	4.916	4.862	4.558	4.420	5.117	4.394	3.616	3.522	3.076	4.150	3.931	2.818	2.555	5.558
Total private alleles, GENALEX	1	1	1	2	3		1			3			1	12
He, GENALEX	0.563	0.524	0.560	0.470	0.593	0.547	0.455	0.454	0.411	0.517	0.483	0.456	0.409	0.4955
Fst, FSTAT, W&C														0.158
Fis, FSTAT W&C	0.012	-0.021	-0.056	0.089	0.038	0.037	0.045	-0.142	-0.039	0.011	0.031	0.082	0.066	
#loci	9	9	9	9	9	9	9	9	9	9	9	9	9	9
Out of HWE? GENEPOP	No	No	No	1 locus	No	No	No	No	No	No	No	1 locus	No	Yes
Het def, GENEPOP	No	No	No	locus	No	No	No	No	No	No	No	locus	No	N/A
Het excess, GENEPOP	No	No	No	No	No	No	No	No	No	No	No	No	No	N/A
Null Alleles? MICROCHECKER	No	No	No	2 loci	No	1 locus	No	No	No	1 locus	No	1 locus	No	N/A
LD? FSTAT	No	No	No	No	No	No	No	No	No	No	No	No	1 pair	N/A
Bottleneck? TPM, alpha <0.05	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	N/A
Bottleneck? TPM, alpha <0.01	No	No	No	No	Yes	No	Yes	Yes	Yes	Yes	No	No	No	N/A

Table 2-3: Summary of pertinent tests for stream-of-origin defined populations. Pairwise-*Fst* was calculated in FSTAT 2.9.3 using 1560 random permutations. Significance values for *Fst,* HWE, heterzygote deficiency and heterozygote excess were obtained using the sequential Bonferroni procedure.

Cb

locus	HWE	Het def	Het exc
Sco_102	0.4266	0.9128	0.1844
Sco_216	0.5179	0.2594	0.7844
Smm_22	0.3846	0.4076	0.6603
Sfo_18	1	0.7651	0.5557
Sco_105	1	1	0.6548
Sco_106	0.2798	0.1192	0.8995
Sco_215	-	-	-
Sco_220	0.5615	0.4901	0.53
Omm_1128	0.2209	0.4923	0.5181

All
chi2   11.6526
Df   16
Prob: 0.7675

Ga

locus	HWE	Het def	Het exc
Sco_102	0.6798	0.5325	0.6117
Sco_216	0.4919	0.9471	0.0881
Smm_22	0.9578	0.8421	0.2548
Sfo_18	1	0.5546	0.7833
Sco_105	1	1	0.6887
Sco_106	0.0312	0.5427	0.4928
Sco_215	-	-	-
Sco_220	0.0871	0.2073	0.8068
Omm_1128	0.0326	0.0203	0.9844

All
chi2   20.9385
Df   16
Prob: 0.1809

Lo

locus	HWE	Het def	Het exc
Sco_102	0.0101	0.9993	0.0084
Sco_216	0.3061	0.9201	0.1127
Smm_22	0.174	0.5802	0.4805
Sfo_18	0.2641	0.2032	0.9477
Sco_105	0.8665	0.8016	0.2378
Sco_106	0.4217	0.4371	0.5796
Sco_215	-	-	-
Sco_220	0.5243	0.2556	0.7649
Omm_1128	0.3756	0.38	0.6764

All
chi2   22.9851
Df   16
Prob: 0.1141

Wca

locus	HWE	Het def	Het exc
Sco_102	0.0334	0.0121	0.996
Sco_216	0.0001	0.0158	0.9843
Smm_22	0.0479	0.2414	0.7428
Sfo_18	0.416	0.2144	0.9448
Sco_105	1	1	0.4407
Sco_106	0.048	0.9279	0.0784
Sco_215	-	-	-
Sco_220	0.73	0.7474	0.2633
Omm_1128	0.0281	0.0594	0.9483

All
chi2   47.4441
Df   16
Prob: 0.0001

Sca

locus	HWE	Het def	Het exc
Sco_102	0.2295	0.0459	0.9629

Sra

locus	HWE	Het def	Het exc
Sco_102	1	1	0.4632
Sco_216	-	-	-
Smm_22	0.9488	0.9202	0.1776
Sfo_18	1	1	0.4637
Sco_105	1	0.7855	0.394
Sco_106	0.8968	0.8746	0.1737
Sco_215	-	-	-
Sco_220	0.4145	0.9641	0.0498
Omm_1128	0.0547	0.5864	0.4703

All
chi2   7.8959
Df   14
Prob: 0.8947

Nra

locus	HWE	Het def	Het exc
Sco_102	0.4211	0.2989	0.9139
Sco_216	-	-	-
Smm_22	0.0526	0.0729	0.9223
Sfo_18	1	0.7271	0.5976
Sco_105	0.0098	0.998	0.0054
Sco_106	0.0729	0.0809	0.9266
Sco_215	-	-	-
Sco_220	0.7353	1	0.0554
Omm_1128	0.07	0.8046	0.2009

All
chi2   28.0423
Df   14
Prob: 0.014

Du

locus	HWE	Het def	Het exc
Sco_102	0.1394	1	0.1019
Sco_216	0.5497	1	0.2804
Smm_22	0.7663	0.7694	0.3116
Sfo_18	0.0005	0.0004	1
Sco_105	0.0874	0.2179	0.7914
Sco_106	0.1607	0.3042	0.7003
Sco_215	-	-	-
Sco_220	0.7229	0.4918	0.5141
Omm_1128	0.296	1	0.034

All
chi2   32.3181
Df   16
Prob: 0.0091

Hi

locus	HWE	Het def	Het exc
Sco_102	0.5633	1	0.4075
Sco_216	1	1	0.6851
Smm_22	0.1822	0.4024	0.6017
Sfo_18	0.2781	0.1702	0.9543
Sco_105	1	0.6603	0.4132
Sco_106	0.6436	0.8802	0.1532
Sco_215	-	-	-
Sco_220	0.0281	0.0022	0.9968
Omm_1128	0.1238	0.1693	0.8563

All
chi2   19.318
Df   16
Prob: 0.2525

Lli

locus	HWE	Het def	Het exc
Sco_102	0.0654	0.0359	0.9966

locus	HWE	Het def	Het exc
Sco_216	0.0381	0.1377	0.8671
Smm_22	0.2551	0.1165	0.8806
Sfo_18	0.4535	0.9111	0.2821
Sco_105	0.7369	0.3159	0.8653
Sco_106	0.0979	0.0428	0.9555
Sco_215	-	-	-
Sco_220	0.0925	0.0341	0.9614
Omm_1128	0.2701	0.8848	0.1156

All
chi2   26.4307
Df   16
Prob: 0.0483

locus	HWE	Het def	Het exc
Sco_216	1	0.5892	0.7384
Smm_22	0.1343	0.0146	0.9861
Sfo_18	0.0942	0.9933	0.0671
Sco_105	0.6937	0.4008	0.8493
Sco_106	0.1448	0.9775	0.023
Sco_215	-	-	-
Sco_220	0.0029	0.0396	0.9632
Omm_1128	0.0394	0.0398	0.9978

All
chi2   36.9427
Df   16
Prob: 0.0021

Mi

locus	HWE	Het def	Het exc
Sco_102	0.3885	0.8107	0.2501
Sco_216	0.4748	0.4924	0.5328
Smm_22	0.0083	0.0015	0.9975
Sfo_18	0.2631	0.9679	0.1382
Sco_105	0.5077	0.1943	0.9123
Sco_106	0.0965	0.065	0.9382
Sco_215	-	-	-
Sco_220	0.0752	0.2376	0.7759
Omm_1128	0.522	0.5172	0.5164

All
chi2   28.1453
Df   16
Prob: 0.0304

Uli

locus	HWE	Het def	Het exc
Sco_102	0.1403	0.1109	0.9752
Sco_216	0.0563	0.0323	0.9867
Smm_22	0.1159	0.3533	0.7485
Sfo_18	1	0.7435	0.5281
Sco_105	0.8471	0.4938	0.6412
Sco_106	0.1339	0.8129	0.2268
Sco_215	-	-	-
Sco_220	0.6902	0.3859	0.6446
Omm_1128	-	-	-

All
chi2   19.0886
Df   14
Prob: 0.1616

Ra

locus	HWE	Het def	Het exc
Sco_102	0.2865	0.1666	0.9759
Sco_216	-	-	-
Smm_22	0.0436	0.9961	0.0207
Sfo_18	1	1	0.8222
Sco_105	0.3982	0.1919	0.8385
Sco_106	0.008	0.4754	0.526
Sco_215	-	-	-
Sco_220	0.4713	0.1391	0.8648
Omm_1128	0.8066	0.5307	0.5052

All
chi2   22.2053
Df   14
Prob: 0.0745

Table 2-4: *P*-value results from Hardy-Weinberg exact tests across all loci and 13 stream-of-origin populations as calculated in GENEPOP 4. Highlighted values are significant after sequential Bonferroni correction.

Locus pair	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli	All
Sco102 X Sco216	0.0203	0.84754	0.26987	0.46987	0.79103	0.23504	NA	1	NA	0.11763	0.67415	0.39103	0.55246	0.25524
Sco102 X Smm22	0.21816	0.26111	0.04145	0.98536	0.15694	0.9015	0.11282	0.46506	0.81485	0.3328	0.62019	0.54541	0.70962	0.20118
Sco102 X Sfo18	0.54455	0.68868	0.70545	0.58045	0.72981	0.23472	0.28301	0.20235	0.87436	0.57885	0.57329	0.17788	0.59679	0.72083
Sco102 X Sco105	0.96987	0.10524	0.12853	0.66688	0.73045	0.99936	0.32981	0.28921	0.43355	0.74925	0.49637	0.87831	0.52179	0.86806
Sco102 X Sco106	0.11207	0.49124	0.13248	0.625	0.92917	0.23729	0.64338	0.06934	0.68002	0.97457	0.36603	0.66613	0.18889	0.29252
Sco102 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco102 X Sco220	0.3562	0.19295	0.87329	0.4297	0.62254	0.0187	0.29038	0.38825	0.11026	0.14145	0.07041	0.80491	0.19936	0.00962
Sco102 X Omm1128	0.38803	0.49915	0.911	0.54434	0.22019	0.52917	0.12425	0.89968	0.10951	0.63846	0.52073	0.3969	1	0.42073
Sco216 X Smm22	1	1	0.3203	0.49006	0.78665	1	NA	0.64476	NA	0.11635	0.01656	0.96763	0.91624	0.58365
Sco216 X Sfo18	0.49081	0.35321	0.26763	0.85342	0.57596	0.6859	NA	0.44498	NA	0.86848	0.87746	0.61197	0.94765	0.9125
Sco216 X Sco105	0.25705	0.24071	0.02115	0.04476	0.15043	0.67682	NA	0.11549	NA	0.13034	0.21368	0.66966	0.58045	0.00951
Sco216 X Sco106	0.04487	0.35064	0.28077	0.00395	1	0.09252	NA	0.28013	NA	0.8984	0.17468	0.45182	0.40427	0.01517
Sco216 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco216 X Sco220	1	0.83376	0.74573	0.01004	0.45395	0.41549	NA	0.22126	NA	0.46934	0.0735	0.48419	0.64434	0.19092
Sco216 X Omm1128	0.01912	0.48536	0.5297	0.36175	0.98237	0.89241	NA	0.40769	NA	0.32874	0.63013	0.68056	0.80192	0.70726
Smm22 X Sfo18	1	1	0.14551	0.75032	0.8281	0.86966	0.7563	0.8172	0.33515	0.0828	0.39605	0.66517	0.05791	0.31111
Smm22 X Sco105	0.95118	0.60844	0.08387	0.00556	0.83141	0.7078	0.31036	0.68184	0.81368	0.00759	0.19701	0.2422	0.18045	0.01154
Smm22 X Sco106	1	1	0.61763	0.01421	1	1	0.90118	1	0.01795	0.08878	0.03237	0.51912	0.22831	0.01699
Smm22 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Smm22 X Sco220	0.15673	1	0.22906	0.01325	1	0.15919	0.72372	1	0.00021	0.03291	0.35449	0.12105	0.63269	0.00085
Smm22 X Omm1128	1	0.21538	0.5859	0.86453	0.10694	1	0.92468	0.09487	0.28739	0.05043	0.36068	0.00823	0.10021	0.00662
Sfo18 X Sco105	0.07895	1	0.07874	0.71966	0.34765	0.08793	0.82596	0.8235	0.51944	0.01464	0.14733	0.41752	0.07393	0.01496
Sfo18 X Sco106	0.94327	0.264	0.69231	0.01389	0.48665	0.08141	0.0906	0.12404	0.10096	0.00524	0.53216	0.68216	0.06902	0.00278
Sfo18 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sfo18 X Sco220	0.44177	0.47564	0.23226	0.89658	0.07051	0.1484	0.01004	0.39498	0.42425	0.00064	0.39669	0.54626	0.00267	0.00021
Sfo18 X Omm1128	0.83547	0.98184	0.83835	0.76902	0.94551	0.86806	0.19423	0.11474	0.69316	0.06592	0.4328	0.33707	0.35897	0.8313
Sco105 X Sco106	0.0906	0.53024	0.06058	0.18568	0.50053	0.639	0.31774	0.16442	0.01688	0.06165	0.95417	0.49947	0.23109	0.00556
Sco105 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco105 X Sco220	0.13632	0.0047	0.0797	0.00032	0.07212	0.00278	0.45235	0.00118	0.04124	0.00075	0.17575	0.00064	0.00011	0.00011
Sco105 X Omm1128	0.06368	0.87511	0.92105	0.1078	0.48269	0.74263	0.69071	0.07682	0.29882	0.08045	0.87596	0.65406	0.12115	0.26218
Sco106 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco106 X Sco220	0.439	0.27863	0.8031	0.01923	1	0.04893	0.90748	0.0703	0.1735	0.01442	0.97746	0.12254	0.29274	0.00235
Sco106 X Omm1128	1	0.70919	0.00609	0.5312	0.95374	0.92126	0.42874	0.80267	0.00812	0.06944	0.53515	0.04765	0.24263	0.02329
Sco215 X Sco220	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco215 X Omm1128	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sco220 X Omm1128	0.31293	0.9187	0.60427	0.03013	0.98397	0.7344	0.40417	0.12094	0.15919	0.3562	0.54615	0.12639	0.2265	0.1594

Table 2-5: *P*-value results from genotypic disequilibrium tests across all loci and 13 stream-of-origin populations as calculated in FSTAT 2.9.3 using 9630 random permutations. Highlighted values are significant after sequential Bonferroni correction.
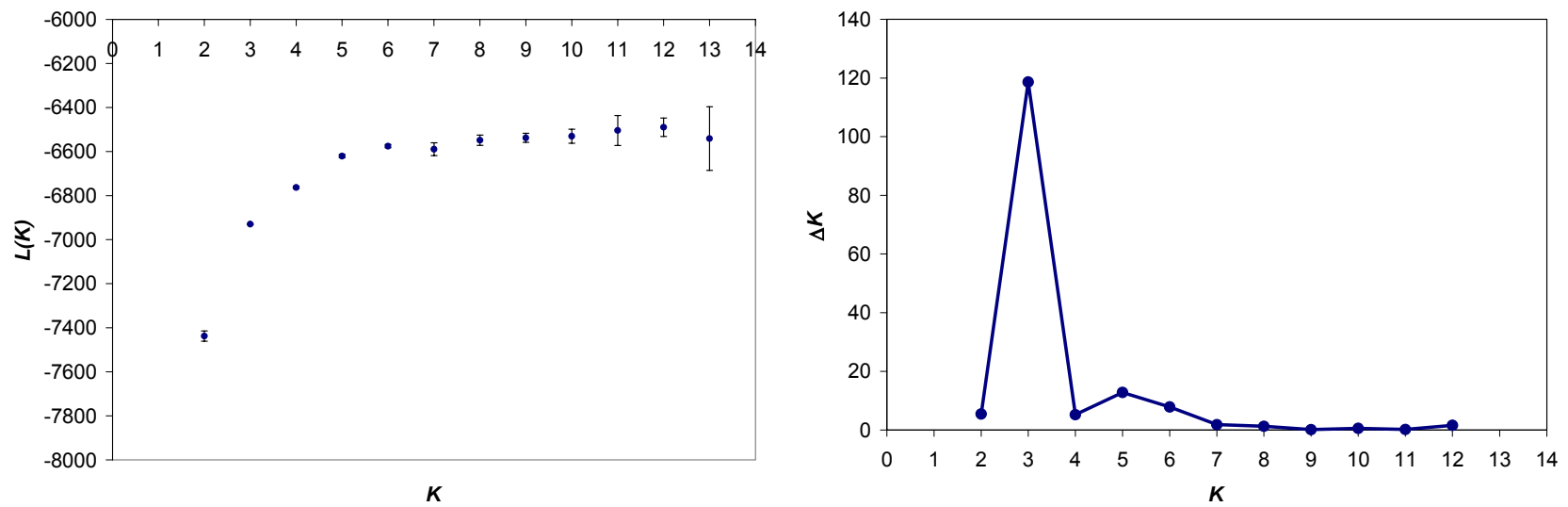
Figure 2-4: Determining the correct *K* value for the first round of hierarchical STRUCTURE analysis. Left: Values represent the mean values of $\ln \Pr(X|K)$ (*L(K)*) over 10 runs in STRUCTURE 2.2 +/- the s.d. Right: True value of *K* is 3, as discerned from $\Delta K = m($I *L''K* I$) / s[L(K)]$ for sample. Overall length = 200,000 in STRUCTURE 2.2 (admixture model, allele frequencies correlated).
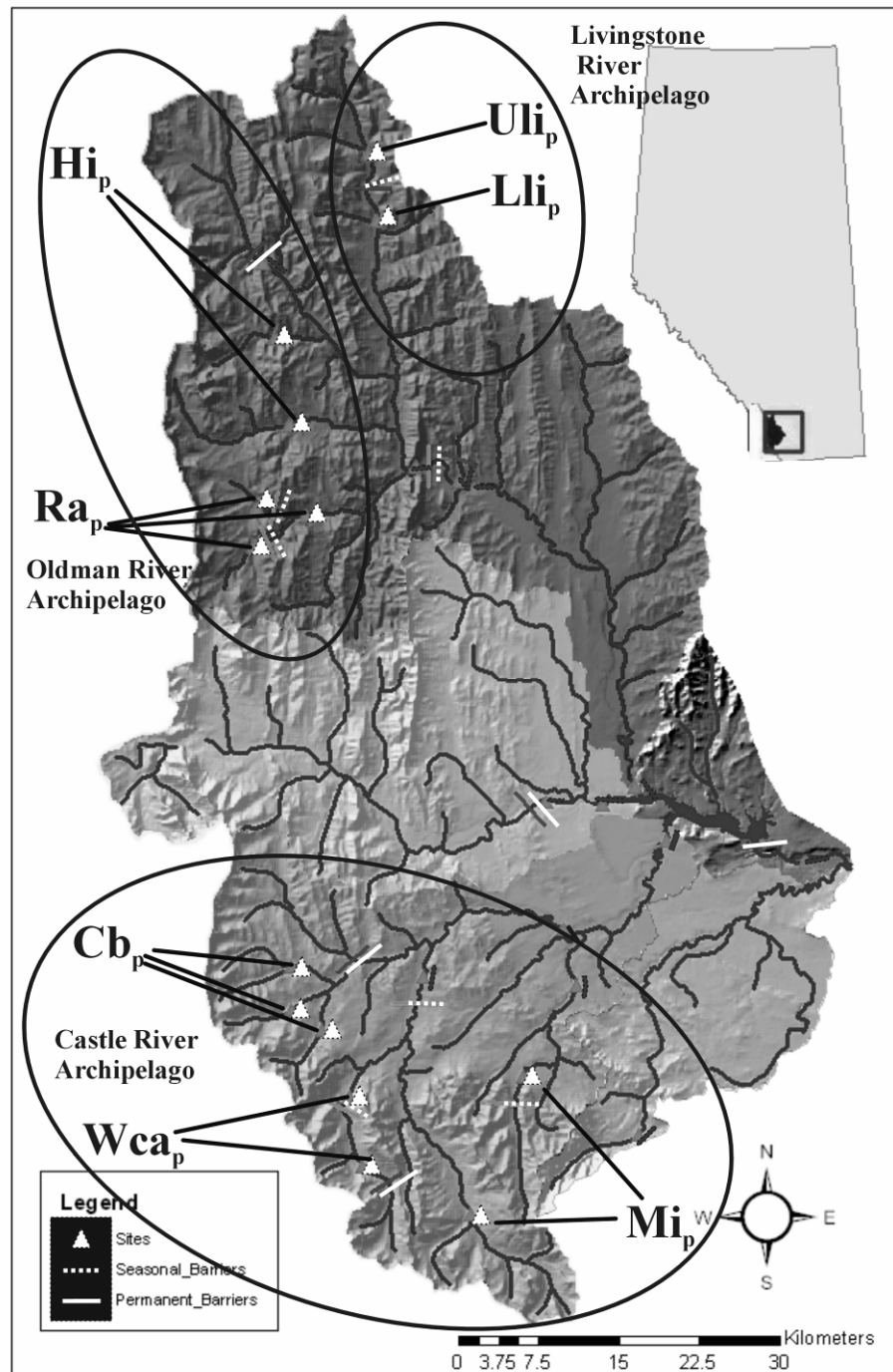
Figure 2-5: Archipelagos and populations found by hierarchical STRUCTURE analysis.

Li archpelago

K	L(K)	s.d.	L'(K)	I L"K I	ΔK
1	-907.48	0.220101			
2	-875.8	1.43217	31.68	83.44	58.26124
3	-927.56	51.89995	-51.76	39.58	0.762621
4	-939.74	11.38031	-12.18	19.96	1.753906
5	-931.96	12.8577	7.78		

OMR archipelago

K	L(K)	s.d.	L'(K)	I L"K I	ΔK
1	-1948.94	0.371782			
2	-1839.95	0.460072	108.99	118.79	258.1985
3	-1849.75	10.16095	-9.8	11.1	1.092417
4	-1870.65	8.164592	-20.9	54.39	6.661693
5	-1945.94	19.00685	-75.29	19.84	1.043834
6	-2001.39	19.18069	-55.45	6.1	0.318028
7	-2062.94	24.94858	-61.55	30.99	1.242155
8	-2093.5		-30.56		

Ca archipelago

K	L(K)	s.d.	L'(K)	I L"K I	ΔK
1	-3835.74	0.254733			
2	-3696.5	2.002776	139.24	97.22	48.54263
3	-3654.48	7.339967	42.02	46.39	6.320192
4	-3658.85	7.508551	-4.37	16.53	2.20149
5	-3679.75	26.95149	-20.9	79.93	2.965699
6	-3620.72	94.35206	59.03	58.87	0.62394
7	-3620.56	42.39217	0.16	22.01	0.5192
8	-3642.41	13.35777	-21.85	0.34	0.025453
9	-3664.6	8.066908	-22.19		

West Ca,
3rd round

K	L(K)	s.d.	L'(K)	I L"K I	ΔK
1	-1631.84	0.374759			
2	-1621.72	2.785578	10.11778	179.2756	64.35848
3	-1790.88	65.09762	-169.158	299.0578	4.59399
4	-1660.98	12.50802	129.9	190.87	15.2598
5	-1721.95	33.81789	-60.97		

Table 2-6: Determining the correct K value for the second and third rounds of hierarchical STRUCTURE analysis. L(K) Values represent the mean values of $\ln \Pr(X|K)$ over 10 runs in STRUCTURE 2.2 +/- the s.d. True value of K is discerned by the highest value of ΔK for each analysis (data table). Overall length = 600,000 in Li archipelago and West Ca group; 1,000,000 in OMR and Ca archipelagos. Values obtained in STRUCTURE 2.2 (admixture model, allele frequencies correlated).
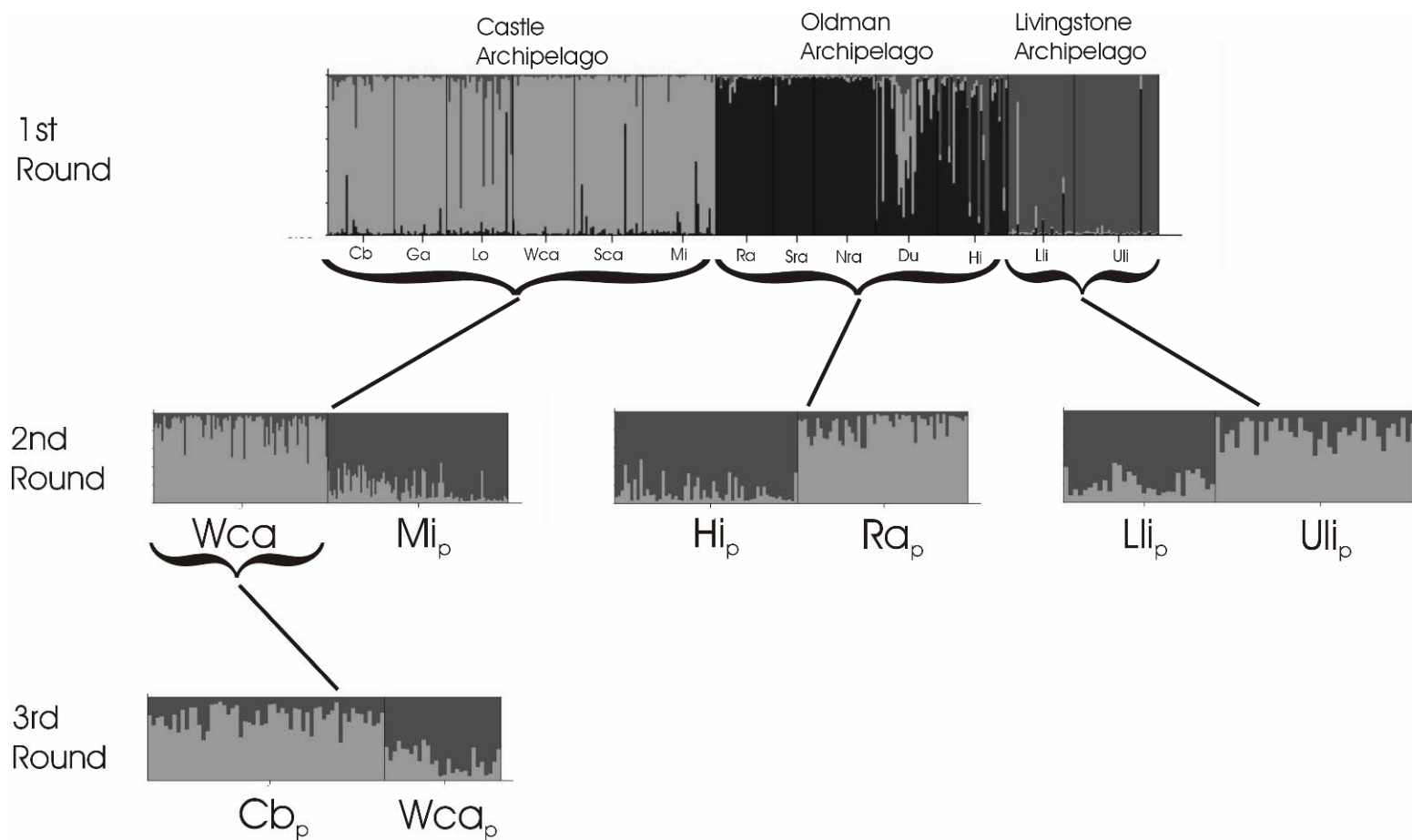
Figure 2-6: Flow chart of hierarchical STRUCTURE analysis. Each chart represents the output from an analysis, with thin vertical columns each representing individual fish on the x-axis. On the y-axis is the likelihood of assignment to any given cluster K, which are represented by different shades in the chart. In the first (topmost) chart, sampling locations are separated by black lines, while in subsequent charts, populations found by previous rounds of analysis are separated.

Locus	Allele	Ca	OMR	Li	Locus	Allele	Ca	OMR	Li
Sco_102	168	0.403	0.750	0.674	**Sco_105**	176	0.227	0.344	0.007
	172	0.541	0.250	0.326		180	0.680	0.536	0.465
	176	0.056	0.000	0.000		184	0.090	0.121	0.521
Sco_216	222	0.066	0.000	0.000		188	0.003	0.000	0.007
	226	0.446	0.040	0.500	**Sco_106**	186	0.003	0.000	0.000
	230	0.413	0.951	0.486		190	0.186	0.186	0.000
	234	0.066	0.009	0.014		194	0.000	0.114	0.000
	238	0.009	0.000	0.000		198	0.006	0.000	0.000
Smm_22	213	0.000	0.005	0.000		206	0.028	0.000	0.000
	217	0.003	0.106	0.000		210	0.022	0.000	0.000
	221	0.000	0.000	0.007		214	0.075	0.000	0.000
	225	0.003	0.000	0.014		222	0.189	0.110	0.023
	229	0.000	0.009	0.600		226	0.286	0.057	0.523
	233	0.068	0.000	0.221		230	0.034	0.019	0.208
	237	0.065	0.000	0.029		234	0.115	0.152	0.246
	241	0.071	0.000	0.000		238	0.003	0.333	0.000
	245	0.000	0.041	0.007		242	0.003	0.000	0.000
	249	0.034	0.069	0.086		246	0.050	0.000	0.000
	253	0.037	0.005	0.000		250	0.000	0.029	0.000
	257	0.105	0.239	0.014	**Sco_215**	288	1.000	1.000	1.000
	261	0.068	0.028	0.000	**Sco_220**	324	0.137	0.000	0.000
	265	0.093	0.032	0.007		328	0.006	0.000	0.000
	269	0.065	0.060	0.014		332	0.003	0.000	0.000
	273	0.015	0.252	0.000		344	0.003	0.000	0.000
	277	0.031	0.096	0.000		348	0.080	0.458	0.000
	281	0.046	0.005	0.000		352	0.172	0.250	0.397
	285	0.022	0.000	0.000		356	0.213	0.023	0.000
	289	0.065	0.000	0.000		360	0.057	0.014	0.048
	293	0.037	0.000	0.000		364	0.089	0.139	0.103
	297	0.015	0.000	0.000		368	0.191	0.000	0.000
	301	0.009	0.000	0.000		372	0.022	0.000	0.016
	305	0.006	0.000	0.000		376	0.013	0.060	0.000
	309	0.003	0.000	0.000		380	0.000	0.000	0.437
	313	0.025	0.000	0.000		388	0.013	0.051	0.000
	317	0.025	0.000	0.000		392	0.000	0.005	0.000
	321	0.043	0.000	0.000	**Omm_1128**	340	0.034	0.000	0.000
	325	0.022	0.046	0.000		344	0.159	0.018	0.000
	329	0.022	0.005	0.000		348	0.241	0.289	0.067
	333	0.003	0.005	0.000		352	0.544	0.509	0.933
Sfo_18	148	0.440	0.668	0.500		356	0.022	0.183	0.000
	154	0.560	0.332	0.500					

Table 2-7: Allele frequencies of all loci in the 3 archipelagos. Calculated in GENALEX 6.

Pop	Locus	Na	Ho	He	F
Ca	**Sco_102**	3.000	0.563	0.542	-0.038
	Sco_216	5.000	0.554	0.622	0.109
	Smm_22	27.000	0.883	0.941	0.062
	Sfo_18	2.000	0.473	0.493	0.040
	Sco_105	4.000	0.372	0.478	0.221
	Sco_106	13.000	0.739	0.824	0.103
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	13.000	0.732	0.851	0.139
	Omm_1128	5.000	0.581	0.619	0.062
OMR	**Sco_102**	2.000	0.355	0.375	0.055
	Sco_216	3.000	0.097	0.093	-0.043
	Smm_22	16.000	0.853	0.845	-0.010
	Sfo_18	2.000	0.375	0.443	0.154
	Sco_105	3.000	0.607	0.580	-0.046
	Sco_106	8.000	0.695	0.802	0.133
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	8.000	0.685	0.701	0.023
	Omm_1128	4.000	0.679	0.623	-0.089
Li	**Sco_102**	2.000	0.304	0.440	0.308
	Sco_216	3.000	0.391	0.514	0.239
	Smm_22	10.000	0.529	0.582	0.092
	Sfo_18	2.000	0.545	0.500	-0.091
	Sco_105	4.000	0.465	0.512	0.093
	Sco_106	4.000	0.738	0.622	-0.187
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	5.000	0.476	0.639	0.255
	Omm_1128	2.000	0.075	0.125	0.404

Table 2-8: Basic statistical data on number of alleles (*Na*), observed (*Ho*) and expected (*He*) heterozygosity and fixation index (*F*) for each locus in each archipelago. Calculated in GENALEX 6.

Population	Ca	Omr	Li		Global
n	174	113	72		359
pw *Fst*, Ca		*	*		N/A
pw *Fst*, OMR	0.1212		*		N/A
pw *Fst*, Li	0.1335	0.2177			N/A
Mean allelic richness, FSTAT	7.185	4.884	3.604		7.433
Total private alleles, GENALEX	28	4	2		
He, GENALEX	0.597	0.496	0.437		0.510
Fst, FSTAT, W&C					0.146
Fis, FSTAT W&C	0.091	0.031	0.112		
#loci	9	9	9		9
Out of HWE? GENEPOP	Yes, 4 loci	Yes, 1 locus	Yes, 1 locus		Yes
Het def, GENEPOP	Yes, 4 loci	No	Yes, 1 locus		N/A
Het excess, GENEPOP	No	No	No		N/A
Null Alleles? MICROCHECKER	Yes, 4 loci	Yes, 1 locus	Yes, 4 loci		N/A
LD? FSTAT	Yes, 2 pairs	Yes, 2 pairs	Yes, 1 pair		N/A
Bottleneck? TPM, alpha <0.05	No	Yes	No		N/A
Bottleneck? TPM, alpha <0.01	No	Yes	No		N/A

Table 2-9: Summary of pertinent tests for archipelagos. Pairwise-*Fst* was calculated in FSTAT 2.9.3 using 60 random permutations. Significance values for *Fst*, HWE, heterzygote deficiency and heterozygote excess were obtained using the sequential Bonferroni procedure.

Castle archipelago

locus	HWE	Het def	Het exc
Sco_102	0.008	0.0416	0.9594
Sco_216	0.1773	0.0154	0.985
Smm_22	0.0002	0.0004	1
Sfo_18	0.6419	0.3379	0.7574
Sco_105	0.0011	0.0002	0.9998
Sco_106	0.0001	0	1
Sco_215	-	-	-
Sco_220	0	0	1
Omm_1128	0.0174	0.0081	0.9915

All
chi2   91.7827
Df   16
Prob: 0

Livingstone archipelago

locus	HWE	Het def	Het exc
Sco_102	0.0136	0.0097	0.9978
Sco_216	0.0336	0.0307	0.9777
Smm_22	0.0544	0.003	0.9942
Sfo_18	0.6223	0.8231	0.3321
Sco_105	0.434	0.226	0.8471
Sco_106	0.2436	0.9815	0.0217
Sco_215	-	-	-
Sco_220	0	0.002	0.9981
Omm_1128	0.0204	0.0208	0.9995

All
chi2   59.2912
Df   16
Prob: 0

Oldman archipelago

locus	HWE	Het def	Het exc
Sco_102	0.612	0.3476	0.8146
Sco_216	1	1	0.7718
Smm_22	0.0302	0.5604	0.4586
Sfo_18	0.1245	0.0826	0.9679
Sco_105	0.0715	0.3833	0.6146
Sco_106	0.0001	0.1181	0.8651
Sco_215	-	-	-
Sco_220	0.4013	0.2478	0.7425
Omm_1128	0.2746	0.928	0.0778

All
chi2   41.4158
Df   16
Prob: 0.0005

Table 2-10: *P*-value results from Hardy-Weinberg exact tests across all loci and archipelagos as calculated in GENEPOP 4. Highlighted values are significant after sequential Bonferroni correction.

	Ca	OMR	Li	All
Sco102 X Sco216	0.00278	0.56898	0.33519	0.00648
Sco102 X Smm22	0.14028	0.09167	0.5838	0.07593
Sco102 X Sfo18	0.88796	0.82407	0.16343	0.78796
Sco102 X Sco105	0.83843	0.68935	0.33889	0.83241
Sco102 X Sco106	0.21111	0.04861	0.03287	0.01713
Sco102 X Sco215	NA	NA	NA	NA
Sco102 X Sco220	0.2787	0.13796	0.60648	0.18704
Sco102 X Omm1128	0.08704	0.15602	0.27361	0.03935
Sco216 X Smm22	0.85278	0.06204	0.86806	0.66204
Sco216 X Sfo18	0.01852	0.91481	0.40046	0.11898
Sco216 X Sco105	0.01759	0.00185	0.57593	0.00231
Sco216 X Sco106	`0.00046`	0.20278	0.08472	`0.00046`
Sco216 X Sco215	NA	NA	NA	NA
Sco216 X Sco220	0.01759	0.02176	0.34722	0.00463
Sco216 X Omm1128	0.45787	0.20231	0.50833	0.36481
Smm22 X Sfo18	0.96574	0.71157	0.57593	0.94769
Smm22 X Sco105	0.18796	0.05278	0.37917	0.03981
Smm22 X Sco106	0.20509	0.00139	0.32454	0.00787
Smm22 X Sco215	NA	NA	NA	NA
Smm22 X Sco220	0.00139	0.05139	0.4125	0.0037
Smm22 X Omm1128	0.85556	0.00694	0.00093	0.00185
Sfo18 X Sco105	0.0713	0.10648	0.13472	0.01806
Sfo18 X Sco106	0.11111	0.00139	0.7963	0.00787
Sfo18 X Sco215	NA	NA	NA	NA
Sfo18 X Sco220	0.26944	0.00324	0.28472	0.01343
Sfo18 X Omm1128	0.63472	0.26574	0.49861	0.49491
Sco105 X Sco106	0.26204	`0.00046`	0.70139	0.00417
Sco105 X Sco215	NA	NA	NA	NA
Sco105 X Sco220	`0.00046`	`0.00046`	`0.00046`	`0.00046`
Sco105 X Omm1128	0.61111	0.1412	0.48565	0.3912
Sco106 X Sco215	NA	NA	NA	NA
Sco106 X Sco220	`0.00046`	0.00972	0.58519	`0.00046`
Sco106 X Omm1128	0.52315	0.01806	0.07315	0.04167
Sco215 X Sco220	NA	NA	NA	NA
Sco215 X Omm1128	NA	NA	NA	NA
Sco220 X Omm1128	0.94907	0.24398	0.28009	0.71898

Table 2-11: *P*-value results from genotypic disequilibrium tests across all loci and archipelagos as calculated in FSTAT 2.9.3 using 2160 random permutations. Highlighted values are significant after sequential Bonferroni correction.

Locus	Allele	Cb	Wca	Mi	Ra	Du	Lli	Uli
Sco_102	**168**	0.575	0.217	0.345	0.709	0.796	0.724	0.638
	172	0.406	0.783	0.560	0.291	0.204	0.276	0.363
	176	0.019	0.000	0.095	0.000	0.000	0.000	0.000
Sco_216	**222**	0.170	0.058	0.006	0.000	0.000	0.000	0.000
	226	0.396	0.769	0.379	0.000	0.083	0.468	0.526
	230	0.396	0.135	0.506	1.000	0.898	0.516	0.461
	234	0.038	0.038	0.092	0.000	0.019	0.016	0.013
	238	0.000	0.000	0.017	0.000	0.000	0.000	0.000
Smm_22	**213**	0.000	0.000	0.000	0.009	0.000	0.000	0.000
	217	0.000	0.000	0.006	0.167	0.046	0.000	0.000
	221	0.000	0.000	0.000	0.000	0.000	0.017	0.000
	225	0.000	0.000	0.006	0.000	0.000	0.017	0.013
	229	0.000	0.000	0.000	0.009	0.009	0.350	0.788
	233	0.091	0.000	0.071	0.000	0.000	0.267	0.188
	237	0.118	0.087	0.024	0.000	0.000	0.050	0.013
	241	0.027	0.000	0.119	0.000	0.000	0.000	0.000
	245	0.000	0.000	0.000	0.065	0.019	0.017	0.000
	249	0.064	0.000	0.024	0.093	0.046	0.200	0.000
	253	0.045	0.000	0.042	0.000	0.009	0.000	0.000
	257	0.136	0.109	0.083	0.241	0.241	0.033	0.000
	261	0.055	0.065	0.077	0.028	0.019	0.000	0.000
	265	0.027	0.000	0.161	0.019	0.046	0.017	0.000
	269	0.009	0.087	0.095	0.019	0.102	0.033	0.000
	273	0.000	0.000	0.030	0.315	0.185	0.000	0.000
	277	0.009	0.022	0.048	0.028	0.167	0.000	0.000
	281	0.045	0.043	0.048	0.009	0.000	0.000	0.000
	285	0.009	0.043	0.024	0.000	0.000	0.000	0.000
	289	0.055	0.152	0.048	0.000	0.000	0.000	0.000
	293	0.027	0.196	0.000	0.000	0.000	0.000	0.000
	297	0.009	0.043	0.012	0.000	0.000	0.000	0.000
	301	0.018	0.000	0.006	0.000	0.000	0.000	0.000
	305	0.000	0.022	0.006	0.000	0.000	0.000	0.000
	309	0.009	0.000	0.000	0.000	0.000	0.000	0.000
	313	0.036	0.022	0.018	0.000	0.000	0.000	0.000
	317	0.036	0.000	0.024	0.000	0.000	0.000	0.000
	321	0.082	0.087	0.006	0.000	0.000	0.000	0.000
	325	0.045	0.000	0.012	0.000	0.093	0.000	0.000
	329	0.045	0.022	0.006	0.000	0.009	0.000	0.000
	333	0.000	0.000	0.006	0.000	0.009	0.000	0.000
Sfo_18	**148**	0.445	0.500	0.418	0.778	0.550	0.589	0.434
	154	0.555	0.500	0.582	0.222	0.450	0.411	0.566
Sco_105	**176**	0.018	0.000	0.432	0.509	0.167	0.016	0.000
	180	0.860	0.963	0.477	0.491	0.583	0.177	0.688
	184	0.123	0.037	0.085	0.000	0.250	0.790	0.313
	188	0.000	0.000	0.006	0.000	0.000	0.016	0.000

Sco_106	186	0.000	0.000	0.006	0.000	0.000	0.000	0.000
	190	0.137	0.519	0.108	0.102	0.270	0.000	0.000
	194	0.000	0.000	0.000	0.176	0.040	0.000	0.000
	198	0.000	0.000	0.012	0.000	0.000	0.000	0.000
	206	0.069	0.000	0.012	0.000	0.000	0.000	0.000
	210	0.010	0.000	0.036	0.000	0.000	0.000	0.000
	214	0.088	0.037	0.078	0.000	0.000	0.000	0.000
	222	0.078	0.056	0.301	0.000	0.230	0.054	0.000
	226	0.343	0.259	0.259	0.056	0.060	0.554	0.500
	230	0.039	0.074	0.018	0.000	0.040	0.179	0.230
	234	0.235	0.056	0.060	0.083	0.230	0.214	0.270
	238	0.000	0.000	0.006	0.583	0.070	0.000	0.000
	242	0.000	0.000	0.006	0.000	0.000	0.000	0.000
	246	0.000	0.000	0.096	0.000	0.000	0.000	0.000
	250	0.000	0.000	0.000	0.000	0.060	0.000	0.000
Sco_215	288	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Sco_220	324	0.085	0.420	0.082	0.000	0.000	0.000	0.000
	328	0.000	0.000	0.012	0.000	0.000	0.000	0.000
	332	0.000	0.000	0.006	0.000	0.000	0.000	0.000
	344	0.000	0.000	0.006	0.000	0.000	0.000	0.000
	348	0.021	0.020	0.129	0.625	0.265	0.000	0.000
	352	0.106	0.420	0.135	0.125	0.392	0.648	0.208
	356	0.000	0.000	0.394	0.000	0.049	0.000	0.000
	360	0.138	0.000	0.029	0.009	0.020	0.074	0.028
	364	0.191	0.000	0.059	0.214	0.059	0.130	0.083
	368	0.426	0.140	0.076	0.000	0.000	0.000	0.000
	372	0.021	0.000	0.029	0.000	0.000	0.037	0.000
	376	0.000	0.000	0.024	0.000	0.127	0.000	0.000
	380	0.000	0.000	0.000	0.000	0.000	0.111	0.681
	388	0.011	0.000	0.018	0.027	0.078	0.000	0.000
	392	0.000	0.000	0.000	0.000	0.010	0.000	0.000
Omm_1128	340	0.029	0.000	0.047	0.000	0.000	0.000	0.000
	344	0.275	0.109	0.105	0.018	0.020	0.000	0.000
	348	0.176	0.326	0.256	0.307	0.275	0.161	0.000
	352	0.480	0.565	0.576	0.386	0.637	0.839	1.000
	356	0.039	0.000	0.017	0.289	0.069	0.000	0.000

Table 2-12: Allele frequencies of all loci in 7 populations found by hierarchical STRUCTURE analysis. Calculated in GENALEX 6.

Pop	Locus	Na	Ho	He	F
Cb	Sco_102	3.000	0.623	0.504	-0.236
	Sco_216	4.000	0.698	0.656	-0.065
	Smm_22	22.000	0.964	0.929	-0.038
	Sfo_18	2.000	0.382	0.494	0.227
	Sco_105	3.000	0.281	0.246	-0.143
	Sco_106	8.000	0.804	0.788	-0.020
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	8.000	0.787	0.744	-0.059
	Omm_1128	5.000	0.510	0.660	0.228
Wca	Sco_102	2.000	0.435	0.340	-0.278
	Sco_216	4.000	0.308	0.385	0.202
	Smm_22	14.000	0.913	0.892	-0.023
	Sfo_18	2.000	0.556	0.500	-0.111
	Sco_105	2.000	0.074	0.071	-0.038
	Sco_106	6.000	0.741	0.651	-0.138
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	4.000	0.560	0.627	0.107
	Omm_1128	3.000	0.522	0.562	0.072
Mi	Sco_102	3.000	0.560	0.559	-0.002
	Sco_216	5.000	0.540	0.592	0.087
	Smm_22	25.000	0.821	0.920	0.107
	Sfo_18	2.000	0.506	0.486	-0.040
	Sco_105	4.000	0.523	0.578	0.096
	Sco_106	13.000	0.699	0.809	0.137
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	13.000	0.753	0.791	0.048
	Omm_1128	5.000	0.640	0.590	-0.084
Ra	Sco_102	2.000	0.400	0.413	0.030
	Sco_216	1.000	0.000	0.000	#N/A
	Smm_22	12.000	0.833	0.800	-0.042
	Sfo_18	2.000	0.333	0.346	0.036
	Sco_105	2.000	0.596	0.500	-0.193
	Sco_106	5.000	0.593	0.608	0.026
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	5.000	0.625	0.547	-0.143
	Omm_1128	4.000	0.807	0.673	-0.200
Du	Sco_102	2.000	0.296	0.324	0.087
	Sco_216	3.000	0.204	0.186	-0.095
	Smm_22	14.000	0.870	0.854	-0.020
	Sfo_18	2.000	0.420	0.495	0.152
	Sco_105	3.000	0.611	0.569	-0.073
	Sco_106	8.000	0.800	0.806	0.007
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	8.000	0.765	0.747	-0.023
	Omm_1128	4.000	0.549	0.513	-0.069
Lli	Sco_102	2.000	0.276	0.400	0.310
	Sco_216	3.000	0.387	0.515	0.248
	Smm_22	10.000	0.733	0.761	0.036
	Sfo_18	2.000	0.536	0.484	-0.107
	Sco_105	4.000	0.387	0.343	-0.127
	Sco_106	4.000	0.750	0.613	-0.224
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	5.000	0.407	0.544	0.251
	Omm_1128	2.000	0.179	0.270	0.338
Uli	Sco_102	2.000	0.325	0.462	0.297
	Sco_216	3.000	0.395	0.511	0.227
	Smm_22	4.000	0.375	0.344	-0.089
	Sfo_18	2.000	0.553	0.491	-0.125
	Sco_105	2.000	0.525	0.430	-0.222
	Sco_106	3.000	0.730	0.624	-0.169
	Sco_215	1.000	0.000	0.000	#N/A
	Sco_220	4.000	0.528	0.486	-0.087
	Omm_1128	1.000	0.000	0.000	#N/A

Table 2-13: Basic statistical data on number of alleles (*Na*), observed (*Ho*) and expected (*He*) heterozygosity and fixation index (*F*) for each locus in each population found by hierarchical STRUCTURE analysis. Calculated in Genalex 6.

Population	Cb$_p$	Wca$_p$	Mi$_p$	Ra$_p$	Hi$_p$	Lli$_p$	Uli$_p$	Global
N	57	28	89	58	54	31	41	358
pw *Fst*, Cb$_p$		*	*	*	*	*	*	N/A
pw *Fst*, Wca$_p$	0.0988		*	*	*	*	*	N/A
pw *Fst*, Mi$_p$	0.0782	0.1218		*	*	*	*	N/A
pw *Fst*, Ra$_p$	0.2206	0.3355	0.1787		*	*	*	N/A
pw *Fst*, Hi$_p$	0.1168	0.2198	0.1015	0.1125		*	*	N/A
pw *Fst*, Lli$_p$	0.1708	0.2644	0.1648	0.3036	0.1569		*	N/A
pw *Fst*, Uli$_p$	0.1776	0.2734	0.191	0.3555	0.2269	0.1485		N/A
Mean allelic richness, FSTAT	5.433	4.207	5.952	3.339	4.454	3.462	2.292	6.133111
Total private alleles, GENALEX	1	0	8	1	2	1	0	13
He, GENALEX	0.558	0.448	0.592	0.432	0.499	0.437	0.372	0.476712
Fst, FSTAT W&C								0.178
Fis, FSTAT W&C	0.004	0.001	0.059	-0.069	0.005	0.087	-0.011	
#loci	9	9	9	9	9	9	9	9
Out of HWE? GENEPOP	No	No	Yes	No	No	No	No	Yes
Het def, GENEPOP	No	No	Yes, 3 loci	No	No	No	No	N/A
Het excess, GENEPOP	No	No	No	No	No	No	No	N/A
Null Alleles? MICROCHECKER	Yes, 1 locus	No	Yes, 2 loci	No	No	No	No	N/A
LD? FSTAT	No	No	Yes, 1 pair	No	No	No	Yes, 1 pair	N/A
Bottleneck? TPM, alpha <0.05	No	No	No	Yes	Yes	No	Yes	N/A
Bottleneck? TPM, alpha <0.01	No	No	No	Yes	Yes	No	No	N/A

Table 2-14: Summary of pertinent tests for each population found by hierarchical STRUCTURE analysis. Pairwise-*Fst* was calculated in FSTAT 2.9.3 using 420 random permutations. Significance values for *Fst,* HWE, heterozygote deficiency and heterozygote excess were obtained using the sequential Bonferroni procedure.

Cb

locus	HWE	Het def	Het exc
Sco_102	0.0461	0.9713	0.0363
Sco_216	0.0986	0.5041	0.5096
Smm_22	0.4406	0.841	0.2127
Sfo_18	0.1045	0.0683	0.9794
Sco_105	0.6864	1	0.2943
Sco_106	0.6823	0.4014	0.6085
Sco_215	-	-	-
Sco_220	0.6328	0.9143	0.1117
Omm_1128	0.006	0.0025	0.9978

All
chi2   29.6251
Df   16
Prob: 0.02

Hi

locus	HWE	Het def	Het exc
Sco_102	0.6708	0.3681	0.8717
Sco_216	1	1	0.5683
Smm_22	0.0607	0.7661	0.257
Sfo_18	0.2707	0.1945	0.9247
Sco_105	0.9734	0.7108	0.3132
Sco_106	0.1273	0.3438	0.6568
Sco_215	-	-	-
Sco_220	0.5936	0.6936	0.3369
Omm_1128	0.8498	0.6606	0.3693

All
chi2   14.5611
Df   16
Prob: 0.557

Wca

locus	HWE	Het def	Het exc
Sco_102	0.5379	1	0.2874
Sco_216	0.2023	0.1141	0.9069
Smm_22	0.0122	0.089	0.9207
Sfo_18	0.7102	0.8082	0.4622
Sco_105	1	1	0.9811
Sco_106	0.6377	0.9466	0.1129
Sco_215	-	-	-
Sco_220	0.3514	0.1828	0.8863
Omm_1128	0.8884	0.4105	0.6727

All
chi2   17.1624
Df   16
Prob: 0.3752

Lli

locus	HWE	Het def	Het exc
Sco_102	0.1515	0.0959	0.9863
Sco_216	0.146	0.0895	0.9578
Smm_22	0.4581	0.0259	0.9737
Sfo_18	0.7078	0.8065	0.4705
Sco_105	0.239	0.7966	0.3265
Sco_106	0.2985	0.9004	0.1088
Sco_215	-	-	-
Sco_220	0.0009	0.0215	0.9798
Omm_1128	0.1133	0.1131	0.9919

All
chi2   33.4272
Df   16
Prob: 0.0065

Mi

locus	HWE	Het def	Het exc
Sco_102	0.0607	0.0745	0.9262
Sco_216	0.1663	0.198	0.8199
Smm_22	0.014	0	0.9986
Sfo_18	0.8244	0.7043	0.4645
Sco_105	0.0047	0.0016	0.9984
Sco_106	0.0021	0	0.9999
Sco_215	-	-	-
Sco_220	0.0251	0	1
Omm_1128	0.4725	0.3543	0.6552

All
chi2   50.0436
Df   16
Prob: 0

Uli

locus	HWE	Het def	Het exc
Sco_102	0.0844	0.0523	0.9896
Sco_216	0.1389	0.0841	0.9548
Smm_22	1	0.8014	0.4493
Sfo_18	0.5285	0.8471	0.3608
Sco_105	0.2711	0.9615	0.1681
Sco_106	0.5829	0.9263	0.0869
Sco_215	-	-	-
Sco_220	0.0345	0.0145	0.9862
Omm_1128	-	-	-

All
chi2   20.5907
Df   14
Prob: 0.1126

Ra

locus	HWE	Het def	Het exc
Sco_102	0.7544	0.5051	0.7379
Sco_216	-	-	-
Smm_22	0.2364	0.712	0.3182
Sfo_18	0.7056	0.5074	0.7743
Sco_105	0.1914	0.9525	0.1285
Sco_106	0.189	0.7479	0.2521
Sco_215	-	-	-
Sco_220	0.1892	0.7534	0.2621
Omm_1128	0.0323	0.9847	0.0156

All
chi2   20.9828
Df   14
Prob: 0.1021

Table 2-15: *P*-value results from Hardy-Weinberg exact tests across all loci and each population found by hierarchical STRUCTURE analysis, as calculated in GENEPOP 4. Highlighted values are significant after sequential Bonferroni correction.

	Cbp	Wcap	Mip	Rap	Hip	Llip	Ulip	All
Sco102 X Sco216	0.06131	0.54782	0.15754	NA	0.46944	0.69206	0.28929	0.11448
Sco102 X Smm22	0.23929	1	0.26825	0.22421	0.13056	0.67083	0.60675	0.25099
Sco102 X Sfo18	0.19444	0.88095	0.98948	0.41667	0.97738	0.04226	0.28294	0.66548
Sco102 X Sco105	0.86329	0.65417	0.67143	0.43413	0.63413	0.05635	0.51825	0.65179
Sco102 X Sco106	0.37857	0.3996	0.12242	0.13214	0.04722	0.2623	0.28075	0.01448
Sco102 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Sco102 X Sco220	0.68333	0.41548	0.35238	0.38988	0.01726	0.56806	0.66131	0.22103
Sco102 X Omm1128	0.5502	0.12083	0.56667	0.59782	0.09028	0.48433	NA	0.33056
Sco216 X Smm22	0.95556	0.6748	0.92024	NA	0.05933	0.925	0.71766	0.86171
Sco216 X Sfo18	0.01964	0.86885	0.27222	NA	0.88135	0.47679	0.41171	0.23155
Sco216 X Sco105	0.22361	0.02778	0.5369	NA	0.08929	0.84663	0.63869	0.12262
Sco216 X Sco106	0.30218	0.0881	0.00179	NA	0.98651	0.14782	0.1627	0.00516
Sco216 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Sco216 X Sco220	0.51171	0.0127	0.63373	NA	0.65556	0.07937	0.48909	0.21448
Sco216 X Omm1128	0.01151	0.90317	0.97222	NA	0.45278	0.18333	NA	0.45099
Smm22 X Sfo18	0.57282	0.63433	0.96806	0.48075	0.92659	0.9	0.11766	0.92718
Smm22 X Sco105	0.68413	0.26567	0.35417	0.19187	0.68492	0.13889	0.70298	0.20714
Smm22 X Sco106	0.50675	0.00675	1	0.03313	0.04266	0.13452	0.4756	0.00476
Smm22 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Smm22 X Sco220	0.19206	0.4373	0.03988	0.0621	0.45714	0.60397	0.90218	0.0619
Smm22 X Omm1128	0.87063	1	0.31468	0.03234	0.54325	0.04226	NA	0.10734
Sfo18 X Sco105	0.65615	0.58095	0.27917	0.63591	0.09028	0.22738	0.1994	0.14841
Sfo18 X Sco106	0.33016	0.06845	0.0377	0.00476	0.38829	0.58175	0.03075	0.00119
Sfo18 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Sfo18 X Sco220	0.2129	0.84067	0.12758	0.0119	0.27599	0.72341	0.10794	0.02778
Sfo18 X Omm1128	0.60694	0.78393	0.40476	0.1625	0.1369	0.62718	NA	0.34365
Sco105 X Sco106	0.03433	0.61448	0.84206	0.0004	0.74425	0.25119	0.91786	0.14008
Sco105 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Sco105 X Sco220	0.00179	0.22917	0.0002	0.0121	0.00754	0.02599	0.0002	0.0002
Sco105 X Omm1128	0.73433	0.44008	0.4996	0.04365	0.89683	0.5371	NA	0.54484
Sco106 X Sco215	NA	NA	NA	NA	NA	NA	NA	NA
Sco106 X Sco220	0.05456	0.03333	0.00357	0.4125	0.99663	0.48115	0.55377	0.03472
Sco106 X Omm1128	0.48512	0.12758	0.73968	0.03413	0.26687	0.05099	NA	0.05694
Sco215 X Sco220	NA	NA	NA	NA	NA	NA	NA	NA
Sco215 X Omm1128	NA	NA	NA	NA	NA	NA	NA	NA
Sco220 X Omm1128	0.89702	0.2	0.95238	0.09563	0.85853	0.71329	NA	0.90694

Table 2-16: *P*-value results from genotypic disequilibrium tests across all loci and each population found by hierarchical STRUCTURE analysis, as calculated in FSTAT 2.9.3 using 5040 random permutations. Highlighted values are significant after sequential Bonferroni correction.
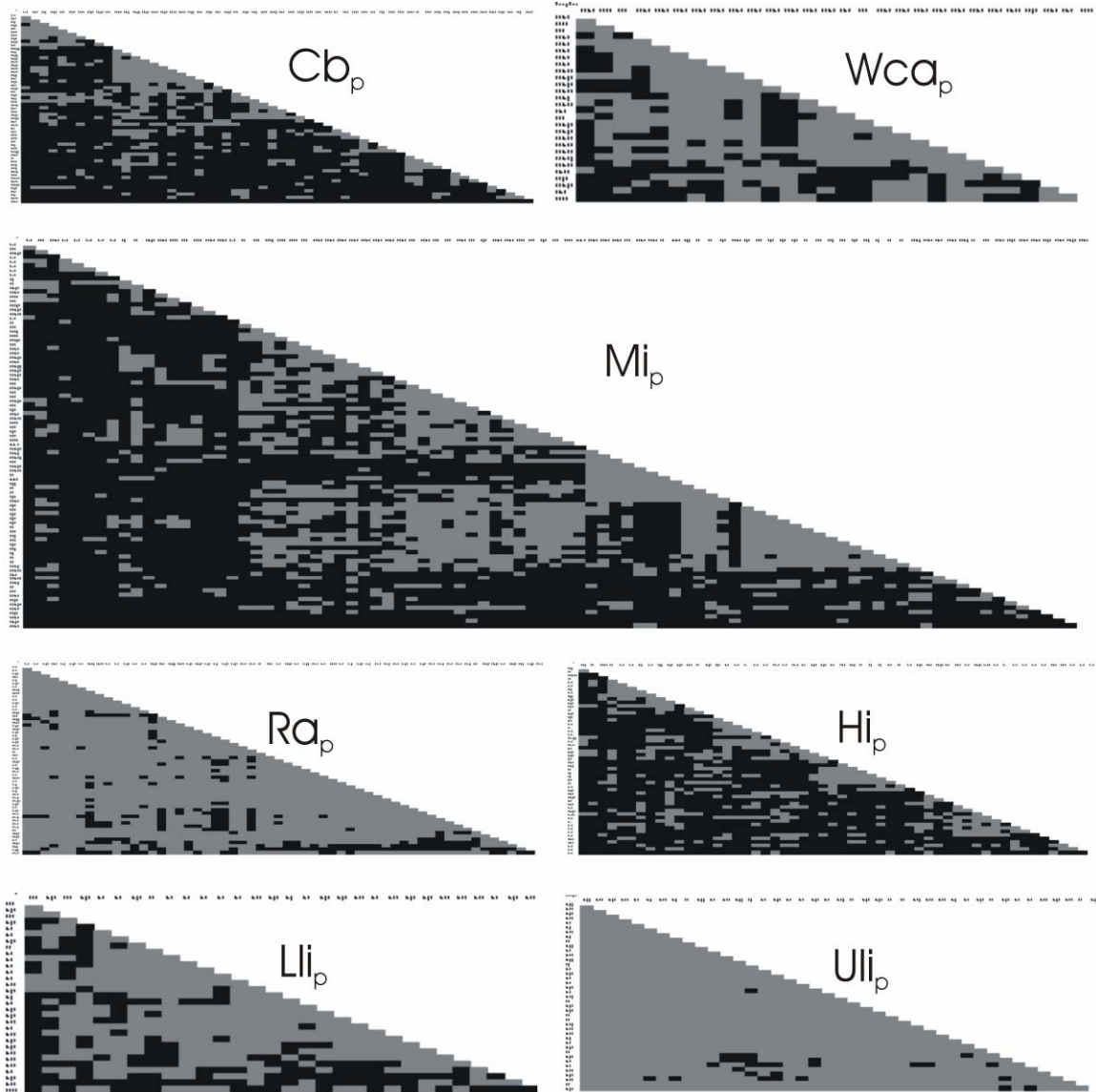
Figure 2-7: Family structure in each population found by hierarchical STRUCTURE analysis, as calculated in KINSHIP 1.3.1 using 1000 simulated pairs. Each half-matrix contains blocks indicating significance or non-significance for fullsib relatedness between individuals in a population. Grey indicates significant values at $\alpha = 0.05$. These individuals are assumed to be fullsibs. Black indicates non-significance between individuals for fullsib occurrence. These individuals are assumed to be unrelated.
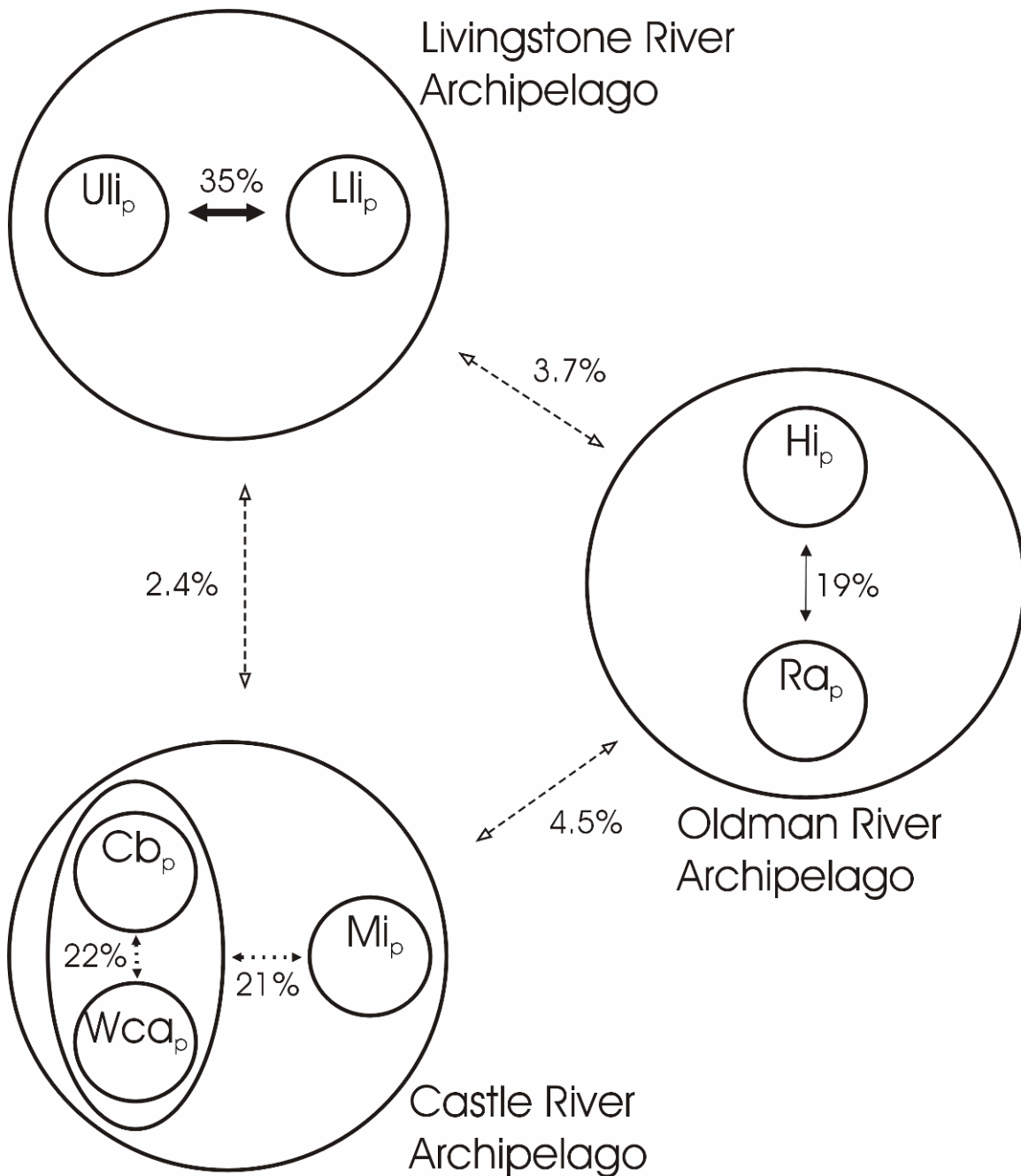
Figure 2-8: Summary of admixture rates between and within archipelagos. Dashed lines with open arrowheads represent the lowest rates at 0-10%. Solid, thin lines with closed arrowheads represent rates between 10 and 20%. Dotted lines with closed arrowheads represent rates between 20 and 30%. Bold solid lines with closed arrowheads represent the highest rates, at 30-40%.

Locus	Sco102	Sco216	Smm22	Sfo18	Sco105	Sco106	Sco215	Sco220	Omm1128	Global
LD w/Sco102		No	No	No	No	No	No	No	No	N/A
LD w/Sco216	-		No	No	No	No	No	No	No	N/A
LD w/Smm22	-	-		No	No	No	No	No	No	N/A
LD w/Sfo18	-	-	-		No	No	No	No	No	N/A
LD w/Sco105	-	-	-	-		No	No	2 pops	No	N/A
LD w/Sco106	-	-	-	-	-		No	No	No	N/A
LD w/Sco215	-	-	-	-	-	-		No	No	N/A
LD w/Sco220	-	-	-	-	-	-	-		No	N/A
LD w/Omm1128	-	-	-	-	-	-	-	-		N/A
Total Allele #	3	5	28	2	3	11	1	13	5	71
Mean allelic richness, FSTAT	2.300	3.070	12.360	2.000	2.680	6.510	1.000	6.360	3.390	4.408
Total Private alleles	0	1	3	0	0	5	0	4	0	13.000
Out of HWE? GENEPOP	No	No	No	No	No	No	No	No	No	
Het def, GENEPOP	No	No	1 pop	No	No	1 pop	No	1 pop	No	
Het excess, GENEPOP	No	No	No	No	No	No	No	No	No	
Null Alleles? MICROCHECKER	No	No	1 pop	No	No	1 pop	No	No	1 pop	
Fst	0.133	0.255	0.147	0.064	0.26	0.166	#N/A	0.247	0.119	0.177
#pops	7	7	7	7	7	7	7	7	7	8
#loci	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	9

Table 2-17: Summary of pertinent tests for each locus, using the 7 populations found by hierarchical STRUCTURE analysis. Significance values for HWE, heterozygote deficiency and heterozygote excess were obtained using the sequential Bonferroni procedure.
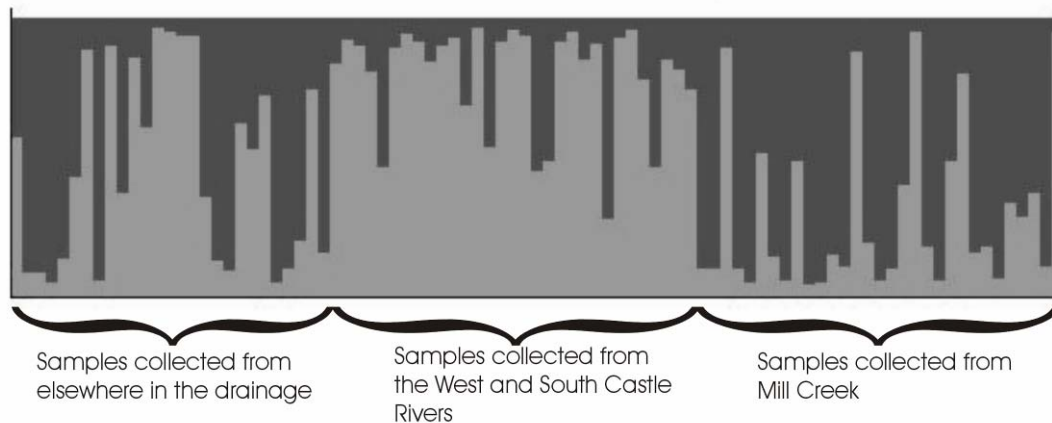
K	L(K)	s.d.	L'(K)	I L"K I	ΔK
1	-1974.46	1.128618			
2	-1984.09	7.684682	-9.63	66.06	8.596322
3	-2059.78	36.81581	-75.69	158.37	4.301684
4	-1977.1	5.529919	82.68	88.51	16.00566
5	-1982.93	7.634723	-5.83		

Figure 2-10: Output from STRUCTURE analysis of the Mip population. The chart represents assignment values if K=2, which appears to separate Mill Creek samples from those collected in (mostly) the South Castle River. As shown in the table, neither ΔK or the highest mean value of $\ln \Pr(X|K)$ suggested that this was the most likely solution.

Sca group

locus	HWE
Sco_102	0.3421
Sco_216	0.0701
Smm_22	0.0539
Sfo_18	0.3278
Sco_105	0.6945
Sco_106	0.0722
Sco_215	-
Sco_220	0.0192
Omm_1128	0.5916

All (Fisher's method)
chi2 : 30.4734
Df : 16
Prob: 0.0157

Mi group

locus	HWE
Sco_102	0.0812
Sco_216	0.8571
Smm_22	0.1509
Sfo_18	0.1362
Sco_105	0.043
Sco_106	0.3244
Sco_215	-
Sco_220	0.0261
Omm_1128	0.3798

All (Fisher's method)
chi2 : 30.8684
Df : 16
Prob: 0.0140

Table 2-18: *P*-value results from Hardy-Weinberg exact tests across all loci for separated Mill and South Castle-assigned groups within the Mi$_p$ population, as calculated in GENEPOP 4. No values are significant after sequential Bonferroni correction.

Chapter 3

Combining clustering and assignment tests: a refined approach to performing mixed-migrant assignment analyses using bull trout (*Salvelinus confluentus*) as a model species.

Introduction

   Mixed-migrant assignment tests attempt to assign individuals of unknown origin back to a given reference population. This can be an extremely powerful tool for a wide realm of ecological studies or applied wildlife management. In many previous mixed-migrant assignment studies, given reference populations are defined *a priori* by site-of-origin (Castric and Bernatchez, 2004; Taylor and Costello, 2006). Such a definition may become problematic if individuals from those sites are not genetically divergent or if dispersal between sites is common. Because assignment tests attempt to match genotypes of individuals to a "genotypic profile" of a reference population, it stands to reason that reference populations defined by an evolutionary or genetic framework (Waples and Gagiotti, 2006; Chapter 2) should be more suitable as given reference populations. In fisheries management, such an attempt has been made by defining major "regions and subregions" as reference populations. These analyses are conducted by genetic distance analyses used in tandem with self-assignment tests (Van Doornik *et al*., 2007). Clustering methods, such as those performed in hierarchical STRUCTURE analysis (Pritchard *et al*., 2000; Vaha *et al*., 2007), cluster individuals into populations based on genotype alone using no *a priori* information of sampling location; therefore, this may represent another method used to define reference populations from a genetic framework, possibly

increasing success of mixed-migrant assignment tests. The usefulness of genetic

clustering methods performed in tandem with self assignment (Cegelski *et al.*, 2003) and

assignment of individuals of unknown origin (Frantz *et al.*, 2006) tests has been shown.

However, such studies are rare and have not addressed the relative assignment success

between defining populations by sampling location, and genetic clustering methods.

Chapter 2 successfully defined populations from a genetic framework using

hierarchical STRUCTURE analysis. In this chapter, mixed-migrant and self-assignment

tests wil be performed using reference populations defined by stream-of-origin and by

hierarchical STRUCTURE analysis methods. It is possible that populations defined by

hierarchical STRUCTURE analysis will provide higher success and confidence for

assignment tests. Secondary aims are to further understand migratory tendencies of

populations of bull trout defined in the previous chapter. These results will be used for

interpretation of how stream access (in particular, seasonal barriers) shapes population

structure and migratory tendencies in salmonids and for valuable local conservation

information.

This study outlines a refined approach to performing assignment tests in order to

maximize assignment success with spatial resolution. Such results may be used to further

appropriate conservation efforts for management of highly migratory species.

Methods

Self-assignment tests

For sampling methods and genetic analyses of reference populations, see the methods section of Chapter 2.

Self-assignment tests were used in the context of this study to determine the reliability of reference populations defined by STRUCTURE or stream-of-origin for the mixed-migrant analysis. The leave-one-out procedure was used in GENECLASS 2 (Piry *et al.*, 2004) to determine the population of origin and probability of an individual being sampled from any given reference population (Ranalla and Mountain, 1997) without the individual's genotype included in any reference sample (Efron, 1983).

Rates of self assignment and mean probability of assignment for individuals belonging to each population (i.e. mean assignment values) were evaluated as measures of assignment success. A paired t-test was used to determine if rates of self assignment were significantly different between populations defined by the two methods. A Mann-Whitney U-test was used to determine if mean individual self-assignment probabilities were different between populations defined by the two methods. Secondly, assignment confidence was evaluated by performing exclusion-based assignment tests (Paetkau *et al.*, 2004) using a pre-defined alpha of 0.05 (i.e. there is a 5% chance of a false-positive for excluding a population) and 10,000 simulated individuals. From this, rates of exclusion to assigned population were examined between the two methods.

Mixed-migrant tests

A total of 85 migrant bull trout were sampled from the main-stem of the Oldman (OMR) and Castle (CA) Rivers by angling, from the Oldman River Reservoir (ORR) by gillnet and from below the dam in the Oldman River tailwater (TW) by jetboat electrofisher in the summers of 2006 and 2007 (Figure 3-1). All fish were processed and genotyped in the same manner as those caught in reference populations (Chapter 2: Methods) except for tailwater fish, which received spaghetti tags for future conservation efforts.

Confidence of assignment tests was evaluated by comparing mean assignment values of individuals between the two methods with a Mann-Whitney U-test. Exclusion-based methods were also used to determine if individuals assigned to any specific population were also excluded from the population to which they were assigned. Exclusion rates were then compared between the two methods.

Differences between tests were also examined by geographical assignment concordance. The geographical regions of reference populations defined by stream-of-origin and STRUCTURE analysis can be seen in Figures 2-2 and 2-5, respectively. Fish that were not assigned to the same geographical region of origin between the two methods were identified.

Results

Self-assignment tests

A summary of self-assignment analyses can be found in Tables 3-1 and 3-2

Rates of self-assignment for populations defined by STRUCTURE were 27.3%

higher than those defined by stream-of-origin (Figure 3-2). The t-test revealed these to be

significantly different (Figure 3-2)

Self-assignment probabilities were also higher for populations defined by

STRUCTURE (*M*= 91.46%, *SE*=0.76%) than by stream-of-origin (*M*=82.04%,

SE=0.96%). The Mann-Whitney U-test revealed that these differences were significant

(Figure 3-3).

Exclusion-based methods revealed low exclusion rates to assigned population in

both methods. Only 3 out of 358 individuals were excluded from the population to which

they were assigned when populations were defined by STRUCTURE, while 14 of 364

individuals were excluded from the population to which they were assigned when

populations were defined by stream-of-origin.

Mixed-migrant tests

A summary of mixed-migrant analyses can be found in Tables 3-3 and 3-4.

Probabilities of assignment of migrants to a specific reference population were

higher for reference populations defined by STRUCTURE (*M*= 91.97%, *SE*=1.51%) than

by stream-of-origin (*M*=79.69%, *SE*=1.87%). The Mann-Whitney U-test revealed that

these differences were significant (Figure 3-4).

Exclusion-based methods revealed low exclusion rates to assigned population in

both methods. Only 1 out of 85 migrant individuals were excluded from the population to

which they were assigned when populations were defined by STRUCTURE, while no

migrant individuals were excluded from the population to which they were assigned when populations were defined by stream-of-origin.

Geographical assignment concordance was 80% between the two methods (68 out of 85 individuals assigned to the same geographical area).

Assignment tests were further analyzed to determine migratory tendencies of populations defined by STRUCTURE. Collectively, the stock composition of the pooled migrant group was largely dominated by Castle archipelago fish, most of which originated from the Mi_p population (Figure 3-5). When broken into composition by sampling location, The TW and CA river sampling sites had nearly identical stock compositions, with ~2/3 of fish originating from the Mi_p and ~1/3 originating from the the Cb_p population (Figure 3-6). The ORR reservoir site consisted of similar proportions of fish originating from the Mi_p and Cb_p populations, with a single fish (6%) assigned back to the Wca_p population. All fish in these three sites were assigned back to populations originating in the Castle River archipelago. In the OMR site, 30% of fish caught were assigned back to the Castle River archipelago, with fish assigning to all three populations within. Of the remaining 70%, all but 2 fish (6%) were assigned to the Oldman River archipelago, with the vast majority assigning to the Hi_p population (Figure 3-6).

When plotted together, expected heterozygosity (Chapter 2) and number of migrants assigned to specific populations defined by STRUCTURE were significantly correlated (Linear regression ANOVA, p=0.004, r=0.9316) (Figure 3-7a,b).

Discussion

Self-assignment tests

Self-assignment rates and probabilities are dependent on genetic divergence and dispersal between given reference populations (Berry *et al*., 2004). Hierarchical STRUCTURE analysis removed all migrants and placed them within their own population, while grouping those streams with low genetic divergence. Such procedures should lead to high self-assignment rates and probabilities in reference populations. While assignment tests generated using STRUCTURE and GENECLASS were not identical (i.e., self-assignment rate was not 100% in all populations defined by STRUCTURE), they were highly concordant. As expected, self-assignment rates for stream-of origin were lower, and of similar value to those found in bull trout by Taylor and Costello (2006). Although self-assignment probability was lower for stream-of-origin, the difference was not as low as self-assignment rate. This is likely because probability is more sensitive to genetic divergence than mis-assignment. As a result, the stream-of-origin method likely yielded more admixed individuals between sites that are not particularly genetically divergent.

Mixed-migrant tests

As predicted, reference populations defined by STRUCTURE led to higher assignment probabilities for mixed-migrant stock assignment. This is not surprising, given that STRUCTURE methods yielded higher self-assignment rates and probabilities in given reference populations. The 20% disparity between congruent geographical

assignments implies assignment error in either method. Unfortunately, without real observational data it is impossible to explicitly determine the accuracy of each method; however, the higher assignment probabilities associated with the STRUCTURE method, together with the underlying assumptions by which populations are defined warrant higher confidence in this method. As such, the following discussion of conservation implications and migratory tendencies of bull trout populations will assume that the population structure outlined by the hierarchical STRUCTURE analysis, represents the true population structure for the system.

Conservation implications

Mixed-migrant tests indicate that there is a gradient of migratory tendencies between the three archipelagos. The Castle archipelago is by far the most migratory of the three, with fish assigning to populations with spawning locations up to hundreds of kilometers from where they were caught. All of the lower reaches in the drainage (e.g., the TW and the ORR) contained only Castle River fish.

The Oldman River Reservoir has existed for less than 20 years. In this time, it is possible that an incipient adfluvial life-history strategy has developed. Such a life-history shift has been documented in bull trout populations following reservoir creation (Bruce and Starr, 1985), but it is unknown how the migratory tendencies of such a life history will affect the genetics of reference populations, if at all (Neraas and Spruell, 2001). Salmonid fishes use a variety of cues to guide themselves back to their stream of origin, but depend mainly on olfactory stimuli (Dittman and Quinn, 1996). Finding olfactory or alternate cues for natal tributaries from a lentic environment, such as a reservoir may be

more or less difficult for bull trout undergoing spawning migrations, compared to those in fluvial environments. As a result, adfluvial or fluvial fish that must pass through the reservoir during spawning migrations might be expected to stray at rates different from their ancestors, affecting patterns of gene flow in the drainage. It is therefore unknown whether the system will undergo population divergence, homogenization or remain the same as a result, but future monitoring should continue in order to answer this question.

Tailwater fish were entirely composed of Castle origin fish. This supports the findings of local radio-telemetry studies, which observed that fish relocated from the tailwater to above the dam primarily, migrated up the Castle River (Fernet and O'Neil, 1997; Golder, 1998a). These highly fluvial fish may represent ancestral stocks that were historically found in the lower reaches of the Oldman River as far downstream as Lethbridge (Fitch, 1997). The fish that once occupied this stretch of river would have been ecologically important as apex predators spatially connecting cool (downstream) and coldwater (upstream) ecosystems (McCann *et al*., 2005) and genetically important as potential sources of gene flow with far-downstream bull trout populations (Waterton and St. Mary's) in the entire Oldman River drainage. The fish currently found in this stretch are commonly considered "stranded" fish which are not able to migrate to upstream sources above the dam to reproduce, and for which a mitigation plan is required (Brewin *et al*., 1999); however, it is possible that the migrants found in this stretch of river are not Castle archipelago fish *per se*, but the F1 or F2 offspring of fish that were stocked into Pincher Creek during relocation efforts of stranded fish in 1996 and 1997 (Golder, 1998b); if such a scenario were true, the assignment of the progeny to Castle River origin indicates that the these fish are products of a successful mitigation project using stranded

fish from Castle River ancestral stocks. Whether the original project was successful or not may never be known, but Pincher Creek currently contains a robust cohort of a single year class of juvenile fish that were products of a 2005 spawning event (Warnock, personal observations). It is unknown if such spawning events will continue in this creek, as 2005 contained high flows in late summer months, during typical bull trout migration periods. The creek should continue to be monitored for presence of juveniles and redd surveys should be conducted annually to determine if a self-sustaining incipient population is developing. An alternate strategy for mitigation involves strategic relocation to above-dam sources. Fish may be individually recaptured and identified by tags implanted during this study. These fish could then be relocated into the home range of the population to which they were individually assigned during late summer or fall spawning periods. Such a strategic approach could help to mitigate the loss of the genetic component of highly migratory groups, and supplement spawning aggregations of local populations.

Yet another interesting finding is that 10 out of 30 fish caught in the Oldman River (the stream that the Livingstone and Oldman archipelago tributaries drain into) assigned to Castle sources. This trend was not reciprocated, in that no Oldman or Livingstone archipelago fish were sampled in the Castle River. This implies that contemporary gene flow within the drainge may occur in a unidirectional manner from the Castle archipelago to the Oldman and Livingstone archipelagos.

Within the Castle archipelago, the fact that most migrants originate from the Mi_p population is not surprising as this population is of the highest effective size. Conversely, few migrants originated from the Wca_p population. This stresses even further caution in

protection of West Castle River bull trout, as addressed in the previous chapter. From our data, it is likely that the West Castle River receives very small runs of migrant bull trout.

The Oldman archipelago is the second most migratory archipelago, contributing most of the migrants sampled in the Oldman River. Unlike the Castle archipelago, however, migrants were all short-range, as none were found more than 50 km from potential spawning sites. Management of migrants from this archipelago should therefore only focus on the Oldman River above the Reservoir, and not on distal streams in the drainage or the reservoir. Within the archipelago, the same trends were observed as in the Castle in that the larger population, Hi_p, contributed the most migrants.

The Livingstone archipelago was found to contribute very few migrants (2), none of which were found outside of the Oldman River. One fish was assigned to each the above (Uli_p) and below-barrier (Lli_p) populations, however, assignment probabilities indicate that the Uli_p fish was of admixed ancestry with the Lli_p population (Potvin and Bernatchez, 2001). This provides evidence for the theory outlined in the previous chapter that the Uli_p may exchange genetic material with populations in other archipelagos via the Lli_p as a stepping-stone. Although this archipelago does not contribute a large number of migrants, the migrants it does contribute may be very important to the persistence of its populations.

Overall stock composition of migrants in the study area seems to largely be dominated by Castle archipelago fish, with the Mill Creek population representing the majority of these. As an apex predator, overall numbers of adult migrant individuals are comparatively low for this species. Since several populations contribute very small proportions of fish to this pool, it is likely that they depend upon very small returns for

spawning events. Such small effective sizes put these populations at higher risk of

extinction (Dunham and Rieman, 1999). Management strategies of the entire migrant

pool, particularly in the Oldman River area, should therefore be conservative, treating

each individual fish as if it were a member of one of those sensitive populations.

Stream access, population structure and migration

When expected heterozygosity was plotted against overall number of migrants

assigned to each population defined by STRUCTURE, a striking correlation was found

(Figures 3-7a,b). Effective population size in migrant salmonid populations is determined

by the overall number of migrants which return to reproduce successfully. It is likely that

the numbers of bull trout I sampled that were assigned to any given population are

reflective of such effective population size. Since effective population size is intimately

linked to intrapopulation diversity, this is the most likely explanation for the correlation

observed. This of course leads to the possibility that heterozygosity may be a powerful

predictor of adult migrant abundance or population estimates used in fisheries

management, whether the underlying cause be due to residency or reduced population

size in given reference populations.

The underlying causes of this observed variability in the inextricably linked

factors of migratory tendency, intra and inter-population diversity and effective

population size are likely governed by a suite of contemporary and historic variables.

Variables such as landscape features (Castric *et al.*, 2001), postglacial colonization routes

(Taylor *et al.*, 1999) anthropogenic disturbances and invasive species (Dunham and

Rieman, 1999) and distances between spawning areas of populations (Castric and

Bernatchez, 2003) have been examined as possible contributors to this observed variability. Qualitatively, some of the patterns observed can be related to a specific landscape feature: stream accessibility.

Collectively, a gradient of overall migrant numbers was observed within and between each archipelago. The Livingstone is a poorly accessible region, providing the least number of migrants to the drainage. The Castle is the most accessible, providing the greatest numbers, life-history flexibility and mobility of migrants. With exception of the West and, to a lesser extent, South Castle Rivers, the Castle archipelago has a variety of spawning grounds which are not located behind significant seasonal barriers, and of which, the major tributary drains into the lower stretch of the river. Conversely, all spawning tributaries of the Oldman River are located above a major seasonal barrier. The Livingstone archipelago is located above this barrier as well, but is more remote from the barrier than the Oldman tributaries and contains a large secondary barrier above one of the populations. Likewise, those populations that have the least accessible spawning grounds within each archipelago tend to provide the least migrants. There are two reasons why difficult access may tend to reduce the number of migrants observed. First, presence of barriers may lead to the evolution of a resident life history strategy (Northcote, 1992) as has occurred in the Uli_p. Second, the presence of barriers may reduce population size (Costello $et\ al.$, 2003) by decreasing the odds of migrants reaching spawning grounds. Indeed, populations such as Wca_p and Ra_p, which have difficult access to spawning grounds, tend to show fewer migrants.

Although this study cannot quantify "difficulty of access" to spawning locations, especially over long temporal scales, such an analysis may correlate extremely well with

measured population structure, size and migratory tendencies. Intimate analyses of migratory routes themselves may therefore be powerful predictors of the fine-scale genetic and migratory trends of salmonid populations over spatial networks. Several metrics that may be used for prediction in future studies are: linear stream distance to spawning sites, stream gradient, number of barriers, and degree of negotiability of each barrier therein.

Conclusions

In the previous chapter, the advantages of defining populations by hierarchical STRUCTURE analysis were outlined. The chapter demonstrates that defining populations by such methods have an additional advantage when performing mixed-migrant assignment analyses. By defining reference populations using a genetic clustering method, assignment confidence may be maximized with spatial resolution, providing the investigator with interpretable data on the genetic origin of individuals of unknown ancestry.

Using assignment tests performed in this manner, the migratory tendencies of populations within the study area were identified, and populations with strong and weak migrant presence were distinguished. The strength of the migratory contribution was correlated with intrapopulation diversity and presumably population size. Such methods may be used to guide conservation efforts on salmonid populations and the mixed-migrant groups they contribute to.

Previous studies have been concerned with describing how landscape features affect genetic structure of populations (Castric *et al*., 2001; Costello *et al*., 2003). Results of this study suggest the importance of further considering migration barriers which may operate on a scale of "negotiability" that is temporally variable in nature. Such landscape features have the potential to be quantified in order to predict patterns of contemporary salmonid population structure, size and migratory tendencies in a fine-scale spatial network.

References

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology, 13*(3), 551-561.

Brewin MK, Baayens DM, Hall C, Norris K. (1999). Identifying the need to mitigate the stranding of bull trout below the Oldman River dam, Alberta. Pp. 187-190. *In*: M.K. Brewin, A.J. Paul & M. Montina (ed.) Bull Trout II Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Bruce, P., & Starr, P. (1985). Vol. 11, Fisheries resources and fisheries potential of Williston Reservoir and its tributary streams - a preliminary overview. British Columbia Ministry of Environment, Fish. Tech. Circ. No. 69. 101 pp. + appendices.

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., & Bernatchez, L. (2004). Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. *Molecular Ecology, 13*(5), 1299-1312.

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

Cegelski, C. C., Waits, L. P., & Anderson, N. J. (2003). Assessing population structure and gene flow in Montana wolverines (Gulo gulo) using assignment-based approaches. *Molecular Ecology, 12*(11), 2907-2918.

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.

DeHaan, P., & Arden, W. (2007). Genetic analysis of bull trout in the Jarbidge River watershed, Nevada/Idaho. U.S. Fish and Wildlife Service. Reno, NV. 22 pp.

Deiner, K., Garza, J. C., Coey, R., & Girman, D. J. (2007). Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. *Conservation Genetics, 8*(2), 437-454.

Dittman, A. H., & Quinn, T. P. (1996). Homing in Pacific salmon: Mechanisms and ecological basis. *Journal of Experimental Biology, 199*(1), 83-91.

Dunham, J. B., & Rieman, B. E. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications, 9*(2), 642-655.

Efron, B. (1983). Estimating the error rate of a prediction rule - improvement on cross-validation. *Journal of the American Statistical Association, 78*(382), 316-331.

Fernet, D.A. & J.O. O'Neil. (1997). Use of radio telemetry to document seasonal movements, and spawning locations for bull trout in relation to a newly created reservoir. Pp. 427-434. *In:* W.C. Mackay, M.K. Brewin & M. Montina (ed.) Friends of the Bull Trout Conference Proceedings, Trout Unlimited Canada, Calgary, AB.

Frantz, A. C., Pourtois, J. T., Heuertz, M., Schley, L., Flamand, M. C., Krier, A., *et al*. (2006). Genetic structure and assignment tests demonstrate illegal translocation of red deer (*Cervus elaphus*) into a continuous population. *Molecular Ecology, 15*(11), 3191-3203.

Golder, & Associates. (1998a). The Oldman River bull trout radio telemetry study 1993-1996. Prepared by Golder Associates Limited. Calgary, AB. 31 pp. + 2 app.

Golder & Associates. (1998b). Bull trout (*Salvelinus confluentus*) transfer from the lower Oldman River and Picher Creek redd surveys (1996 and 1997). Prepared by Golder Associates Limited. Calgary, AB. 16 pp.

Kassler, T., & Mendel, G. (2007). Genetic characterization of bull trout from the Walla Walla River basin. Washington Dept. of Fish and Wildlife. Olympia, WA. 42 pp. + app.

McCann, K. S., Rasmussen, J. B., & Umbanhowar, J. (2005). The dynamics of spatially coupled food webs. *Ecology Letters, 8*(5), 513-523.

Neraas, L. P., & Spruell, P. (2001). Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology, 10*(5), 1153-1164.

Northcote, T. G. (1992). Migration and residency in stream salmonids some ecological considerations and evolutionary consequences. *Nordic Journal of Freshwater Research*(67), 5-17.

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.

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.

Potvin, C., & Bernatchez, L. (2001). Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed-stock analyses. *Molecular Ecology, 10*(10), 2375-2388.

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

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

Taylor, E. B., & Costello, A. B. (2006). Microsatellite DNA analysis of coastal populations of bull trout (*Salvelinus confluentus*) in British Columbia: zoogeographic implications and its application to recreational fishery management. *Canadian Journal of Fisheries and Aquatic Sciences, 63*(5), 1157-1171.

Taylor, E. B., Pollard, S., & Louie, D. (1999). Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Molecular Ecology, 8*(7), 1155-1170.

Vaha, J. P., Erkinaro, J., Niemela, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology, 16*(13), 2638-2654.

Van Doornik, D. M., Teel, D. J., Kuligowski, D. R., Morgan, C. A., & Casillas, E. (2007). Genetic analyses provide insight into the early ocean stock distribution and survival of juvenile coho salmon off the coasts of Washington and Oregon. *North American Journal of Fisheries Management, 27*(1), 220-237.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology, 15*(6), 1419-1439.
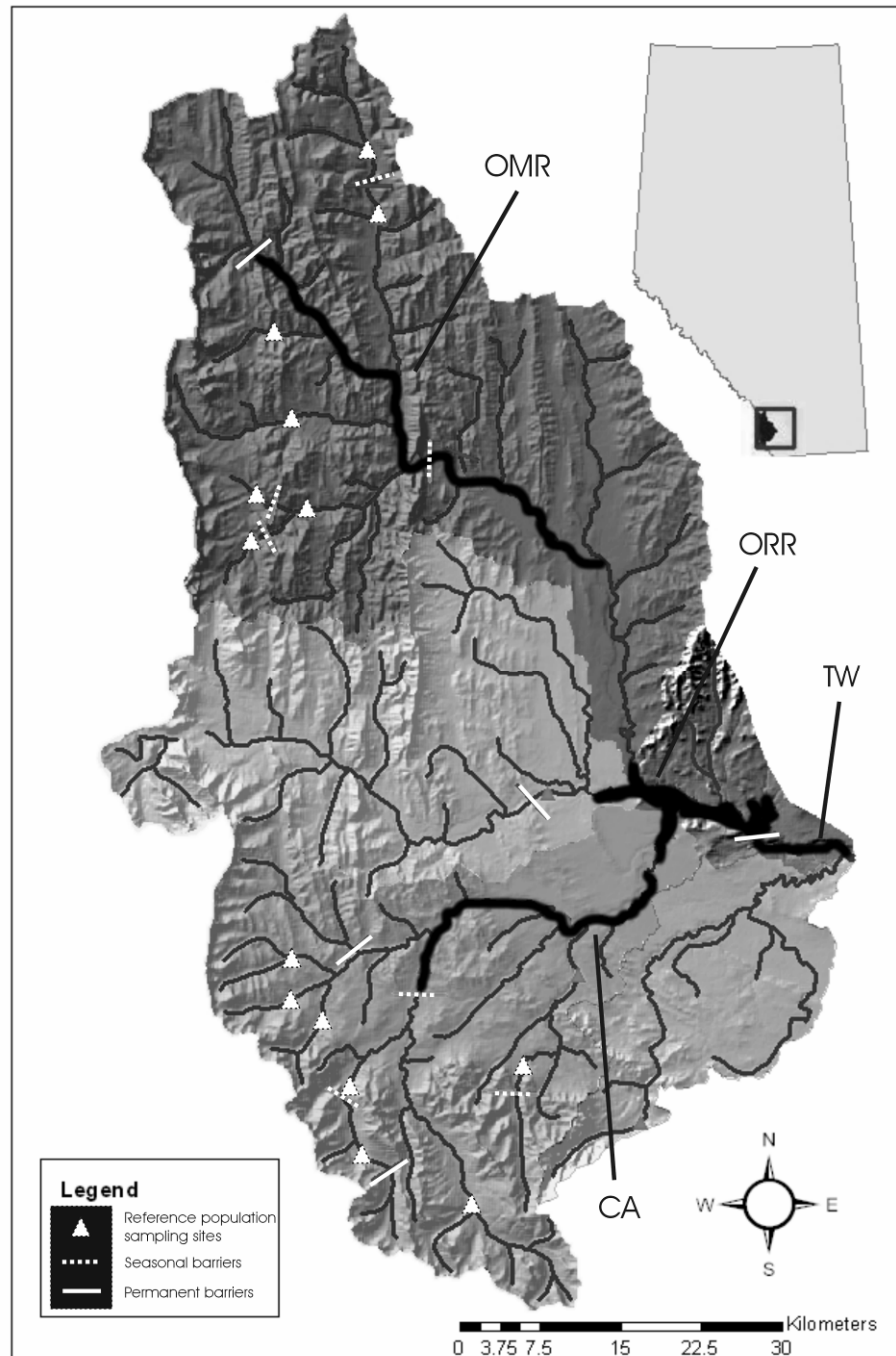
Appendix

Figure 3-1: Map of migrant sampling sites in the Oldman River complex.

Origin	Assigned sample	rank 1	score %	rank 2	score %
Cb	/CB_1	Ga	93.405	Cb	3.495
Cb	/CB_2	Ga	99.521	Cb	0.420
Cb	/CB_3	Cb	99.973	Lli	0.013
Cb	/CB_4	Ga	96.696	Cb	2.149
Cb	/CB_5	Lo	50.910	Cb	48.735
Cb	/CB_6	Cb	68.649	Wca	27.118
Cb	/CB_7	Cb	38.446	Ga	25.539
Cb	/CB_8	Cb	99.196	Sca	0.330
Cb	/CB_9	Sca	76.293	Sra	9.021
Cb	/CB_10	Cb	92.453	Wca	7.388
Cb	/CB_11	Cb	99.972	Lo	0.027
Cb	/CB_12	Du	80.143	Lo	8.156
Cb	/CB_13	Cb	78.057	Lo	11.882
Cb	/CB_14	Cb	99.991	Lo	0.005
Cb	/CB_16	Cb	98.282	Lo	1.423
Cb	/CB_17	Lo	34.939	Cb	33.053
Cb	/CB_18	Cb	65.191	Ga	34.056
Cb	/CB_19	Ga	87.699	Cb	6.829
Cb	/CB_20	Lo	92.571	Cb	6.955
Cb	/CB_21	Cb	34.070	Lo	33.610
Cb	/CB_22	Cb	99.961	Ga	0.022
Cb	/CB_23	Cb	99.645	Ga	0.315
Cb	/CB_24	Cb	53.958	Sca	42.512
Cb	/CB_25	Cb	96.264	Ga	3.732
Cb	/CB_26	Sca	82.676	Lo	15.718
Cb	/CB_28	Ga	54.546	Cb	42.160
Cb	/CB_29	Cb	89.278	Lo	5.890
Cb	/CB_30	Cb	99.218	Lo	0.608
Cb	/Cb_31	Cb	62.111	Lo	37.857
Ga	/J-GA-20	Ga	97.249	Cb	2.697
Ga	/J-GA-8	Cb	86.487	Ga	13.503
Ga	/R-GA-1	Lo	54.593	Mi	42.950
Ga	/J-GA-9	Ga	77.894	Sca	22.039
Ga	/R-GA-2	Cb	73.727	Ga	26.237
Ga	/J-GA-10	Ga	45.495	Sca	34.787
Ga	/R-GA-3	Lo	63.381	Cb	18.955
Ga	/J-GA-11	Cb	66.411	Ga	33.581
Ga	/R-GA-4	Ga	99.813	Cb	0.183
Ga	/J-GA-12	Ga	78.462	Cb	21.477
Ga	/R-GA-5	Ga	95.492	Wca	3.688
Ga	/J-GA-13	Ga	76.615	Wca	22.085
Ga	/J-GA-14	Wca	70.435	Sca	29.425
Ga	/J-GA-15	Sca	53.358	Ga	44.818
Ga	/J-GA-1	Ga	64.128	Cb	35.857
Ga	/J-GA-16	Cb	99.338	Ga	0.657
Ga	/J-GA-3	Cb	69.477	Ga	16.114
Ga	/J-GA-17	Ga	94.155	Cb	5.054
Ga	/J-GA-5	Wca	94.023	Cb	4.788
Ga	/J-GA-18	Cb	82.616	Ga	11.186
Ga	/J-GA-6	Ga	94.570	Cb	4.414
Ga	/J-GA-19	Ga	98.711	Cb	0.816
Ga	/J-Ga-7	Wca	95.471	Ga	4.507
Lo	/R-LO-14	Cb	99.977	Lo	0.013
Lo	/R-LO-2	Lo	85.865	Hi	13.887
Lo	/F-LO-1	Sca	75.347	Ga	14.752
Lo	/J-LO-1	Lo	81.362	Cb	17.043
Lo	/F-LO-2	Lo	96.500	Cb	2.310
Lo	/R-LO-4	Cb	88.252	Lo	11.461
Lo	/J-LO-2	Li	98.344	Lli	1.317
Lo	/F-LO-3	Cb	84.694	Lo	14.827

Origin	Assigned sample	rank 1	score %	rank 2	score %
Ra	/RA_13	Ra	44.417	Sra	33.557
Ra	/RA_14	Sra	70.460	Ra	25.467
Ra	/RA_15	Du	69.167	Sra	22.757
Ra	/RA_16	Du	95.208	Nra	3.449
Ra	/RA_17	Ra	88.975	Sra	10.928
Ra	/RA_18	Nra	66.720	Ra	30.675
Ra	/RA_19	Nra	95.525	Sra	3.995
Ra	/RA_20	Ra	46.506	Nra	41.338
Ra	/RA_21	Sra	51.040	Ra	48.884
Ra	/RA_22	Nra	96.938	Ra	1.582
Ra	/RA_23	Sra	70.210	Ra	29.734
Ra	/RA_24	Ra	64.803	Sra	34.795
Ra	/RA_25	Sra	55.060	Nra	44.617
Sra	/Sra-1	Sra	60.058	Ra	24.187
Sra	/Sra-2	Du	95.683	Hi	3.919
Sra	/Sra-3	Nra	35.259	Sra	33.319
Sra	/Sra-4	Sra	51.904	Ra	42.565
Sra	/Sra-5	Sra	74.990	Ra	24.958
Sra	/Sra-6	Ra	49.931	Nra	32.344
Sra	/Sra-7	Ra	68.056	Sra	30.510
Sra	/Sra-8	Ra	98.378	Sra	0.714
Sra	/Sra-9	Sra	75.172	Ra	21.558
Sra	/Sra-10	Ra	66.936	Sra	31.824
Sra	/Sra-11	Ra	73.396	Sra	24.949
Sra	/Sra-12	Sra	51.965	Ra	34.499
Sra	/Sra-13	Sra	80.181	Ra	16.435
Sra	/Sra-14	Sra	71.210	Ra	28.053
Sra	/Sra-16	Sra	45.130	Ra	38.129
Sra	/Sra-17	Sra	80.255	Ra	19.624
Sra	/Sra-18	Nra	91.365	Ra	6.724
Sra	/Sra-19	Ra	54.455	Sra	44.782
Nra	/J-Nra-1	Nra	50.540	Sra	31.909
Nra	/J-Nra-2	Nra	84.373	Sra	9.360
Nra	/J-Nra-3	Ra	73.122	Nra	23.388
Nra	/J-Nra-4	Nra	96.062	Ra	2.293
Nra	/J-Nra-5	Ra	66.098	Nra	28.904
Nra	/J-Nra-6	Nra	94.340	Ra	3.808
Nra	/J-Nra-7	Du	40.009	Ra	25.938
Nra	/J-Nra-8	Ra	86.242	Nra	8.009
Nra	/J-Nra-9	Ra	58.211	Du	27.854
Nra	/J-Nra-10	Nra	91.758	Ra	5.674
Nra	/J-Nra-11	Nra	49.928	Sra	33.105
Nra	/J-Nra-12	Nra	96.385	Sra	2.041
Nra	/J-Nra-13	Nra	97.148	Sra	1.547
Nra	/J-Nra-14	Nra	87.408	Ra	10.789
Nra	/R-Nra-1	Nra	99.205	Ra	0.752
Nra	/R-Nra-2	Ra	72.140	Nra	13.641
Nra	/R-Nra-3	Nra	96.045	Sra	3.061
Nra	/R-Nra-4	Ra	61.064	Nra	37.476
Nra	/R-Nra-5	Nra	93.141	Sra	3.840
Nra	/R-Nra-6	Sra	48.050	Nra	36.056
Nra	/R-Nra-7	Nra	91.386	Ra	8.391
Nra	/R-Nra-8	Nra	85.862	Sra	10.290
Nra	/R-Nra-9	Nra	91.758	Ra	5.674
Nra	/R-Nra-10	Nra	93.757	Sra	3.296
Nra	/R-Nra-11	Du	50.817	Hi	48.430
Nra	/R-Nra-12	Nra	94.127	Ra	4.340
Nra	/F-Nra-1	Nra	90.700	Du	5.040
Du	/Du-A	Ga	99.664	Du	0.236
Du	/Du-B	Hi	39.944	Ra	36.549

Lo	/R-LO-5	Lo	94.956	Cb	3.997
Lo	/J-LO-3	Mi	68.177	Sca	31.806
Lo	/F-LO-4	Lo	97.949	Du	1.238
Lo	/R-LO-6	Lo	85.191	Mi	8.194
Lo	/J-LO-4	Lo	53.263	Sca	37.527
Lo	/F-LO-5	Lo	92.920	Cb	5.652
Lo	/R-LO-7	Lo	94.525	Cb	5.084
Lo	/J-LO-5	Lo	99.371	Sca	0.456
Lo	/F-LO-6	Lo	99.330	Cb	0.614
Lo	/R-LO-8	Lo	97.314	Mi	2.586
Lo	/J-LO-6	Lo	51.321	Du	48.501
Lo	/R-LO-9	Lo	63.588	Cb	28.730
Lo	/J-LO-7	Lo	59.599	Du	18.461
Lo	/R-LO-10	Cb	95.903	Sca	2.973
Lo	/J-LO-8	Lo	95.984	Du	3.695
Lo	/R-LO-11	Lo	98.948	Cb	0.933
Lo	/J-LO-9	Mi	81.725	Sca	11.032
Lo	/R-LO-12	Lo	46.142	Ga	45.461
Lo	/J-LO-10	Hi	65.288	Lo	20.622
Lo	/R-LO-13	Lo	99.258	Sca	0.458
Lo	/R-LO-1	Lo	96.610	Cb	3.276
Wca	/WCA_1	Sca	74.851	Lo	15.939
Wca	/WCA_2	Wca	96.448	Sca	3.454
Wca	/WCA_3	Ga	49.858	Wca	14.637
Wca	/WCA_4	Wca	97.994	Sca	1.469
Wca	/WCA_5	Wca	99.685	Sca	0.289
Wca	/WCA_6	Wca	99.515	Mi	0.415
Wca	/WCA_7	Wca	97.580	Mi	2.054
Wca	/WCA_8	Wca	99.992	Sca	0.008
Wca	/WCA_9	Wca	99.926	Ga	0.071
Wca	/WCA_10	Wca	94.289	Ga	5.375
Wca	/WCA_11	Cb	92.892	Ga	4.737
Wca	/WCA_12	Wca	99.815	Ga	0.133
Wca	/WCA_13	Wca	94.892	Cb	3.836
Wca	/WCA_14	Sca	82.240	Mi	17.586
Wca	/WCA_15	Wca	96.694	Sca	2.645
Wca	/WCA_16	Ga	90.834	Cb	2.928
Wca	/WCA_17	Wca	98.679	Mi	0.703
Wca	/WCA_18	Wca	85.076	Sca	11.011
Wca	/WCA_19	Ga	71.915	Wca	20.588
Wca	/WCA_20	Wca	99.787	Ga	0.190
Wca	/WCA_21	Cb	48.557	Ga	32.923
Wca	/WCA_22	Wca	88.245	Cb	4.344
Wca	/WCA_23	Wca	80.601	Ga	18.414
Wca	/WCA_24	Wca	99.971	Sca	0.015
Wca	/WCA_25	Wca	99.787	Sca	0.126
Wca	/WCA_26	Wca	86.922	Ga	12.834
Wca	/WCA_27	Ga	67.766	Sca	29.522
Sca	/SCA_1	Sca	98.327	Mi	1.635
Sca	/SCA_2	Ga	78.094	Cb	18.586
Sca	/SCA_3	Lo	72.153	Cb	22.560
Sca	/SCA_4	Mi	83.401	Sca	13.092
Sca	/SCA_5	Sca	92.080	Wca	6.638
Sca	/SCA_6	Wca	36.354	Sca	34.330
Sca	/SCA_7	Sca	96.414	Cb	1.743
Sca	/SCA_8	Lo	74.093	Wca	15.234
Sca	/SCA_9	Mi	67.837	Sca	22.542
Sca	/SCA_10	Wca	61.214	Sca	21.595
Sca	/SCA_11	Sca	99.976	Lo	0.012
Sca	/SCA_12	Sca	72.882	Cb	23.037
Sca	/SCA_13	Mi	96.928	Wca	2.378
Sca	/SCA_14	Sca	85.850	Mi	13.273
Sca	/SCA_15	Sca	99.814	Wca	0.133
Sca	/SCA_16	Cb	76.603	Lo	19.133

Du	/Du-C	Hi	75.589	Du	24.397
Du	/Du-D	Du	99.967	Ga	0.016
Du	/Du-E	Ra	58.927	Hi	18.905
Du	/Du-F	Ra	58.927	Hi	18.905
Du	/Du-G	Du	84.274	Hi	15.005
Du	/Du-H	Du	99.070	Hi	0.671
Du	/Du-I	Du	97.636	Nra	2.031
Du	/Du-J	Du	94.809	Hi	5.179
Du	/Du-K	Du	99.913	Hi	0.066
Du	/Du-L	Du	99.966	Hi	0.018
Du	/Du-M	Du	99.481	Lo	0.373
Du	/Du-N	Du	99.956	Lo	0.038
Du	/Du-O	Du	100.000	Lo	0.000
Du	/Du-P	Du	85.649	Hi	14.341
Du	/Du-Q	Lo	84.538	Du	13.585
Du	/Du-R	Du	99.472	Sra	0.412
Du	/Du-S	Du	92.857	Hi	7.141
Du	/Du-T	Du	99.462	Lo	0.426
Du	/Du-U	Sra	86.612	Ra	7.615
Du	/Du-V	Ra	62.524	Sra	37.288
Du	/Du-R-2	Ra	70.276	Du	23.364
Du	/R-Du-3	Du	99.571	Hi	0.427
Du	/Du-4	Nra	30.150	Du	29.474
Du	/Du-5	Du	99.905	Hi	0.075
Du	/Du-6	Ra	55.423	Sra	44.419
Hi	/Hi-2	Hi	99.853	Ra	0.092
Hi	/Hi-3	Hi	99.932	Li	0.067
Hi	/Hi-4	Hi	99.787	Ra	0.184
Hi	/Hi-5	Hi	68.637	Ga	15.394
Hi	/Hi-6	Sra	50.622	Ra	48.101
Hi	/Hi-7	Hi	99.441	Li	0.558
Hi	/Hi-8	Ra	84.223	Sra	15.326
Hi	/Hi-9	Hi	99.700	Mi	0.270
Hi	/Hi-10	Hi	99.977	Sra	0.012
Hi	/Hi-11	Du	99.788	Hi	0.136
Hi	/Hi-12	Hi	94.085	Du	2.932
Hi	/Hi-1	Hi	96.010	Ra	3.864
Hi	/Hi_1	Hi	72.886	Du	24.881
Hi	/Hi_2	Ra	45.837	Hi	42.393
Hi	/Hi_4	Hi	94.571	Lo	5.216
Hi	/Hi_5	Du	41.368	Ra	34.118
Hi	/Hi_6	Hi	99.982	Ra	0.011
Hi	/Hi_7	Hi	97.680	Sra	1.567
Hi	/Hi_8	Hi	85.395	Lo	8.577
Hi	/Hi_9	Du	51.266	Hi	48.653
Hi	/Hi_10	Hi	99.958	Mi	0.016
Hi	/Hi_11	Li	93.836	Lli	6.164
Hi	/Hi_12	Li	93.869	Lli	5.864
Hi	/Hi_13	Ra	45.463	Du	22.058
Hi	/Hi_14	Hi	98.537	Sra	0.621
Hi	/Hi_15	Hi	100.000	Lo	0.000
Hi	/Hi_16	Hi	99.123	Ra	0.750
Hi	/Hi_17	Hi	97.864	Du	1.786
Hi	/Hi_18	Hi	99.872	Ra	0.081
Hi	/Hi_19	Hi	96.972	Du	1.285
Hi	/Hi_20	Hi	99.923	Mi	0.058
Lli	/Lli_1	Lli	99.960	Li	0.031
Lli	/Lli_2	Lli	56.834	Hi	38.107
Lli	/Lli_3	Lli	99.993	Lo	0.007
Lli	/Lli_4	Lli	91.088	Li	8.789
Lli	/Lli_5	Cb	71.406	Lo	25.248
Lli	/Lli_6	Lli	70.250	Li	29.743
Lli	/Lli_7	Li	96.462	Lli	3.536
Lli	/Lli_8	Lli	99.119	Li	0.881

Sca	/SCA_17	Sca	99.902	Ga	0.043
Sca	/SCA_18	Sca	67.743	Wca	16.266
Sca	/SCA_19	Sca	67.469	Mi	32.363
Sca	/SCA_20	Sca	55.115	Mi	34.862
Sca	/SCA_21	Sca	96.768	Mi	1.570
Sca	/SCA_22	Sca	99.955	Wca	0.021
Sca	/SCA_23	Nra	32.429	Lo	28.335
Sca	/SCA_24	Sca	95.804	Ga	1.370
Sca	/SCA_25	Sca	92.357	Mi	7.064
Sca	/SCA_26	Sca	99.526	Lo	0.350
Sca	/SCA_27	Sca	79.208	Wca	18.394
Sca	/SCA_28	Lo	44.806	Sca	43.451
Sca	/SCA_29	Lo	86.824	Cb	6.777
Sca	/SCA_30	Sca	99.698	Lo	0.203
Mi	/Mi-25	Mi	53.626	Sca	42.545
Mi	/Mi-13	Mi	64.202	Lo	31.485
Mi	/Mi-1	Mi	83.645	Sca	13.373
Mi	/Mi-26	Mi	73.545	Wca	25.350
Mi	/Mi-14	Mi	99.478	Lo	0.515
Mi	/Mi-2	Sca	74.089	Mi	23.026
Mi	/Mi-27	Mi	99.824	Sca	0.119
Mi	/Mi-15	Mi	99.991	Hi	0.007
Mi	/Mi-3	Mi	97.550	Sca	2.296
Mi	/Mi-28	Mi	99.999	Ga	0.001
Mi	/Mi-16	Mi	99.991	Sca	0.008
Mi	/Mi-4	Lo	58.896	Sca	35.887
Mi	/Mi-29	Mi	96.956	Sca	2.103
Mi	/Mi-17	Mi	99.418	Lo	0.397
Mi	/Mi-5	Sca	73.859	Ga	20.267
Mi	/Mi-18	Mi	91.359	Lo	5.426
Mi	/Mi-6	Mi	99.413	Hi	0.527
Mi	/F-Mi-1	Mi	98.420	Lo	1.531
Mi	/Mi-19	Mi	87.975	Wca	10.664
Mi	/Mi-7	Sca	80.595	Mi	13.050
Mi	/Mi-20	Mi	99.903	Ga	0.089
Mi	/Mi-8	Mi	99.965	Lo	0.028
Mi	/Mi-21	Mi	99.878	Sca	0.061
Mi	/Mi-9	Sca	77.798	Mi	21.898
Mi	/Mi-22	Mi	98.593	Lo	0.734
Mi	/Mi-10	Mi	98.316	Wca	0.789
Mi	/Mi-23	Mi	90.392	Lo	5.826
Mi	/Mi-11	Mi	99.810	Sca	0.179
Mi	/Mi-24	Mi	78.180	Sca	21.275
Mi	/Mi-12	Mi	98.684	Sca	1.287
Mi	/Mi_2_07	Lo	99.710	Sca	0.272
Mi	/Mi_6_-7	Sca	77.612	Mi	22.340
Ra	/RA_1	Sra	68.876	Ra	30.692
Ra	/RA_2	Sra	79.269	Ra	14.994
Ra	/RA_3	Du	55.702	Ra	30.268
Ra	/RA_4	Ra	61.498	Hi	34.277
Ra	/RA_5	Sra	94.481	Ra	4.168
Ra	/RA_6	Sra	74.828	Ra	21.784
Ra	/RA_7	Du	94.755	Lo	1.791
Ra	/RA_8	Du	54.541	Hi	17.868
Ra	/RA_9	Hi	79.968	Sra	12.480
Ra	/RA_10	Ra	82.806	Hi	13.840
Ra	/RA_11	Sra	56.791	Ra	27.117
Ra	/RA_12	Nra	97.525	Ra	2.128

Lli	/Lli_9	Li	86.833	Lli	13.167
Lli	/Lli_10	Li	80.739	Lli	19.261
Lli	/Lli_11	Li	94.862	Lli	5.135
Lli	/Lli_12	Lli	99.890	Lo	0.042
Lli	/Lli_13	Lli	94.140	Du	3.936
Lli	/Lli_14	Lli	58.277	Li	41.719
Lli	/Lli_15	Lli	55.304	Li	44.639
Lli	/Lli_16	Lli	66.106	Li	33.737
Lli	/Lli_17	Lli	99.962	Lo	0.024
Lli	/Lli_18	Lli	99.512	Li	0.453
Lli	/Lli_19	Li	79.829	Lli	20.170
Lli	/Lli_20	Lli	99.744	Hi	0.241
Lli	/Lli_21	Lli	53.831	Li	46.167
Lli	/Lli_22	Li	90.689	Lli	9.234
Lli	/Lli_23	Li	59.079	Lli	40.920
Lli	/Lli_24	Lli	99.836	Li	0.162
Lli	/Lli_25	Hi	85.963	Li	11.411
Lli	/Lli_26	Lli	98.161	Li	1.819
Lli	/Lli_27	Lli	99.952	Li	0.032
Lli	/Lli_28	Lli	99.970	Lo	0.017
Lli	/Lli_29	Lli	99.541	Li	0.320
Uli	/Uli_1	Li	91.939	Lli	8.060
Uli	/Uli_2	Li	89.837	Lli	10.163
Uli	/Uli_3	Li	66.308	Lli	31.519
Uli	/Uli_4	Lli	89.652	Li	10.338
Uli	/Uli_5	Li	80.835	Lli	10.759
Uli	/Uli_6	Li	73.829	Lli	24.853
Uli	/Uli_7	Li	83.409	Lli	16.439
Uli	/Uli_8	Li	97.977	Lli	2.002
Uli	/Uli_9	Li	92.201	Lli	7.776
Uli	/Uli_10	Li	87.485	Lli	12.515
Uli	/Uli_11	Li	96.463	Lli	3.474
Uli	/Uli_12	Lli	52.896	Li	47.104
Uli	/Uli_13	Li	93.626	Lli	5.229
Uli	/Uli_14	Lli	51.859	Li	47.995
Uli	/Uli_15	Li	98.787	Lli	1.156
Uli	/Uli_16	Li	82.544	Lli	17.242
Uli	/Uli_17	Li	76.755	Lli	23.240
Uli	/Uli_18	Lli	63.169	Li	36.452
Uli	/Uli_19	Lli	54.740	Li	45.259
Uli	/Uli_20	Li	83.668	Lli	16.332
Uli	/Uli_21	Li	83.848	Lli	15.991
Uli	/Uli_22	Li	90.076	Lli	9.923
Uli	/Uli_23	Li	90.324	Lli	9.672
Uli	/Uli_24	Li	80.329	Lli	17.291
Uli	/Uli_25	Li	67.379	Lli	32.621
Uli	/Uli_26	Li	97.645	Lli	2.324
Uli	/Uli_27	Lli	50.347	Li	49.651
Uli	/Uli_28	Lli	52.562	Li	47.418
Uli	/Li-2	Li	61.833	Lli	37.312
Uli	/Li-3	Hi	90.793	Sra	5.845
Uli	/Li-4	Li	51.370	Lli	48.630
Uli	/Li-5	Li	83.286	Lli	16.710
Uli	/Li-6	Li	87.309	Lli	12.622
Uli	/Li-7	Li	91.810	Lli	7.147
Uli	/Li-8	Li	87.351	Lli	12.649
Uli	/Li-9	Li	82.834	Lli	17.159
Uli	/Li-10	Li	97.602	Lli	2.361

Table 3-1: Summary of self-assignment tests for populations defined by stream-of-origin.

Origin	Assigned sample	rank 1	score %	rank 2	score %
Cbp	/Du-A	Cbp	71.952	Mip	27.626
Cbp	/Lli_5	Cbp	92.132	Llip	4.386
Cbp	/SCA_3	Cbp	98.185	Mip	1.800
Cbp	/SCA_8	Cbp	55.514	Wcap	38.422
Cbp	/SCA_28	Mip	34.536	Llip	28.673
Cbp	/Mi-4	Cbp	86.354	Mip	12.618
Cbp	/WCA_11	Cbp	92.862	Wcap	6.952
Cbp	/WCA_16	Cbp	58.912	Mip	40.539
Cbp	/CB_1	Cbp	97.789	Wcap	2.097
Cbp	/CB_2	Cbp	99.185	Wcap	0.766
Cbp	/CB_3	Cbp	99.974	Mip	0.015
Cbp	/CB_5	Cbp	99.725	Mip	0.270
Cbp	/CB_7	Cbp	56.446	Mip	43.206
Cbp	/CB_8	Wcap	60.308	Cbp	39.531
Cbp	/CB_10	Cbp	96.990	Wcap	2.858
Cbp	/CB_11	Cbp	99.999	Mip	0.000
Cbp	/CB_13	Cbp	97.981	Llip	1.256
Cbp	/CB_14	Cbp	100.000	Mip	0.000
Cbp	/CB_16	Cbp	99.984	Mip	0.013
Cbp	/CB_18	Cbp	93.445	Wcap	5.743
Cbp	/CB_19	Cbp	99.627	Mip	0.357
Cbp	/CB_20	Cbp	97.904	Mip	1.500
Cbp	/CB_21	Cbp	58.325	Wcap	39.472
Cbp	/CB_22	Cbp	99.139	Wcap	0.861
Cbp	/CB_23	Cbp	95.894	Wcap	4.105
Cbp	/CB_24	Cbp	56.850	Mip	42.839
Cbp	/CB_25	Cbp	99.771	Wcap	0.227
Cbp	/CB_26	Cbp	81.355	Mip	18.250
Cbp	/CB_29	Cbp	66.926	Wcap	31.267
Cbp	/CB_30	Cbp	96.919	Mip	3.077
Cbp	/CB_31	Cbp	99.907	Mip	0.092
Cbp	/J-GA-20	Cbp	99.644	Mip	0.283
Cbp	/J-GA-8	Cbp	99.833	Mip	0.139
Cbp	/R-GA-2	Cbp	99.780	Wcap	0.178
Cbp	/J-GA-10	Mip	86.411	Cbp	9.802
Cbp	/R-GA-3	Cbp	99.125	Mip	0.871
Cbp	/J-GA-11	Cbp	95.476	Wcap	3.370
Cbp	/R-GA-4	Cbp	97.924	Wcap	1.973
Cbp	/J-GA-12	Cbp	99.605	Wcap	0.367
Cbp	/J-GA-13	Mip	87.770	Wcap	8.247
Cbp	/J-GA-1	Cbp	99.998	Mip	0.001
Cbp	/J-GA-16	Cbp	99.993	Wcap	0.006
Cbp	/J-GA-17	Cbp	98.204	Wcap	1.585
Cbp	/J-GA-18	Cbp	99.648	Mip	0.240
Cbp	/J-GA-6	Cbp	99.812	Hip	0.134
Cbp	/R-LO-14	Cbp	99.958	Mip	0.042
Cbp	/F-LO-1	Mip	91.710	Wcap	5.885
Cbp	/J-LO-1	Cbp	96.501	Mip	3.283
Cbp	/R-LO-4	Cbp	99.888	Mip	0.107
Cbp	/F-LO-3	Cbp	94.894	Mip	5.106
Cbp	/R-LO-5	Mip	52.203	Cbp	41.536
Cbp	/F-LO-5	Cbp	99.752	Llip	0.204
Cbp	/R-LO-7	Cbp	88.879	Mip	10.961
Cbp	/R-LO-9	Cbp	99.130	Mip	0.869
Cbp	/R-LO-10	Cbp	99.018	Mip	0.977
Cbp	/J-LO-8	Cbp	52.383	Mip	47.589
Cbp	/F-LO-1	Cbp	62.909	Mip	30.347
Wcap	/SCA_2	Wcap	76.710	Cbp	12.534
Wcap	/SCA_16	Wcap	61.906	Cbp	37.901

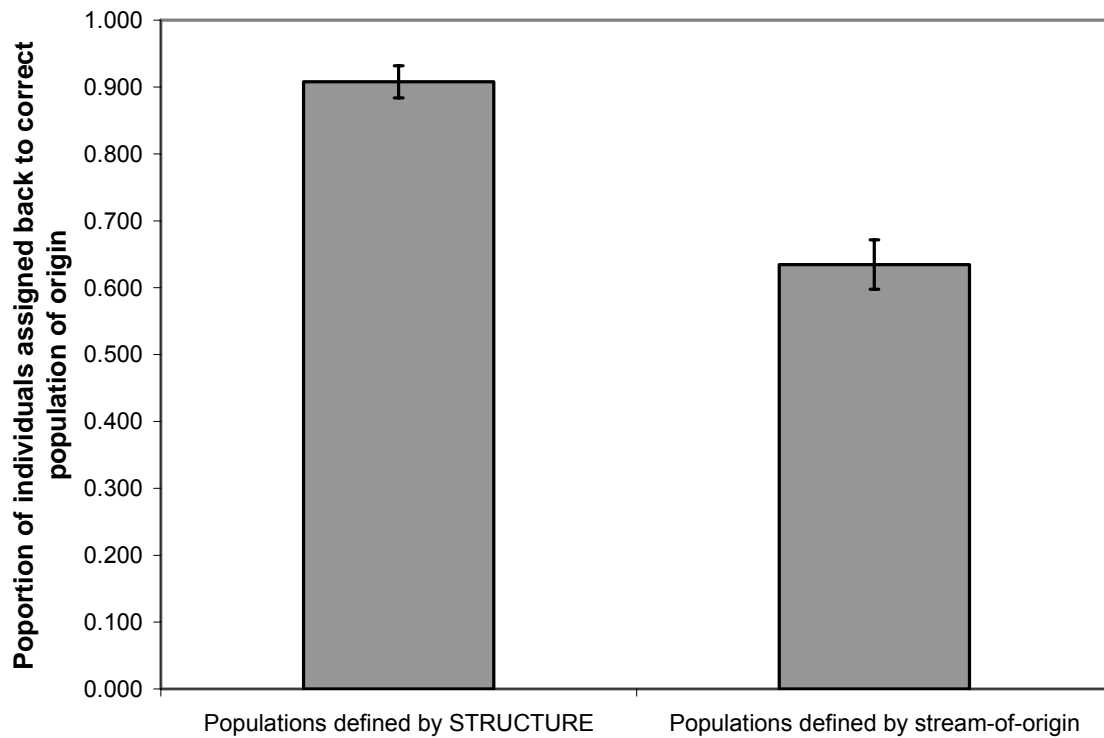| Origin	Assigned sample	rank 1	score %	rank 2	score %
Rap	/RA_2	Rap	99.997	Hip	0.003
Rap	/RA_5	Rap	63.512	Hip	36.220
Rap	/RA_11	Rap	99.981	Hip	0.019
Rap	/RA_12	Rap	72.387	Hip	27.613
Rap	/RA_13	Rap	98.470	Hip	1.530
Rap	/RA_14	Rap	89.614	Hip	10.289
Rap	/RA_17	Rap	99.170	Hip	0.830
Rap	/RA_18	Rap	98.038	Hip	1.961
Rap	/RA_19	Rap	99.999	Hip	0.001
Rap	/RA_20	Rap	99.944	Hip	0.056
Rap	/RA_21	Rap	98.549	Hip	1.451
Rap	/RA_22	Rap	99.847	Hip	0.153
Rap	/RA_23	Rap	99.611	Hip	0.389
Rap	/RA_24	Rap	99.866	Hip	0.134
Rap	/Sra-1	Rap	99.626	Hip	0.374
Rap	/Sra-3	Rap	99.962	Hip	0.038
Rap	/Sra-4	Rap	99.989	Hip	0.011
Rap	/Sra-5	Rap	99.140	Hip	0.860
Rap	/Sra-6	Rap	66.076	Hip	33.888
Rap	/Sra-9	Rap	99.984	Hip	0.016
Rap	/Sra-10	Rap	91.304	Hip	8.696
Rap	/Sra-11	Rap	88.692	Hip	11.308
Rap	/Sra-12	Rap	99.968	Hip	0.032
Rap	/Sra-13	Rap	99.614	Hip	0.386
Rap	/Sra-14	Rap	99.589	Hip	0.411
Rap	/Sra-16	Rap	99.808	Hip	0.192
Rap	/Sra-17	Rap	98.606	Hip	1.394
Rap	/Sra-18	Rap	98.618	Hip	1.382
Rap	/Sra-19	Rap	99.652	Hip	0.348
Rap	/J-Nra-1	Rap	99.986	Hip	0.014
Rap	/J-Nra-2	Rap	99.924	Hip	0.076
Rap	/J-Nra-3	Rap	82.486	Hip	17.513
Rap	/J-Nra-4	Rap	99.996	Hip	0.004
Rap	/J-Nra-5	Rap	93.560	Hip	6.436
Rap	/J-Nra-6	Rap	99.899	Hip	0.101
Rap	/J-Nra-8	Rap	98.295	Hip	1.705
Rap	/J-Nra-10	Rap	99.990	Hip	0.010
Rap	/J-Nra-11	Rap	99.997	Hip	0.003
Rap	/J-Nra-12	Rap	99.953	Hip	0.047
Rap	/J-Nra-13	Rap	99.874	Hip	0.126
Rap	/J-Nra-14	Rap	99.858	Hip	0.140
Rap	/R-Nra-1	Rap	99.782	Hip	0.215
Rap	/R-Nra-2	Rap	95.260	Hip	4.696
Rap	/R-Nra-3	Rap	99.815	Hip	0.185
Rap	/R-Nra-4	Rap	98.721	Hip	1.276
Rap	/R-Nra-5	Rap	99.751	Hip	0.248
Rap	/R-Nra-6	Rap	99.901	Hip	0.099
Rap	/R-Nra-7	Rap	99.750	Hip	0.249
Rap	/R-Nra-8	Rap	99.994	Hip	0.006
Rap	/R-Nra-9	Rap	99.990	Hip	0.010
Rap	/R-Nra-10	Rap	99.954	Hip	0.046
Rap	/R-Nra-12	Rap	99.996	Hip	0.004
Rap	/F-Nra-1	Rap	99.112	Hip	0.864
Hip	/J-LO-10	Hip	87.666	Mip	12.326
Hip	/SCA_23	Hip	88.894	Wcap	4.947
Hip	/Li-3	Hip	99.609	Rap	0.227
Hip	/RA_3	Hip	99.659	Mip	0.268
Hip	/RA_4	Hip	68.118	Rap	31.880
Hip	/RA_6	Hip	96.994	Rap	3.006

Wcap	/SCA_27	Wcap	38.487	Mip	31.575
Wcap	/CB_4	Cbp	53.263	Wcap	46.273
Wcap	/CB_6	Wcap	55.528	Cbp	43.862
Wcap	/CB_28	Wcap	78.001	Cbp	21.798
Wcap	/R-GA-5	Wcap	96.465	Cbp	3.447
Wcap	/J-GA-3	Cbp	68.117	Wcap	31.776
Wcap	/J-GA-5	Wcap	95.479	Cbp	4.232
Wcap	/J-GA-19	Cbp	63.648	Wcap	36.349
Wcap	/F-LO-2	Mip	55.774	Wcap	28.193
Wcap	/WCA_2	Wcap	99.558	Cbp	0.248
Wcap	/WCA_4	Wcap	99.627	Cbp	0.361
Wcap	/WCA_8	Wcap	100.000	Cbp	0.000
Wcap	/WCA_9	Wcap	99.994	Cbp	0.006
Wcap	/WCA_10	Wcap	69.420	Mip	18.357
Wcap	/WCA_12	Wcap	99.962	Mip	0.022
Wcap	/WCA_13	Wcap	99.173	Cbp	0.823
Wcap	/WCA_15	Wcap	99.878	Hip	0.100
Wcap	/WCA_18	Wcap	99.903	Cbp	0.096
Wcap	/WCA_20	Wcap	99.999	Cbp	0.001
Wcap	/WCA_21	Wcap	69.400	Cbp	30.561
Wcap	/WCA_22	Wcap	97.762	Cbp	2.083
Wcap	/WCA_23	Wcap	97.728	Cbp	2.223
Wcap	/WCA_24	Wcap	99.996	Cbp	0.004
Wcap	/WCA_25	Wcap	99.965	Hip	0.024
Wcap	/WCA_26	Wcap	66.090	Cbp	32.168
Wcap	/WCA_27	Wcap	63.211	Mip	21.383
Mip	/Du-D	Hip	55.261	Cbp	28.333
Mip	/Du-K	Mip	53.726	Hip	44.379
Mip	/Du-L	Hip	97.310	Mip	2.679
Mip	/Du-M	Llip	64.063	Mip	20.326
Mip	/Du-O	Hip	99.937	Mip	0.053
Mip	/Du-Q	Cbp	73.979	Mip	22.585
Mip	/Du-R	Hip	80.061	Mip	19.501
Mip	/Hi-9	Hip	92.452	Mip	7.539
Mip	/CB_9	Mip	38.953	Rap	33.758
Mip	/CB_12	Hip	51.396	Llip	26.654
Mip	/CB_17	Cbp	70.259	Mip	29.740
Mip	/R-GA-1	Mip	99.948	Cbp	0.052
Mip	/J-GA-9	Mip	84.192	Wcap	14.047
Mip	/J-GA-14	Mip	66.407	Wcap	33.313
Mip	/J-GA-15	Mip	98.062	Hip	0.757
Mip	/J-GA-7	Cbp	61.748	Mip	31.141
Mip	/J-LO-3	Mip	100.000	Cbp	0.000
Mip	/R-LO-2	Mip	99.368	Llip	0.489
Mip	/F-LO-4	Cbp	59.819	Mip	40.079
Mip	/R-LO-6	Mip	98.819	Cbp	1.181
Mip	/J-LO-4	Mip	98.556	Cbp	1.418
Mip	/J-LO-5	Mip	89.902	Hip	9.365
Mip	/R-LO-8	Mip	99.707	Llip	0.145
Mip	/J-LO-6	Mip	83.356	Cbp	16.578
Mip	/J-LO-9	Mip	96.174	Cbp	3.462
Mip	/R-LO-12	Mip	83.548	Hip	9.312
Mip	/R-LO-13	Mip	91.117	Cbp	8.882
Mip	/WCA_1	Mip	93.819	Cbp	6.148
Mip	/WCA_3	Wcap	61.776	Cbp	22.975
Mip	/WCA_5	Mip	99.893	Wcap	0.105
Mip	/WCA_6	Mip	99.950	Cbp	0.037
Mip	/WCA_7	Mip	99.100	Wcap	0.883
Mip	/WCA_14	Mip	99.967	Wcap	0.030
Mip	/WCA_17	Mip	99.996	Cbp	0.003
Mip	/WCA_19	Mip	83.920	Wcap	15.594
Mip	/SCA_1	Mip	100.000	Cbp	0.000

Hip	/RA_7	Hip	93.150	Llip	2.796
Hip	/RA_8	Hip	98.594	Rap	0.806
Hip	/RA_9	Hip	72.674	Rap	15.126
Hip	/RA_10	Hip	77.252	Rap	22.736
Hip	/RA_15	Hip	99.596	Mip	0.387
Hip	/RA_16	Hip	98.699	Mip	1.237
Hip	/Sra-2	Hip	96.675	Mip	3.322
Hip	/Sra-7	Hip	99.747	Rap	0.252
Hip	/Sra-8	Rap	80.559	Hip	19.434
Hip	/J-Nra-7	Hip	99.899	Rap	0.085
Hip	/J-Nra-9	Hip	90.814	Rap	9.166
Hip	/R-Nra-11	Hip	99.985	Rap	0.009
Hip	/Du-B	Hip	94.763	Rap	3.300
Hip	/Du-C	Hip	100.000	Mip	0.000
Hip	/Du-E	Hip	95.088	Rap	4.907
Hip	/Du-F	Hip	95.088	Rap	4.907
Hip	/Du-G	Hip	99.999	Mip	0.001
Hip	/Du-H	Hip	99.630	Mip	0.312
Hip	/Du-I	Hip	98.839	Mip	0.767
Hip	/Du-J	Hip	94.469	Mip	5.526
Hip	/Du-P	Hip	98.199	Mip	1.800
Hip	/Du-S	Hip	99.988	Mip	0.012
Hip	/Du-T	Hip	99.320	Mip	0.437
Hip	/Du-U	Hip	95.398	Wcap	1.592
Hip	/Du-R-2	Hip	99.749	Rap	0.219
Hip	/R-Du-3	Hip	99.988	Mip	0.012
Hip	/Du-4	Hip	85.752	Rap	11.692
Hip	/Du-5	Hip	99.742	Mip	0.190
Hip	/Hi-2	Hip	98.755	Mip	1.211
Hip	/Hi-4	Hip	99.958	Mip	0.031
Hip	/Hi-5	Hip	99.637	Mip	0.337
Hip	/Hi-6	Hip	61.358	Rap	38.641
Hip	/Hi-10	Hip	99.977	Mip	0.020
Hip	/Hi-11	Hip	99.998	Mip	0.001
Hip	/Hi-12	Hip	99.954	Mip	0.046
Hip	/Hi-1	Hip	94.983	Rap	2.580
Hip	/Hi_1	Hip	99.992	Mip	0.008
Hip	/Hi_5	Hip	97.284	Rap	2.560
Hip	/Hi_6	Hip	99.982	Mip	0.017
Hip	/Hi_7	Hip	98.401	Mip	1.417
Hip	/Hi_9	Hip	99.934	Mip	0.066
Hip	/Hi_13	Hip	99.455	Rap	0.542
Hip	/Hi_14	Hip	99.793	Wcap	0.070
Hip	/Hi_15	Hip	99.944	Mip	0.056
Hip	/Hi_16	Hip	99.779	Rap	0.204
Hip	/Hi_17	Hip	99.699	Mip	0.197
Hip	/Hi_18	Hip	99.989	Mip	0.010
Hip	/Hi_20	Hip	99.746	Mip	0.254
Llip	/J-LO-2	Ulip	75.535	Llip	24.408
Llip	/F-LO-6	Llip	75.972	Cbp	22.119
Llip	/J-LO-7	Llip	77.303	Cbp	11.205
Llip	/Hi_4	Llip	99.906	Hip	0.046
Llip	/Hi_8	Llip	99.788	Ulip	0.152
Llip	/Hi_12	Llip	87.594	Ulip	12.406
Llip	/Hi_19	Hip	71.534	Llip	26.664
Llip	/Uli_3	Llip	98.362	Ulip	1.636
Llip	/Uli_4	Llip	91.168	Ulip	8.832
Llip	/Uli_5	Llip	97.639	Ulip	2.348
Llip	/Uli_6	Llip	52.015	Ulip	47.964
Llip	/Uli_7	Ulip	84.326	Llip	15.674
Llip	/Uli_13	Llip	61.791	Ulip	38.208
Llip	/Uli_28	Llip	80.913	Ulip	19.087

Mip	/SCA_4	Mip	94.944	Hip	4.984	Llip	/Li-7	Llip	58.945	Ulip	41.046
Mip	/SCA_5	Mip	99.996	Cbp	0.002	Llip	/Lli_1	Llip	96.183	Ulip	3.817
Mip	/SCA_6	Mip	98.876	Hip	1.062	Llip	/Lli_2	Llip	99.984	Ulip	0.011
Mip	/SCA_7	Wcap	95.245	Mip	3.019	Llip	/Lli_3	Llip	98.426	Ulip	1.572
Mip	/SCA_9	Mip	99.758	Hip	0.144	Llip	/Lli_12	Llip	99.820	Ulip	0.178
Mip	/SCA_10	Mip	100.000	Cbp	0.000	Llip	/Lli_13	Llip	99.171	Cbp	0.638
Mip	/SCA_11	Mip	99.965	Cbp	0.035	Llip	/Lli_15	Llip	97.333	Ulip	2.666
Mip	/SCA_12	Mip	96.657	Cbp	3.343	Llip	/Lli_16	Llip	97.303	Ulip	2.689
Mip	/SCA_13	Mip	99.731	Hip	0.269	Llip	/Lli_17	Llip	99.999	Ulip	0.000
Mip	/SCA_14	Mip	98.646	Hip	1.353	Llip	/Lli_18	Llip	99.518	Ulip	0.482
Mip	/SCA_15	Mip	99.999	Cbp	0.001	Llip	/Lli_20	Llip	99.995	Hip	0.005
Mip	/SCA_17	Mip	86.874	Cbp	12.985	Llip	/Lli_24	Llip	99.354	Ulip	0.646
Mip	/SCA_18	Mip	60.715	Wcap	33.642	Llip	/Lli_25	Ulip	71.536	Llip	11.586
Mip	/SCA_19	Mip	100.000	Cbp	0.000	Llip	/Lli_26	Ulip	50.666	Llip	49.330
Mip	/SCA_20	Mip	99.240	Hip	0.749	Llip	/Lli_27	Llip	99.977	Ulip	0.019
Mip	/SCA_21	Mip	100.000	Wcap	0.000	Llip	/Lli_28	Llip	99.997	Ulip	0.002
Mip	/SCA_22	Mip	99.701	Wcap	0.210	Llip	/Lli_29	Llip	99.880	Ulip	0.118
Mip	/SCA_24	Mip	99.971	Cbp	0.014	Ulip	/Hi_11	Ulip	99.978	Llip	0.022
Mip	/SCA_25	Mip	99.768	Hip	0.231	Ulip	/Lli_4	Llip	83.758	Ulip	16.234
Mip	/SCA_26	Mip	99.054	Cbp	0.894	Ulip	/Lli_6	Ulip	92.063	Llip	7.937
Mip	/SCA_29	Mip	81.234	Cbp	18.765	Ulip	/Lli_7	Ulip	98.268	Llip	1.732
Mip	/SCA_30	Mip	96.788	Cbp	2.496	Ulip	/Lli_8	Llip	63.788	Ulip	36.212
Mip	/Mi-25	Mip	99.878	Hip	0.117	Ulip	/Lli_9	Ulip	99.811	Llip	0.189
Mip	/Mi-13	Mip	99.657	Hip	0.339	Ulip	/Lli_10	Ulip	99.953	Llip	0.047
Mip	/Mi-1	Mip	99.797	Cbp	0.203	Ulip	/Lli_11	Ulip	99.878	Llip	0.122
Mip	/Mi-26	Mip	99.980	Wcap	0.015	Ulip	/Lli_14	Llip	57.861	Ulip	42.139
Mip	/Mi-14	Mip	99.930	Hip	0.069	Ulip	/Lli_19	Ulip	99.486	Llip	0.514
Mip	/Mi-2	Mip	99.639	Cbp	0.361	Ulip	/Lli_21	Ulip	87.528	Llip	12.472
Mip	/Mi-27	Mip	99.998	Hip	0.002	Ulip	/Lli_22	Ulip	99.394	Llip	0.605
Mip	/Mi-15	Mip	99.999	Hip	0.001	Ulip	/Lli_23	Ulip	99.964	Llip	0.036
Mip	/Mi-3	Mip	100.000	Cbp	0.000	Ulip	/Uli_1	Ulip	95.206	Llip	4.794
Mip	/Mi-28	Mip	99.984	Hip	0.016	Ulip	/Uli_2	Ulip	99.961	Llip	0.039
Mip	/Mi-16	Mip	100.000	Hip	0.000	Ulip	/Uli_8	Ulip	97.619	Llip	2.381
Mip	/Mi-29	Mip	99.998	Hip	0.002	Ulip	/Uli_9	Ulip	89.277	Llip	10.723
Mip	/Mi-17	Mip	100.000	Hip	0.000	Ulip	/Uli_10	Ulip	99.972	Llip	0.028
Mip	/Mi-5	Mip	99.794	Wcap	0.122	Ulip	/Uli_11	Ulip	99.444	Llip	0.556
Mip	/Mi-18	Mip	97.594	Hip	2.405	Ulip	/Uli_12	Ulip	93.563	Llip	6.437
Mip	/Mi-6	Mip	99.663	Hip	0.337	Ulip	/Uli_14	Llip	54.370	Ulip	45.629
Mip	/F-Mi-1	Mip	100.000	Cbp	0.000	Ulip	/Uli_15	Ulip	92.833	Llip	7.167
Mip	/Mi-19	Mip	100.000	Hip	0.000	Ulip	/Uli_16	Ulip	99.084	Llip	0.914
Mip	/Mi-7	Mip	95.424	Wcap	3.787	Ulip	/Uli_17	Ulip	71.657	Llip	28.343
Mip	/Mi-20	Mip	99.996	Cbp	0.004	Ulip	/Uli_18	Ulip	64.548	Llip	35.447
Mip	/Mi-8	Mip	100.000	Cbp	0.000	Ulip	/Uli_19	Ulip	95.455	Llip	4.545
Mip	/Mi-21	Mip	100.000	Cbp	0.000	Ulip	/Uli_20	Ulip	98.512	Llip	1.488
Mip	/Mi-9	Mip	99.928	Hip	0.072	Ulip	/Uli_21	Ulip	94.753	Llip	5.247
Mip	/Mi-22	Mip	98.064	Hip	1.936	Ulip	/Uli_22	Ulip	97.747	Llip	2.253
Mip	/Mi-10	Mip	100.000	Cbp	0.000	Ulip	/Uli_23	Ulip	99.984	Llip	0.016
Mip	/Mi-23	Mip	99.999	Hip	0.001	Ulip	/Uli_24	Ulip	82.368	Llip	17.632
Mip	/Mi-11	Mip	100.000	Llip	0.000	Ulip	/Uli_25	Ulip	99.979	Llip	0.021
Mip	/Mi-24	Mip	99.966	Hip	0.034	Ulip	/Uli_26	Ulip	99.866	Llip	0.134
Mip	/Mi-12	Mip	99.898	Hip	0.102	Ulip	/Uli_27	Ulip	87.768	Llip	12.232
Mip	/Mi_2_07	Mip	99.992	Cbp	0.008	Ulip	/Li-2	Ulip	63.313	Llip	36.685
Mip	/Mi_6_-7	Mip	100.000	Cbp	0.000	Ulip	/Li-4	Ulip	96.708	Llip	3.292
Rap	/Du-V	Rap	99.532	Hip	0.467	Ulip	/Li-5	Ulip	99.080	Llip	0.920
Rap	/Du-6	Rap	99.252	Hip	0.748	Ulip	/Li-6	Ulip	92.926	Llip	7.074
Rap	/Hi-8	Rap	98.180	Hip	1.820	Ulip	/Li-8	Ulip	99.963	Llip	0.037
Rap	/Hi_2	Rap	52.635	Hip	47.365	Ulip	/Li-9	Ulip	79.408	Llip	20.592
Rap	/RA_1	Rap	91.205	Hip	8.794	Ulip	/Li-10	Ulip	99.586	Llip	0.414

Table 3-2: Summary of self-assignment tests for populations defined by STRUCTURE.

a)

b)

Paired Samples Test

| | | Paired Differences | | | | | | | |
| | | | | | 95% Confidence Interval of the Difference | | | | |
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Proportion_self_ass igned_STRpops - Proportion_self_ass igned_sites	.24643	.14734	.05569	.11016	.38270	4.425	6	.004

Figure 3-2) a) Total proportions of individuals successfully self-assigned to their population, and b) statistical analysis of comparison. Difference was found to be significant at $p < 0.01$

a)

b)
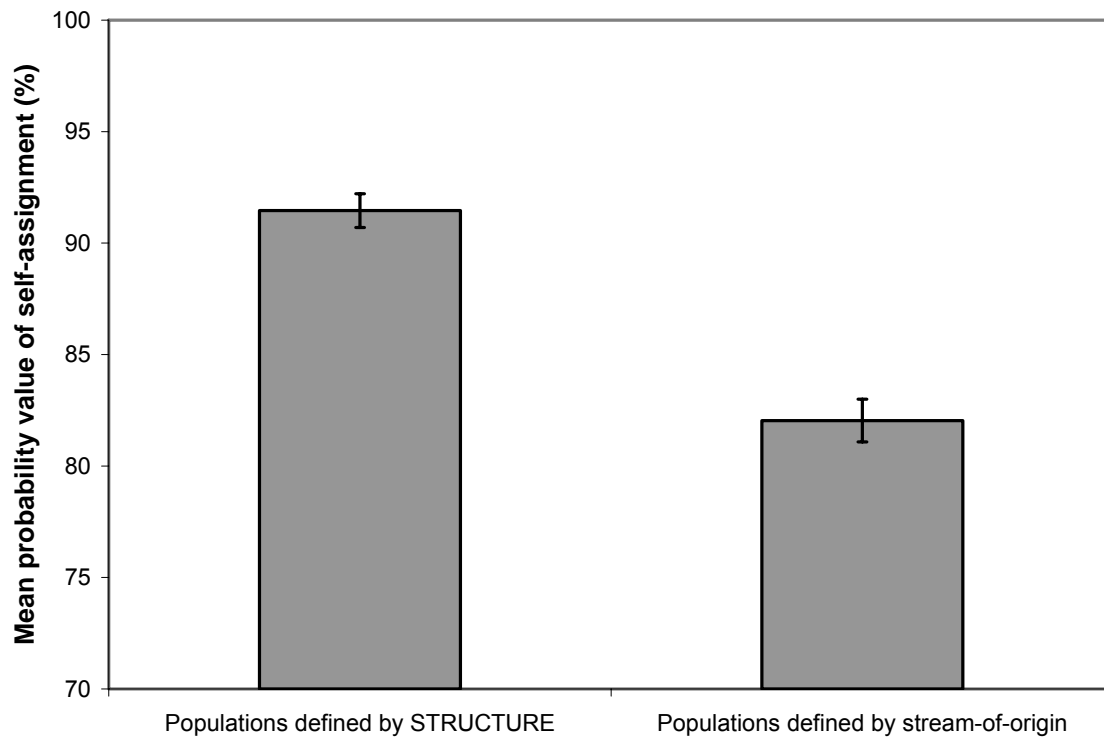
Ranks

	Population_definition	N	Mean Rank	Sum of Ranks
Prob_assignment	STRpops	358	435.15	155782.00
	Sites	364	289.07	105221.00
	Total	722		

Test Statistics[a]

	Prob_assignment
Mann-Whitney U	38791.000
Wilcoxon W	105221.000
Z	-9.409
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Population_definition

Figure 3-3: a) Mean probability values of self-assignment tests are higher when populations are defined by STRUCTURE analysis rather than stream-of-origin.b) Differences were found to be statistically significant by a Mann-Whitney U-test.

| Origin	Assigned sample	rank 1	score %	rank 2	score %
CA	/F-CA-6	Mi	57.769	Sca	27.459
CA	/F-CA-1	Mi	88.282	Sca	11.700
CA	/F-CA-2	Sca	43.118	Mi	25.244
CA	/F-CA-3	Sca	61.175	Mi	35.350
CA	/F-CA-4	Sca	99.993	Mi	0.006
CA	/F-CA-5	Sca	81.728	Mi	18.258
CA	/CA_1	Lo	53.967	Cb	43.564
CA	/CA_2	Sca	87.670	Mi	10.285
CA	/CA_3	Cb	61.558	Lo	34.477
CA	/CA_4	Ga	99.770	Cb	0.229
CA	/CA_5	Mi	93.776	Sca	5.506
CA	/CA_6	Ga	99.957	Cb	0.041
CA	/CA_7	Wca	77.019	Ga	17.538
CA	/CA_8	Cb	77.099	Ga	17.548
CA	/CA_9	Sca	99.096	Mi	0.802
CA	/CA_10	Sca	99.830	Mi	0.125
CA	/CA_11	Ga	91.749	Cb	7.843
CA	/CA_12	Mi	89.207	Hi	6.928
CA	/CA_13	Cb	71.008	Lo	19.463
CA	/CA_14	Cb	99.883	Ga	0.115
CA	/CA_15	Sca	65.971	Ga	21.851
CA	/CA_16	Mi	42.077	Sca	30.185
CA	/CA_17	Mi	49.888	Sca	48.849
CA	/CA_18	Mi	87.698	Sca	11.897
CA	/CA_19	Sca	99.700	Mi	0.132
CA	/CA_20	Sca	51.367	Lo	31.310
CA	/CA_21	Mi	72.370	Wca	26.715
CA	/CA_22	Mi	77.132	Lo	17.076
CA	/CA_23	Ga	85.039	Cb	14.026
ORR	/Res-1	Sca	67.176	Mi	28.694
ORR	/Res-2	Cb	77.124	Ga	12.431
ORR	/Res-3	Ga	99.614	Cb	0.311
ORR	/Res-4	Wca	78.142	Sca	17.722
ORR	/Res-5	Sca	69.747	Lo	21.998
ORR	/Res-6	Lo	73.219	Sca	22.257
ORR	/Res-7	Mi	99.148	Sca	0.359
ORR	/Res-8	Ga	83.217	Lo	6.504
ORR	/Res-9	Lo	56.948	Cb	25.581
ORR	/Res-10	Ga	99.466	Cb	0.533
ORR	/Res-11	Sca	95.003	Mi	4.614
ORR	/Res-12	Sca	98.169	Lo	0.760
ORR	/Res-13	Sca	98.909	Lo	1.061
ORR	/Res-14	Mi	58.495	Sca	41.475
ORR	/Res-15	Wca	97.391	Sca	0.956
ORR	/Res-16	Sca	62.090	Lo	21.937
ORR	/Res-17	Sca	77.705	Mi	9.097
OMR	/OMR-1	Ra	37.573	Hi	23.729
OMR	/OMR-2	Lo	61.092	Hi	17.904
OMR	/OMR-3	Hi	96.221	Du	3.358
OMR	/OMR-4	Hi	99.605	Lo	0.385
OMR	/OMR-5	Ga	60.648	Mi	36.664
OMR	/OMR-6	Du	95.689	Hi	2.582
OMR	/OMR-8	Du	99.622	Hi	0.376
OMR	/OMR-9	Lo	80.367	Uli	10.917
OMR	/OMR-10	Ga	86.257	Cb	13.622
OMR	/OMR-11	Ra	76.401	Sra	21.707
OMR	/OMR-12	Hi	99.987	Uli	0.007
OMR	/OMR-14	Sca	99.968	Ga	0.012
OMR	/OMR-15	Sra	80.336	Ra	15.785
OMR	/OMR-16	Hi	71.812	Du	14.728
OMR	/OMR-17	Ra	73.007	Du	25.360
OMR	/CM-1	Du	82.329	Wca	17.412
OMR	/CM-2	Ra	54.363	Sra	44.233
OMR	/CM-3	Mi	96.739	Sca	2.851
OMR	/CM-4	Ra	61.802	Du	25.944
OMR	/CM-5	Ra	76.681	Sra	20.135
OMR	/CM-6	Du	85.803	Lo	9.695
OMR	/CM-7	Hi	85.373	Lo	9.844
OMR	/CM-8	Du	56.307	Hi	21.704
OMR	/CM-9	Du	92.791	Sra	3.669
OMR	/CM-10	Hi	94.785	Du	4.839
OMR	/CM-11	Nra	96.814	Ra	1.680
OMR	/OMR_1	Hi	92.159	Sca	4.428
OMR	/OMR_2	Hi	71.544	Sra	16.412
OMR	/OMR_3	Sca	61.104	Ga	20.269
OMR	/OMR_4	Sca	73.055	Wca	26.086
TW	/TW_1	Mi	78.620	Sca	17.952
TW	/TW_2	Sca	41.450	Du	22.393
TW	/TW_3	Sca	95.924	Wca	1.573
TW	/TW_4	Cb	86.986	Lo	9.294
TW	/TW_5	Mi	58.117	Cb	25.545
TW	/TW_6	Ga	91.591	Cb	5.394
TW	/TW_7	Mi	99.795	Sca	0.186
TW	/TW_8	Cb	76.273	Ga	23.357
TW	/TU	Wca	88.196	Sca	11.274

Table 3-3: Summary of assignment tests of mixed migrant groups with reference populations defined by stream-of-origin

Origin	Assigned sample	rank 1	score %	rank 2	score %
CA	/F-CA-6	Mip	99.919	Cbp	0.080
CA	/F-CA-1	Mip	99.992	Hip	0.008
CA	/F-CA-2	Mip	99.988	Cbp	0.012
CA	/F-CA-3	Mip	99.953	Cbp	0.047
CA	/F-CA-4	Mip	99.998	Llip	0.001
CA	/F-CA-5	Mip	99.905	Hip	0.095
CA	/CA_1	Cbp	79.463	Llip	20.352
CA	/CA_2	Mip	99.422	Cbp	0.575
CA	/CA_3	Cbp	96.766	Mip	3.207
CA	/CA_4	Cbp	95.628	Wcap	4.368
CA	/CA_5	Mip	100.000	Hip	0.000
CA	/CA_6	Cbp	99.271	Wcap	0.520
CA	/CA_7	Mip	50.517	Wcap	47.773
CA	/CA_8	Cbp	97.212	Mip	2.467
CA	/CA_9	Mip	99.989	Cbp	0.011
CA	/CA_10	Mip	99.999	Cbp	0.001
CA	/CA_11	Cbp	96.515	Wcap	3.478
CA	/CA_12	Mip	90.407	Hip	9.108
CA	/CA_13	Cbp	99.993	Mip	0.006
CA	/CA_14	Cbp	100.000	Wcap	0.000
CA	/CA_15	Mip	59.453	Hip	32.609
CA	/CA_16	Mip	65.641	Hip	34.294
CA	/CA_17	Mip	99.989	Hip	0.011
CA	/CA_18	Mip	99.980	Hip	0.015
CA	/CA_19	Mip	99.985	Hip	0.009
CA	/CA_20	Mip	81.812	Wcap	12.137
CA	/CA_21	Mip	94.038	Wcap	5.544
CA	/CA_22	Mip	99.753	Cbp	0.150
CA	/CA_23	Cbp	99.674	Wcap	0.318
ORR	/Res-1	Mip	99.965	Cbp	0.017
ORR	/Res-2	Cbp	87.780	Wcap	12.055
ORR	/Res-3	Wcap	96.242	Cbp	3.757
ORR	/Res-4	Mip	91.870	Wcap	7.383
ORR	/Res-5	Mip	99.999	Hip	0.001
ORR	/Res-6	Mip	87.873	Cbp	11.560
ORR	/Res-7	Mip	99.520	Cbp	0.238
ORR	/Res-8	Cbp	98.567	Wcap	0.670
ORR	/Res-9	Cbp	99.097	Mip	0.610
ORR	/Res-10	Cbp	99.208	Wcap	0.770
ORR	/Res-11	Mip	99.961	Hip	0.039
ORR	/Res-12	Mip	98.527	Cbp	1.231
ORR	/Res-13	Mip	99.914	Cbp	0.086
ORR	/Res-14	Mip	100.000	Hip	0.000

Origin	Assigned sample	rank 1	score %	rank 2	score %
ORR	/Res-15	Mip	54.191	Cbp	26.861
ORR	/Res-16	Cbp	89.806	Wcap	9.686
ORR	/Res-17	Mip	96.666	Hip	3.330
OMR	/OMR-1	Hip	78.471	Rap	12.857
OMR	/OMR-2	Cbp	51.584	Llip	31.324
OMR	/OMR-3	Hip	99.997	Rap	0.002
OMR	/OMR-4	Llip	79.070	Hip	14.967
OMR	/OMR-5	Mip	99.586	Hip	0.412
OMR	/OMR-6	Hip	99.044	Mip	0.955
OMR	/OMR-8	Hip	100.000	Mip	0.000
OMR	/OMR-9	Ulip	39.996	Llip	37.056
OMR	/OMR-10	Cbp	99.980	Wcap	0.016
OMR	/OMR-11	Hip	97.550	Rap	2.109
OMR	/OMR-12	Hip	99.781	Mip	0.106
OMR	/OMR-14	Mip	99.149	Wcap	0.388
OMR	/OMR-15	Rap	88.262	Hip	11.674
OMR	/OMR-16	Hip	93.907	Llip	5.902
OMR	/OMR-17	Hip	98.714	Cbp	0.792
OMR	/CM-1	Wcap	95.041	Cbp	2.873
OMR	/CM-2	Rap	89.799	Hip	10.201
OMR	/CM-3	Mip	99.975	Hip	0.023
OMR	/CM-4	Hip	97.389	Mip	2.584
OMR	/CM-5	Rap	99.864	Hip	0.136
OMR	/CM-6	Hip	98.890	Mip	0.790
OMR	/CM-7	Hip	99.703	Mip	0.282
OMR	/CM-8	Hip	99.427	Mip	0.537
OMR	/CM-9	Hip	98.170	Rap	1.691
OMR	/CM-10	Hip	83.959	Cbp	9.131
OMR	/CM-11	Rap	99.977	Hip	0.023
OMR	/OMR_1	Wcap	73.049	Hip	21.822
OMR	/OMR_2	Hip	95.551	Mip	3.850
OMR	/OMR_3	Mip	70.425	Cbp	27.772
OMR	/OMR_4	Wcap	91.237	Mip	7.382
TW	/TW_1	Mip	99.484	Cbp	0.514
TW	/TW_2	Mip	52.443	Hip	47.499
TW	/TW_3	Mip	99.999	Wcap	0.001
TW	/TW_4	Cbp	93.771	Mip	5.802
TW	/TW_5	Mip	91.070	Cbp	8.877
TW	/TW_6	Cbp	95.210	Wcap	3.574
TW	/TW_7	Mip	100.000	Cbp	0.000
TW	/TW_8	Cbp	99.999	Mip	0.001
TW	/TU	Mip	63.279	Wcap	31.618

Table 3-4: Summary of assignment tests of mixed migrant groups with reference populations defined by STRUCTURE

a)

b)
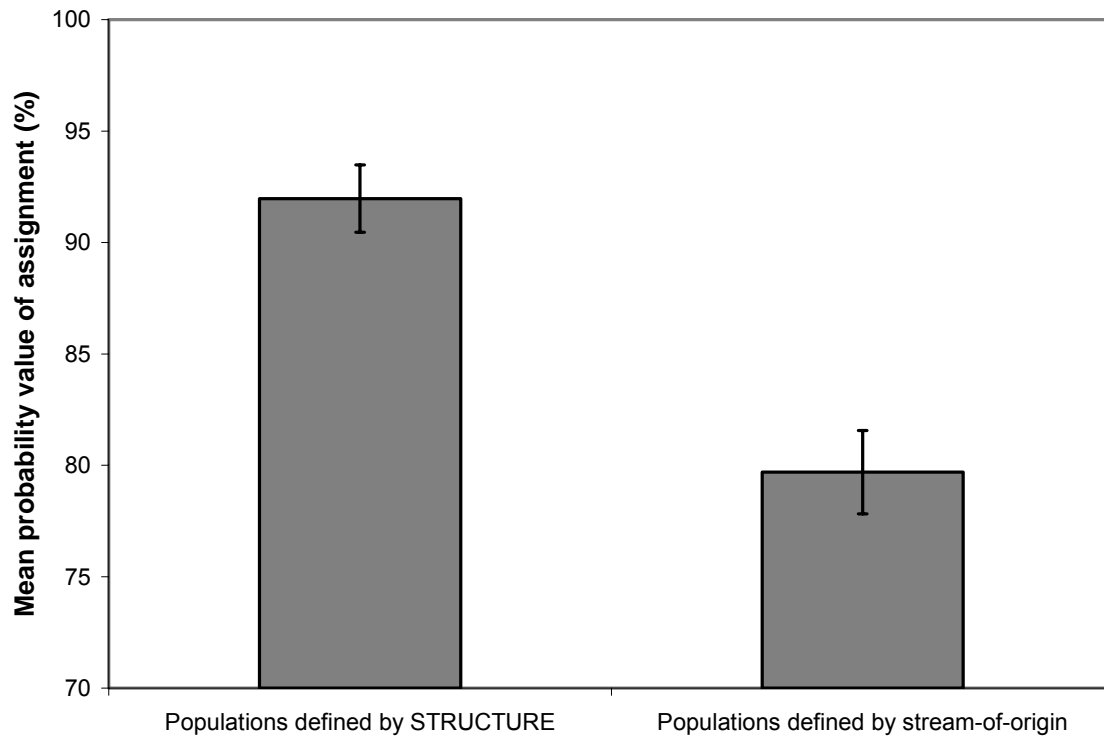
Ranks

	Population_definition	N	Mean Rank	Sum of Ranks
Prob_assignment	STRpops	85	106.84	9081.50
	Sites	85	64.16	5453.50
	Total	170		

Test Statistics[a]

	Prob_assignment
Mann-Whitney U	1798.500
Wilcoxon W	5453.500
Z	-5.654
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Population_definition

Figure 3-4: a) Mean probability values of mixed-migrant assignment tests are higher when populations are defined by STRUCTURE analysis rather than stream-of-origin.b) Differences were found to be statistically significant by a Mann-Whitney U-test.
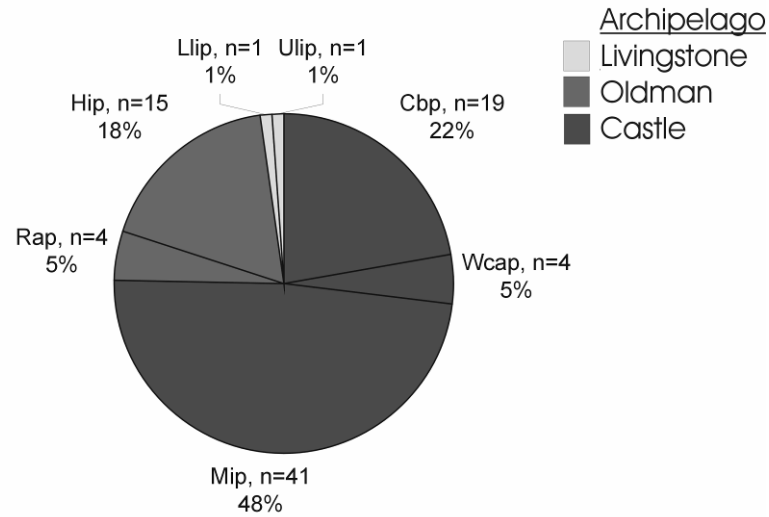
Figure 3-5: Stock composition of the total mixed-migrant pool when populations are defined by hierarchical STRUCTURE analysis.
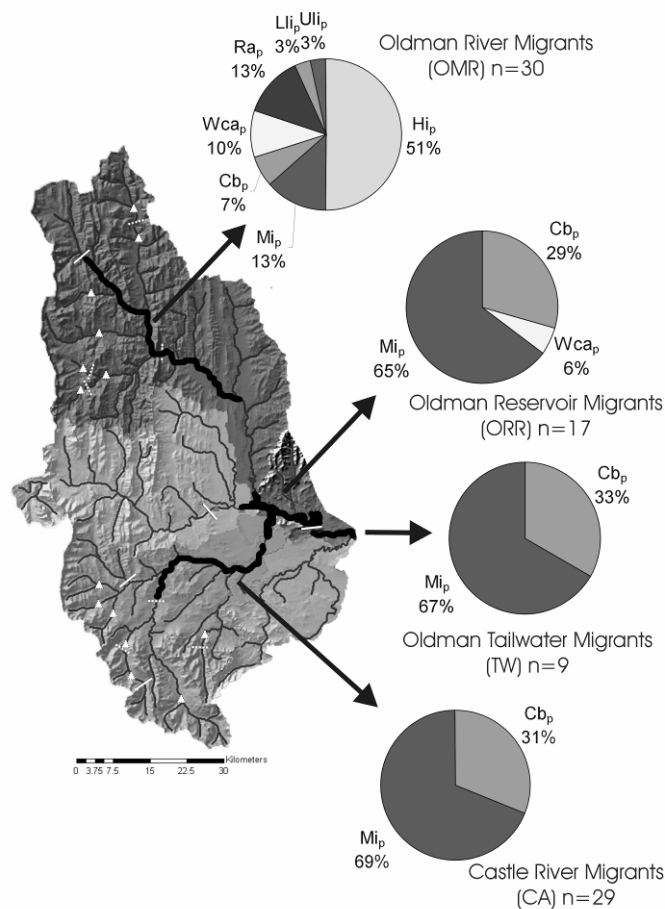
Figure 3-6: Stock composition of mixed-migrants from each of 4 sampling locations.

a)
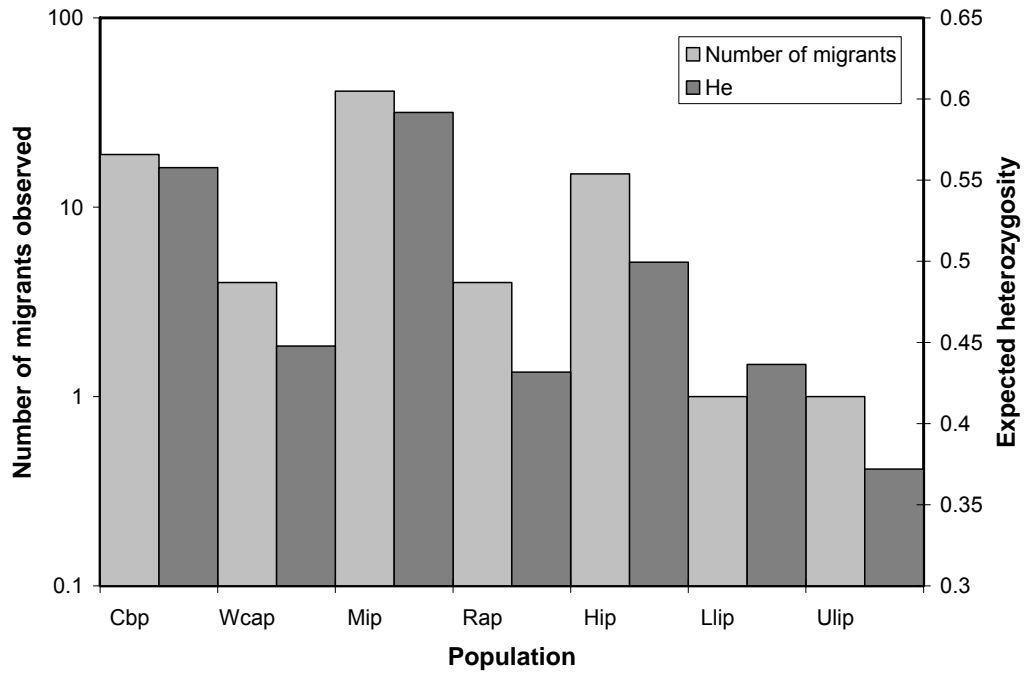
b)
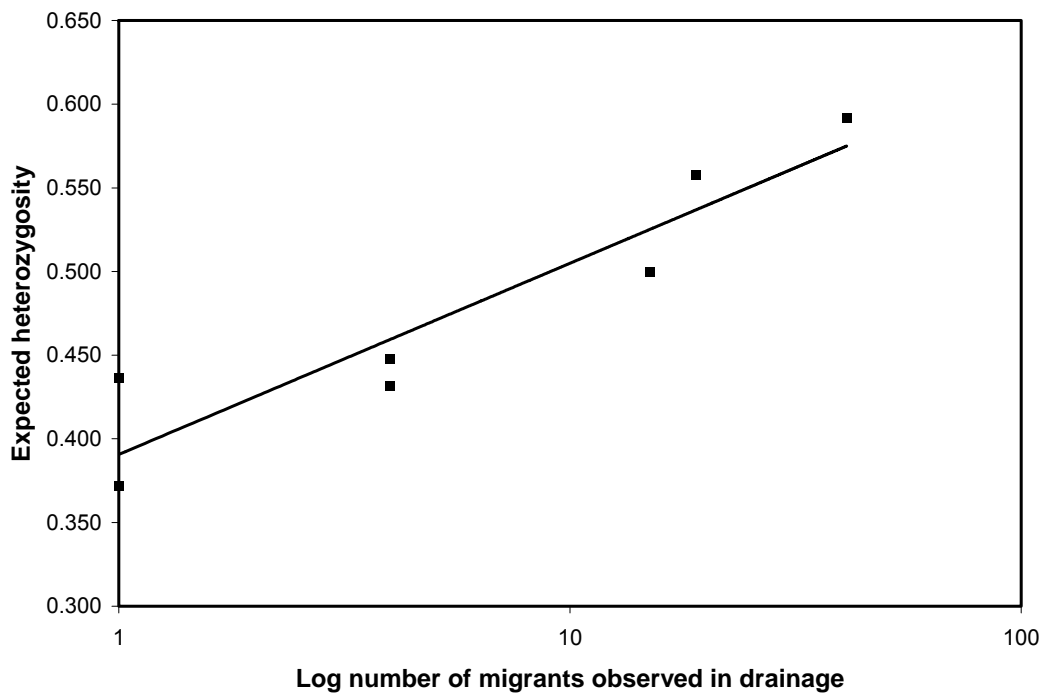
Figure 3-7: a) and b) Patterns of migrant abundance and expected heterozygosity for reference populations are similar. Each point represents a reference population. Quantities of these variables were correlated with high confidence ($y = 0.0496\mathrm{Ln}(x) + 0.3906$, $r^2 = 0.8678$, r=0.9316).

Chapter 4

Determining the spatial scale of juvenile and adult bull trout (*Salvelinus confluentus*)

movement using molecular tools

Introduction

Spatial scale is an important aspect of ecological processes (Levin, 1992), studies that are therefore able to determine the distances that individual and populations of organisms move are warranted. Such methods provide information used for defining the framework by which organisms operate in the spatial scale of their environment; critical considerations for a wide array of ecological disciplines. For the molecular ecologist and population biologist, such knowledge could be used to determine the spatial scale over which populations interact and possibly interbreed. For the ecosystem scientist, such data is critical for an understanding of the spatial scale of predator-prey interactions and flows of nutrients and energy across large areas (Finlay *et al*., 2002; McCann *et al*., 2005). For the ecotoxicologist, fundamental knowledge on the spatial scale of individual movement is vital for environmental impact assessment (Galloway *et al*., 2003). Finally, for the conservation geneticist and manager, the spatial tendencies of individual organisms and the hierarchical populations to which they belong is central to a guided management strategy for conservation efforts (Kareiva, 1990).

Direct methods of examining movements of vertebrates have traditionally been conducted by mark-recapture methods. Such estimators are notoriously labor-intensive and highly biased by sampling design (Gowan and Fausch, 1996; Koenig *et al*., 1996).

Alternate real-time movement trackers, such as radio-telemetry are expensive, can only be conducted on limited sample sizes and may be impractical in smaller organisms to which such tags cannot be affixed. Indirect methods of examining movements have recently become robust with superior analytical techniques. These methods use data contained within individual organisms themselves, including highly variable genetic markers (Rannala and Mountain, 1997), calcified structure chemistry (Elsdon and Gillanders, 2003) or soft tissue isotope tracers (Hobson, 1999).

Determining dispersal or movement between populations via genetics traditionally focused on genetic differentiation between populations (Bunn and Hughes, 1997; Wilson *et al*., 2004). Such studies used *F*-statistics to derive estimates of the numbers of migrants between populations, which likely yields inaccurate results for contemporary dispersal (see Neigel, 2002 for review). Where population differentiation is substantial, genetic assignment tests provide an alternate indirect method to infer long range contemporary migration, without being subject to the woes of *F*-statistics based approaches (Paetkau *et al*., 1995; Pearse and Crandall, 2004). When compared to direct observations of movement, assignment tests have been shown to be accurate indicators of dispersal rates (Berry *et al*., 2004) and have common applications in fisheries research and management (Hansen *et al*., 2001). Although studies of dispersal rates are common, those specifically addressing the spatial scale of movement have not relied on genetic tools.

Salmonid fishes are notorious for the variability they display in scales of movement. No species exemplifies this better than the bull trout (*Salvelinus confluentus*), an inland char species that is known to display a range of migratory life histories. In bull

trout literature, this variation in migratory life history is manifested as one of four categories: stream-resident, fluvial (river migrant), adfluvial (lake migrant) and anadromous (ocean migrant) (McPhail and Baxter, 1996). All life-histories spawn and rear as juveniles in headwater streams of mountainous regions of Western North America, but stream-residents mature and live as adults in the streams in which they were born (McPhail and Baxter, 1996). The other three migrant life histories outmigrate to areas of higher productivity and species richness, becoming piscivorous as subadults and returning to spawn in their natal tributaries as much larger, fecund individuals than their resident counterparts (McPhail and Baxter, 1996). While it is useful to categorize these life-histories, each is subject to variability in the spatial scale of movement. As such, it may be useful not only to categorize migratory life-history of such adult fish, but to quantify the magnitude of migration therein along a continuous gradient.

Previous studies addressing bull trout migration have been generally been concerned with adult movements, while juvenile movements have been studied less (Swanberg, 1997; Nerass and Spruell, 2001; Bahr and Shrimpton, 2004; Taylor *et al*., 2006). Indeed, in all salmonid literature, studies on the movements of juveniles have been largely under-represented. This may reflect a widespread inherent assumption in salmonid biology of rearing juveniles adhering to a restricted movement paradigm (RMP) (Gerking, 1959; Gowan *et al*., 1994). Direct studies on movements of juvenile fishes are inherently difficult, due to their high mortality rate and an inability to affix large external or internal tags used in most mark-recapture and radio-telemetry studies. When such studies are performed, they are also highly biased by sampling design to detect non-movement (Gowan and Fausch, 1996; Koenig *et al*., 1996), or they are aimed at studying

emigration related to life history shifts from juvenile or subadult to adult stage (Byrne *et al*., 2003; Mogen and Kaeding, 2005). The long-range migratory tendencies of non or pre-smolting juvenile salmonid fish between rearing streams therefore largely remain a mystery. Studies challenging the RMP for rearing juvenile salmonids are few, but some of those conducted have used indirect methods of determining movement (Kennedy *et al*., 2002, Rasmussen *et al*., in press). These studies have found that extreme long-range movements (>10km) of rearing juvenile salmonid fish may in fact be quite common. Such long-range movements may even pertain to bull trout (Homel and Budy, 2008). Genetic assignment tests have the potential as yet another indirect tool which may infer such extreme long-range movements with relatively little sampling rigor.

Variation in life history and scale of migration in salmonids may occur at the population level, reflecting locally adaptive traits (Taylor, 1991). Since bull trout display hierarchical population structure at fine spatial scales (Whiteley *et al*., 2006a; Warnock, Chapters 2 and 3), it may be of great importance to quantify the spatial scales of movement of such populations at each level. Such an approach would be valuable for managing salmonid fish species according to the hierarchical population composition they display at differing spatial scales (Whiteley *et al*., 2006b)

In this chapter, the spatial scale of movement in all levels of population structure and life stage (juvenile or adult) of Oldman River bull trout will be quantified using genetic assignment tools.

Methods

A total of 13 tributary streams were sampled by backpack electrofisher in the Oldman River drainage above the Oldman River Dam (Figure 2-2). These streams represented spawning and rearing tributaries with favorable habitat characteristics for bull trout population persistence (Dunham and Rieman, 1999). Within the Castle sub-basin, 6 streams were sampled: the Carbondale River (Cb), Gardiner Creek (Ga), Lost Creek (Lo), the West (Wca) and South (Sca) Castle Rivers, and Mill Creek (Mi). Within the upper Oldman sub-basin, 7 streams were sampled: South (Sra) and North (Nra) and the main-stem Racehorse (Ra) Creeks, Dutch Creek (Du), Hidden Creek (Hi), and the Lower (Lli) and Upper (Uli) Livingstone Rivers. The latter two streams were separated on the basis of a seasonally passable set of falls, which may reduce gene flow between and lead to genetic divergence of the two sites. All streams were sampled with a target sample size of 30 fish in order to minimize bias in stream representation. Fish captured were >80mm FL, so that only juveniles, subadults and adults were sampled, and not young-of-year. Migrant fish were sampled by angling, gillnet or jetboat electrofisher in the main-stems of the Oldman (OMR), Castle and Oldman Tailwater (TW) Rivers (Figure 3-1). Collections of all individuals took place during mid-summer to early fall in 2006 and 2007.

Hierachical STRUCTURE analysis (Vaha *et al*., 2007; Chapter 2) was conducted in the program STRUCTURE 2.2 (Pritchard *et al*., 2000) to determine all levels of population structure in the drainage and provide assignment tests for each individual fish in the tributary reference populations.

To assess scale of movement of tributary-rearing juvenile fishes, all adults and subadults caught in tributaries were removed by excluding fish >200mm FL. The remaining fish were identified by population, and assignment tests from hierarchical STUCTURE analysis were used to determine if each individual fish had dispersed from the geographical area of the population in which it was born. Although sampling efforts were roughly equivalent between streams, the number of streams sampled that contributed to any given population could only be determined a posteriori. To correct for the unequal sampling effort between "home ranges" of populations as a consequence of this, the number of fish self-assigned to a population was modified based on the number of streams present in the population's home range. Self-assigned fish were assigned a conservative estimate of 0 km moved. For fish assigned to an alternate population, migration distance between populations was calculated as the mean distance from the midpoint of all spawning tributaries in a population of origin to the first node in the spatial network of the drainage, added to the linear stream-distance of the node to the site in which it was captured. Because probability of sampling migrant individuals from any given population may be biased by unequal population sizes, sample sizes of migrant fish were corrected based on census size estimates for each stream relative to those for the population of origin:

$$Nm_p = Nm_t \left(\frac{Nc_t}{Nc_p} \right)$$

Where N is the number of juvenile fish, m is migrants, c is census size, $_t$ is the target stream sampled and $_p$ is the population of origin. The census size values were derived from sampling density data as number of fish caught per meter of stream backpack

electrofishing (assuming 100% capture rate), multiplied by the habitable length of the stream. Stream distances were calculated in the Garmin MapSource[TM] computer program.

To assess scale of movement of migrant subadult and adult fish, assignment tests were performed in GENECLASS 2 (Piry *et al*., 2004) in order to determine the most probable (i.e. highest posterior probability) population-of-origin for each migrant fish caught in river main-stems and the reservoir; while adults and subadults (fish >200mm FL) caught in tributaries were assigned to most probable population-of-origin based on assignment test results of hierarchical STRUCTURE analysis. Migration distance for each fish was calculated as the mean distance of the midpoint of all spawning tributaries in the population of origin to the first node in the spatial network of the drainage, added to the linear stream-distance of the node to the site in which it was captured. Self-assigned adult and subadult fish were discarded from analysis because these fish may have been migrants that had undergone early spawning migrations to their stream-of-origin, not stream residents with nil movement as was assumed for juvenile fish.

Average distances moved by population and archipelago were determined for both juvenile and adult migrant fish. This scale of movement for migrants only was compared between populations and archipelagos by a one-way analysis of variance (ANOVA). A Tamhane's T2 (Tamhane, 1979) post-hoc test was used to determine which populations and archipelagos of migrant fish displayed significantly different mean scales of movement. The Upper Livingstone River was not included in this analysis for adult and subadult fish because only a single migrant fish was assigned to this population and the adult population is considered resident (Chapter 3).

Proportions of long-range (>10km displacement) migratory fish by population

and archipelago were also compared by a Tukey-type multiple range test for juvenile fish

only (Zar, 1999 pp 563). By taking the product of the proportion moved and the average

distance moved by juvenile migrants, the mean scale of movement for each population,

archipelago and the total drainage could be determined:

$$\overline{X} = \sum pd$$

where $\overline{X}$ is the mean scale of movement of the population and p is the proportion of

migrants of a particular d distance traveled.

Results

A total of 3 archipelagos containing 7 populations were identified by hierarchical

STURCTURE analysis (Chapter 2, Figure 2-5). A total corrected number of 172.5

juvenile fish were classified to these identified populations and archipelagos of origin and

distances moved for each was calculated (Table 4-1). Long-range, inter-stream migratory

proportions varied from 4% (Upper Livingstone) to 47% (Mill) by population and 11%

(Livingstone) to 34% (Castle) by archipelago (Table 4-1). Tukey-type multiple range

tests revealed many of these differences were statistically significant at both the 0.2 and

0.05 levels (Table 4-3, Figure 4-1). A total of 97 adult and subadult migrant fish were

classified to population and archipelago of origin and distances move for each was

calculated (Table 4-2).

Mean distances moved for each population and archipelago of juvenile migrant

fish was calculated (Table 4-4). An analysis of variance revealed significant differences

between populations and archipelagos for juvenile migrant fish (Table 4-6). Post-hoc tests revealed many significant differences between populations and archipelagos (Table 4-6, Figures 4-3 and 4-4).

Mean distances moved for each population and archipelago of adult and subadult migrant fish was calculated (Table 4-5). Results of the analysis of variance for adult and subadult fish were similar to those conducted on juveniles, with significant differences in mean scale of movement between populations and archipelagos (Table 4-7). Post-hoc tests revealed many significant differences between populations and archipelagos (Table 4-7, Figures 4-5 and 4-6).

Mean scale of movement for juvenile populations ranged from 3.7km (Upper Livingstone) to 35.6km (Mill), and archipelagos from 9.8km (Oldman) to 23.5km (Castle) (Table 4-8, Figures 4-7 and 4-8). The entire dataset revealed a mean scale of movement of 17.1km for juvenile bull trout (Tables 4-7 and 4-8).

Discussion

Bull trout were found to exhibit variable scales of movement as juveniles, adults and subadults. The fact that both proportion and magnitude of movements of migratory individuals differ based on genetic group of origin supports the notion that populations are highly variable with respect to migratory tendencies. It is possible that such migratory tendencies evolve as locally adaptive traits, to aid populations in optimally exploiting the full spatial scale of their ecosystem (Taylor, 1991). Alternatively, variation in observed migratory tendencies may arise due to conditional dispersal strategies based on local

141

habitat features and population size (McPeek and Holt, 1992). Because populations of lotic salmonid fishes are arranged as a nested hierarchy within a spatial network of a drainage system, this variation was manifested at different levels. At the population and archipelago level observed in this drainage, significant differences could be found for both proportion and spatial scale of migration. Results of studies such as these that do examine layers of population structure may therefore be able to provide necessary information for a guided management strategy within the spatial hierarchy (Whiteley *et al.*, 2006b)

Mean scale of movement for long-range migrant fish specifies the spatial scale over which populations may operate and interact with each other and their environment. Proportions of long-range migratory individuals provide data on the dispersal propensities of specific populations. Together, these values provide insight into the migratory tendencies of populations and are important when modeling ecological interactions over large spatial scales (McCann *et al.*, 2005) or for management purposes. The final value determined in juvenile fish, overall mean scale of movement, may be conservative estimates of migration, given that self-assigned individuals were all assigned nil movement values. Caution must be stressed when interpreting these values. The mean scale of movement measured serves mainly as a "benchmark" value by which cross literature comparisons may be made. Because number of non movers are more common that movers, however, the mean is not an ideal indicator of central tendency, therefore the overall "mean scale" of movement found should not be interpreted as a particularly ecologically relevant quantity.

It should be noted that spatial resolution in studies using these tools will be driven by the degree of population divergence at any given spatial scale. For this study, there was a minimum detection limit of 6.5km separating the home ranges of the Upper and Lower Livingstone populations, but more typical values of ~30-40 km of linear stream distance separated the home ranges of most adjacent populations within archipelagos. As such, this technique only has the potential to detect movement of extreme long range for this species. Further limitations may stem from the confidence in the accuracy of assignment tests. In this study, results of all assignment tests were assumed to be correct because of the high degree of divergence between populations; although erroneous conclusions may be drawn from such assignment tests for two main (although unlikely) reasons: first, an admixture event may have occurred in the area of question between parents with majority membership of a different population, leading to inflated rates of migration. Or, secondly, some individuals may be missing up to 4 alleles and could show unequal representation of some homoplasic or population-specific alleles by chance. As such, caution is stressed in interpreting results from these tests, especially in other species where genetic divergence between populations may not be high.

Juvenile fish movement

The restricted movement paradigm has received a considerable amount of attention in the past 15 years from various studies on salmonid fish movement. Gowan *et al*. (1994) argue that the RMP need not apply to stream-resident adult salmonids, as movements within stream-reaches (tens to low hundreds of meters) are very common. In contrast, Rodriguez (2002) argues that movements are usually restricted to a small

proportion of the population, and that the RMP still explains the majority of the migratory state of such fish. While this debate is not currently resolved, considerable gaps in knowledge still exit with respect to juveniles still in the rearing stages of life. There have been several studies assessing dispersal in juvenile bull trout, which have found dispersal for juveniles of all ages in migrant populations to be common (Mogen and Kaeding, 2005; Downs *et al*., 2006) and possibly of large spatial scale (over 45 km downstream (Homel and Budy, 2008); however, these studies have been designed to measure stream emigration dispersal or local-scale movements, possibly related to either life-stage shift in migratory fish (i.e. smolting-equivalent transformation) or habitat shifts for rearing individuals. The technique used in this study specifically has the capability to detect long-range inter-stream movements relating to the latter motive. Such extreme long-range movements may be extremely important when considering ecological interactions among hierarchical population levels over large spatial scales (Gaines and Bertness, 1993). Studies which have found evidence for these movements in juvenile salmonids have used indirect techniques such as otolith microchemistry (Kennedy *et al*., 2002) and tissue stable isotope analysis (Rasmussen *et al*., in press). Results of this study support these findings, and point out the need for an expanding research field into rearing-stage juvenile salmonid fish movement.

It is possible that the observed variation in long-range inter-stream migration rates reflects a conditional dispersal strategy for this life stage of bull trout (McPeek and Holt, 1992). Small-scale movements of fishes may largely be coupled to habitat quality, possibly due to "monitoring" of habitat at a reach-scale (hundreds of meters) in order to maximize foraging efficiency (Gowan and Fausch, 2002). It is possible that the long-

range movements or emigration of juvenile fish observed in this study reflects a future result of this behaviour, as specific stream reaches may provide poor habitat or are saturated past carrying capacity (Homel and Budy, 2008). If monitoring behaviour reveals that it is suboptimal to stay in the reach-scale area, a fish may have to either increase the area it monitors over or undergo long-range dispersal to an alternate stream. This may be of particular importance to bull trout, which mature quickly and have very specialized rearing habitat requirements (McPhail and Baxter, 1996). For example, Mill Creek receives strong spawning runs (Gerrand and Watmough, 1998), but may not be able to support the amount of juveniles produced by these spawning events. Mill Creek may then act as a source population from which juvenile fish are forced to emigrate in order to find suitable rearing habitat. In contrast, the Carbondale area may act as a recipient system, receiving Mill Creek immigrants. A suite of measurable habitat features may even make such phenomena predictable in stream ecosystems: Mill creek contains a large amount of alluvial gravels and cobble with significant groundwater inputs and is easily accessible for migrating adults en route to spawn. These factors are conducive to favorable spawning habitat (Baxter and Hauer, 2000). The stream only contains short stretches with large woody debris, instream boulder and rock cover, has high presence of large piscivorous trout and invasive brook trout, and does not have any large tributaries, all factors which reduce the suitability of this stream as a nursery habitat for bull trout (Dunham and Rieman, 1999); therefore, it may be that the home range of this population contains a high ratio of spawning relative to rearing habitat. The presence of invasive brook trout in particular raises interesting questions. It is possible that the carrying capacity for rearing juveniles in this stream was once greater, but introduced brook trout

competitors may have reduced the amount of habitat and resources available (Gunckel *et al*., 2002). In this respect, the presence of an invasive species may have increased the migratory component of juvenile fish in this stream, a compensatory strategy which may partially explain the resiliency of such robust populations to brook trout invasion (Rieman *et al*., 2006). In contrast to Mill Creek, the Carbondale system contains long sections of river with high instream cover and a dendritic drainage system with several large tributaries of high habitat quality; in effect, creating the opposite effect of a high ratio of rearing relative to spawning habitat. This could explain the significant difference observed between proportions of inter-stream migratory juvenile fish observed in the Mill Creek population relative to that of the Carbondale (Figure 4-1).

Dispersal strategy differences between resident and migrant populations may also reflect inherent bioenergentic differences between these life history strategies, even at the juvenile stage. Juvenile migrant brook trout have higher consumption rates and different habitat requirements than their resident counterparts (Morinville and Rasmussen, 2003; 2006); furthermore, populations containing only residents do not experience outmigration (Morinville and Rasmussen, 2006). In bull trout, a similar phenomenon was observed in that the only resident population in the drainage (Upper Livingstone) had the lowest observed rates and scale of juvenile inter-stream movement. It is conceivable that the local rearing habitat of such resident populations may provide ample resources for such fish, precluding a need to disperse; alternatively, the metabolic demands of migration may be unfeasible for such fish, reflecting an unconditional strategy for limited dispersal in such populations.

From a conservation perspective, the finding of common long-range inter-stream movement, even at this life stage, further emphasizes the importance of reducing habitat fragmentation in coldwater lotic ecosystems (Morita and Yokota, 2002). There may be a natural discontinuity between streams of high spawning suitability in relation to rearing habitat for juvenile bull trout that occupy geologically heterogeneous regions. This may be a driving force in determining juvenile salmonid dispersal patterns within the spatial network of coldwater ecosystems. Such features highlight the importance of habitat conservation and interpretation of streams as spatially connected interdependent ecosystems (Harding *et al*., 1998).

Adult fish movement

Adult fish movement was determined to be variable based on population or archipelago of origin. This stresses the importance of locally and hierarchically applied management, as the degree of residency and spatial scale of migration are two extremely important considerations when determining conservation strategies. Highly migratory archipelagos, such as the Castle and the three populations of which it is composed, must have unobstructed migratory routes to allow robust population persistence (Rieman and Allendorf, 2001). Less migratory archipelagos, such as the Oldman and the two populations of which it is composed must be studied further to determine what mechanisms limit the migratory capacity of these groups.

The spatial scale of movement in resident populations, such as that found in the Upper Livingstone River, could not be examined by the methods of this chapter due to the rare occurrence of migrants originating from this population and the high threshold of

detecting movement. Nevertheless, this population must have high habitat conservation priority in their stream of origin, since high habitat complexity is needed to support a self-sustaining population (Dunham and Rieman, 1999).

Movement scale of migrant fish was observed to be highest in those populations and archipelagos with the easiest access to spawning habitat. This supports the stream-access theory outlined in the discussions of Chapters 2 and 3 as a potential limitation on migratory tendencies. It may be adaptive for bull trout populations that must overcome significant barriers during spawning runs to reduce their spatial scale of migration, so as to maintain optimal access to migration passages that have restricted windows of opportunity for passage. This is conceivable for populations spawning in stream systems which have barriers that are temporally variable in degree of navigability based on stream flows. Since adult bull trout migrate in lower flow late summer months, it may be advantageous to continuously "monitor" barriers which may cease to exist during unpredictable stream level risings. Such a mechanism has been proposed to explain the movements of stream-resident salmonids (Gowan and Fausch, 2002) with respect to optimal foraging, but may also explain movements with respect to migratory route accessibility. Because levels often drop quickly after such rises in these months, a fish that lacks this mechanism may not return to the barrier in time to take advantage of such an event. An alternate explanation simply relies on individual migrant fish having a conditional dispersal strategy based on local population size or density, rather than a "hard wired" mechanism to limit dispersal scale (McPeek and Holt, 1992). This is similar to the explanation for juvenile movements explained in the preceding section. It is possible that higher adult population sizes in areas such as the Castle archipelago lead to

conditions of higher intra-archipelago competition within the home geographical range

for these fish. This may therefore lead to dispersal behaviour in surplus numbers of less

dominant fish (Hughes, 1992), explaining the higher spatial scale of migration seen in

this archipelago relative to the other two. It is interesting to note that this explanation is

theoretically not independent of barriers as the ultimate cause, as it is possible that the

restricted stream access serves to reduce population sizes in the first place (Chapters 2

and 3).

This study did not separate subadult and adult fish in the analysis. It is likely that

these stages do in fact have differing migratory habits. Future studies with greater sample

sizes of each should measure migration separately for these two life stages.

Conclusions

Results of this study indicate that genetic assignment tests may be valuable tools

not only for detecting movement, but for quantifying spatial scale of movement for

organisms displaying significant population structure. This is likely to be a powerful tool

at spatial scales over which organisms display strong population structure.

This paper provides evidence that migratory tendency at all levels of population

structure is variable, both among juvenile and adult bull trout. The mechanism driving

such variation is unclear, and may reflect either unconditional "locally adaptive", or

conditional or "behaviourally plastic" dispersal strategies. Of particular importance is the

result that juvenile long-range inter-stream movement is common in most bull trout

populations, and provides further evidence against a restricted movement paradigm for

salmonid fish of this life stage. Research into the inter-stream migratory trends of such

fish, with direct observational data is needed in future studies.

References

Bahr, M. A., & Shrimpton, J. M. (2004). Spatial and quantitative patterns of movement in large bull trout (*Salvelinus confluentus*) from a watershed in north-western British Columbia, Canada, are due to habitat selection and not differences in life history. *Ecology of Freshwater Fish, 13*(4), 294-304.

Baxter, C. V., & Hauer, F. R. (2000). Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (*Salvelinus confluentus*). *Canadian Journal of Fisheries and Aquatic Sciences, 57*(7), 1470-1481.

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology, 13*(3), 551-561.

Bunn, S. E., & Hughes, J. M. (1997). Dispersal and recruitment in streams: Evidence from genetic studies. *Journal of the North American Benthological Society, 16*(2), 338-346.

Byrne, C. J., Poole, R., Rogan, G., Dillane, M., & Whelan, K. F. (2003). Temporal and environmental influences on the variation in Atlantic salmon smolt migration in the Burrishoole system 1970-2000. *Journal of Fish Biology, 63*(6), 1552-1564.

Downs, C. C., Horan, D., Morgan-Harris, E., & Jakubowski, R. (2006). Spawning demographics and juvenile dispersal of an adfluvial bull trout population in Trestle Creek, Idaho. *North American Journal of Fisheries Management, 26*(1), 190-200.

Dunham, J. B., & Rieman, B. E. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications, 9*(2), 642-655.

Elsdon, T. S., & Gillanders, B. M. (2003). Reconstructing migratory patterns of fish based on environmental influences on otolith chemistry. *Reviews in Fish Biology and Fisheries, 13*(3), 219-235.

Finlay, J. C., Khandwala, S., & Power, M. E. (2002). Spatial scales of carbon flow in a river food web. *Ecology, 83*(7), 1845-1859.

Gaines, S. D., & Bertness, M. (1993). The dynamics of juvenile dispersal - why field ecologists must integrate. *Ecology, 74*(8), 2430-2435.

Galloway, B. J., Munkittrick, K. R., Currie, S., Gray, M. A., Curry, R. A., & Wood, C. S. (2003). Examination of the responses of slimy sculpin (*Cottus cognatus*) and white sucker (*Catostomus commersoni*) collected on the Saint John River (Canada) downstream of pulp mill, paper mill, and sewage discharges. *Environmental Toxicology and Chemistry, 22*(12), 2898-2907.

Gerking, S. D. (1959). The restricted movement of fish populations. *Biological Reviews of the Cambridge Philosophical Society, 34*(2), 221-242.

Gerrand, M., & Watmough, M. (1998). 1998 Spawning survey of bull trout in the Oldman River drainage. Prepared for Alberta Fish and Wildlife. Alberta Conservation Association. Blairmore, AB. 43 Pp.

Gowan, C., & Fausch, K. D. (1996). Mobile brook trout in two high-elevation Colorado streams: Re-evaluating the concept of restricted movement. *Canadian Journal of Fisheries and Aquatic Sciences, 53*(6), 1370-1381.

Gowan, C., & Fausch, K. D. (2002). Why do foraging stream salmonids move during summer? *Environmental Biology of Fishes, 64*(1-3), 139-153.

Gowan, C., Young, M. K., Fausch, K. D., & Riley, S. C. (1994). Restricted movement in stream salmonids - a paradigm lost. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(11), 2626-2637.

Gunckel, S. L., Hemmingsen, A. R., & Li, J. L. (2002). Effect of bull trout and brook trout interactions on foraging habitat, feeding behavior, and growth. *Transactions of the American Fisheries Society, 131*(6), 1119-1130.

Hansen, M. M., Kenchington, E., & Nielsen, E. E. (2001). Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries, 2*(2), 93-112.

Harding, J. S., Benfield, E. F., Bolstad, P. V., Helfman, G. S., & Jones, E. B. D. (1998). Stream biodiversity: The ghost of land use past. *Proceedings of the National Academy of Sciences of the United States of America, 95*(25), 14843-14847.

Hobson, K. A. (1999). Tracing origins and migration of wildlife using stable isotopes: a review. *Oecologia, 120*(3), 314-326.

Homel, K., & Budy, P. (2008). Temporal and spatial variability in the migration patterns of juvenile and subadult bull trout in northeastern Oregon. *Transactions of the American Fisheries Society, 137*(3), 869-880.

Hughes, N. F. (1992). Ranking of feeding positions by drift-feeding arctic grayling (*Thymallus-arcticus*) in dominance hierarchies. *Canadian Journal of Fisheries and Aquatic Sciences, 49*(10), 1994-1998.

Kareiva, P. (1990). Population dynamics in spatially complex environments: theory and data. *Philosphical Transactions of the Royal Society of London, 330*, 175-190.

Kennedy, B. P., Klaue, A., Blum, J. D., Folt, C. L., & Nislow, K. H. (2002). Reconstructing the lives of fish using Sr isotopes in otoliths. *Canadian Journal of Fisheries and Aquatic Sciences, 59*(6), 925-929.

Koenig, W. D., VanVuren, D., & Hooge, P. N. (1996). Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology & Evolution, 11*(12), 514-517.

Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology, 73*(6), 1943-1967.

McCann, K. S., Rasmussen, J. B., & Umbanhowar, J. (2005). The dynamics of spatially coupled food webs. *Ecology Letters, 8*(5), 513-523.

McPeek, M. A., & Holt, R. D. (1992). The evolution of dispersal in spatially and temporally varying environments. *American Naturalist, 140*(6), 1010-1027.

McPhail, J. D., & Baxter, J. S. (1996). A review of bull trout (*Salvelinus confluentus*) life-history and habitat use in relation to compensation and improvement opportunities. University of British Columbia. Fisheris Management Report No. 104. Vancouver, BC. 31 pp.

Morita, K., & Yokota, A. (2002). Population viability of stream-resident salmonids after habitat fragmentation: a case study with white-spotted charr (*Salvelinus leucomaenis*) by an individual based model. *Ecological Modelling, 155*(1), 85-94.

Morinville, G. R., & Rasmussen, J. B. (2003). Early juvenile bioenergetic differences between anadromous and resident brook trout (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences, 60*(4), 401-410.

Morinville, G. R., & Rasmussen, J. B. (2006). Does life-history variability in salmonids affect habitat use by juveniles? A comparison among streams open and closed to anadromy. *Journal of Animal Ecology, 75*(3), 693-704.

Neigel, J. E. (2002). Is F-ST obsolete? *Conservation Genetics, 3*(2), 167-173.

Neraas, L. P., & Spruell, P. (2001). Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology, 10*(5), 1153-1164.

Paetkau, D., Calvert, W., Stirling, I., & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology, 4*(3), 347-354.

Pearse, D. E., & Crandall, K. A. (2004). Beyond F-ST: Analysis of population genetic data for conservation. *Conservation Genetics, 5*(5), 585-602.

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.

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

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

Rasmussen, J.B., Trudeau, V., & Morinville, G. (in press). Estimating the scale of fish feeding movements in rivers using $\delta^{13}C$ signature gradients. *Journal of animal ecology*

Rieman, B. E., & Allendorf, F. W. (2001). Effective population size and genetic conservation criteria for bull trout. *North American Journal of Fisheries Management, 21*(4), 756-764.

Rieman, B. E., Peterson, J. T., & Myers, D. L. (2006). Have brook trout (*Salvelinus fontinalis*) displaced bull trout (*Salvelinus confluentus*) along longitudinal gradients in central Idaho streams? *Canadian Journal of Fisheries and Aquatic Sciences, 63*(1), 63-78.

Rodriguez, M. A. (2002). Restricted movement in stream fish: The paradigm is incomplete, not lost. *Ecology, 83*(1), 1-13.

Swanberg, T. R. (1997). Movements of and habitat use by fluvial bull trout in the Blackfoot River, Montana. *Transactions of the American Fisheries Society, 126*(5), 735-746.

Tamhane, A. C. (1979). Comparison of procedures for multiple comparisons of means with unequal variances. *Journal of the American Statistical Association, 74*(366), 471-480.

Taylor, E. B. (1991). A review of local adaptation in salmonidae, with particular reference to pacific and atlantic salmon. *Aquaculture, 98*(1-3), 185-207.

Taylor, E. B., & Costello, A. B. (2006). Microsatellite DNA analysis of coastal populations of bull trout (*Salvelinus confluentus*) in British Columbia: zoogeographic implications and its application to recreational fishery management. *Canadian Journal of Fisheries and Aquatic Sciences, 63*(5), 1157-1171.

Vaha, J. P., Erkinaro, J., Niemela, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology, 16*(13), 2638-2654.

Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2006). Can common species provide valuable information for conservation? *Molecular Ecology, 15*(10), 2767-2786.

Whiteley, A. R., Spruell, P., Rieman, B. E., & Allendorf, F. W. (2006). Fine-scale genetic structure of bull trout at the southern limit of their distribution. *Transactions of the American Fisheries Society, 135*(5), 1238-1253.

Wilson, A. J., Hutchings, J. A., & Ferguson, M. M. (2004). Dispersal in a stream dwelling salmonid: Inferences from tagging and microsatellite studies. *Conservation Genetics, 5*(1), 25-37.

Zar, J. H. (1999). *Biostatistical Analysis* (4 ed.). Upper Saddle River, N.J.: Prentice-Hall.

Corrected number of migrant fish[1]

population	archipelago	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli	Self	Corrected number of self-assigned fish[2]	proportion of population as migrants >10 km	proportion of archipelago as migrants >10km
Cbp	CA	0.00	0.00	0.00	0.59	1.09	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00	46	15.3	0.12	
Wcap	CA	2.93	0.34	0.64	0.00	1.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15	15	0.28	
Mip	CA	2.51	0.36	6.02	5.14	0.00	0.00	0.00	0.00	0.00	5.11	0.49	0.00	0.00	45	22.5	0.47	0.34
Rap	OMR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.61	0.88	0.00	0.00	40	13.3	0.21	
Hip	OMR	0.00	0.00	0.45	0.00	0.43	0.00	2.31	0.23	0.37	0.00	0.00	0.00	0.00	31	15.5	0.20	0.20
Llip	LI	0.00	0.00	3.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	0.00	6.53	15	15	0.20	
Ulip	LI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.13	9.19	0.00	19	19	0.04	0.11

1. Corrected number of migrant fish for any population is derived from $Nm_p = Nm_t \left(\dfrac{Nc_t}{Nc_p} \right)$ this value corrects for unequal census sizes of streams relative to the size of the population of origin.

2. Corrected number of self-assigned fish for any population is derived from $Ns = n / $ (number of streams in the population's home range). This value standardizes sampling effort of self-assigned fish between populations.

Distances between populations and sites (km)

							Site caught							
	pop	arch	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli
	Cbp	CA	n/a	n/a	n/a	45	37	56.4	138	143	144	140	149	155	161
Population/	Wcap	CA	45	44	46	n/a	47	67.9	149	154	155	152	160	166	173
archipelago	Mip	CA	52	51	53	53	n/a	n/a	134	138	140	136	145	150	157
of	Rap	OMR	142	141	142	153	155	120	n/a	n/a	n/a	32	41	47	53
origin	Hip	OMR	145	144	145	156	158	123	32.9	37	39	n/a	n/a	37	43
	Llip	LI	150	149	151	166	168	133	43	48	49	37	36	n/a	6.5
	Ulip	LI	157	156	157	173	175	139	49.5	54	56	44	42	6.5	n/a

Table 4-1: Numbers and distances moved of juvenile fish by population and archipelago.

Site caught

population	archipelago	CA3	CA2	CA1	ORR	TW	OMR1	OMR2	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli
Cbp	CA		5	4	5	3	2					1		1				1			
Wcap	CA				1		2	1													
Mip	CA	9	6	5	11	6	4					2									
Rap	OMR						4												1		
Hip	OMR						12	3													1
Llip	LI						1													2	3
Ulip	LI						1														

Distances between populations and sites (km)

Site caught

pop	arch	CA3	CA2	CA1	ORR	TW	OMR1	OMR2	Cb	Ga	Lo	Wca	Sca	Mi	Ra	Sra	Nra	Du	Hi	Lli	Uli
Cbp	CA	19	26	37	65	83	107	134	n/a	n/a	n/a	45	37	56	138	143	144	140	149	155	161
Wcap	CA	27	38	49	76	95	119	145	45	44	46	n/a	47	68	149	154	155	152	160	166	173
Mip	CA	35	35	35	60	79	103	130	52	51	53	53	n/a	n/a	134	138	140	136	145	150	157
Rap	OMR	n/a	n/a	n/a	n/a	n/a	34.4	26	142	141	142	153	155	120	n/a	n/a	n/a	32	41	47	53
Hip	OMR	n/a	n/a	n/a	n/a	n/a	37.4	15.8	145	144	145	156	158	123	33	37	39	n/a	n/a	37	43
Llip	LI	n/a	n/a	n/a	n/a	n/a	47.5	20.7	150	149	151	166	168	133	43	48	49	37	36	n/a	6.5
Ulip	LI	n/a	n/a	n/a	n/a	n/a	54	27.2	157	156	157	173	175	139	50	54	56	44	42	6.5	n/a

Table 4-2: Numbers and distances moved of adult and subadult fish by population and archipelago.

a)

Population	Ulip	Cbp	Hip	Llip	Rap	Wcap	Mip
Proportion as migrants >10km (p_i)	0.04	0.12	0.20	0.20	0.21	0.28	0.47
Ranked arcsin transformation, radians (p'_i)	0.23	0.38	0.48	0.47	0.49	0.57	0.75
Ranked arcsin transformation, degrees (p'_i)	13.24	21.64	27.34	27.15	28.20	32.43	43.09

Comparisons	Difference (p'_a-p'_b)	SE = root ((410.35/(n_a+0.5))+(410.35/(n_b+0.5)))	Test statistic (q = diff/SE)	$q_{0.05,\infty,7}$	$q_{0.20,\infty,7}$	Conclusion, α = 0.05 (H_o: p_a=p_b)	Conclusion, α = 0.20 (H_o: p_a=p_b)
Mi vs Uli	29.856	4.836	6.174	4.17	3.39	reject H_o	reject H_o
Mi vs Cb	21.450	5.711	3.756	4.17	3.39	accept H_o	reject H_o
Mi vs Hi				4.17	3.39	accept H_o	accept H_o
Mi vs Lli				4.17	3.39	accept H_o	accept H_o
Mi vs Ra				4.17	3.39	accept H_o	accept H_o
Mi vs Wca				4.17	3.39	accept H_o	accept H_o
Wca vs Uli	19.194	5.750	3.338	4.17	3.39	accept H_o	accept H_o
All other comparisons				4.17	3.39	accept H_o	accept H_o

b)

Population	LI	OMR	CA
Proportion as migrants >10km (p_i)	0.11	0.2	0.34
Ranked arcsin transformation, radians (p'_i)	0.36	0.48	0.63
Ranked arcsin transformation, degrees (p'_i)	20.4	27.2	35.9

Comparisons	Difference (p'_a-p'_b)	SE = root ((410.35/(n_a+0.5))+(410.35/(n_b+0.5)))	Test statistic (q = diff/SE)	$q_{0.05,\infty,3}$	$q_{0.20,\infty,3}$	Conclusion, α = 0.05 (H_o: p_a=p_b)	Conclusion, α = 0.20 (H_o: p_a=p_b)
CA vs LI	15.517	3.511	4.420	3.63	2.42	reject H_o	reject H_o
CA vs OMR	8.645	4.036	2.142	3.63	2.42	accept H_o	accept H_o
OMR vs LI	6.872	4.295	1.600	3.63	2.42	accept H_o	accept H_o

Table 4-3: Tukey-type test for multiple comparisons on proportions of migratory juvenile fish by a) population and b) archipelago.

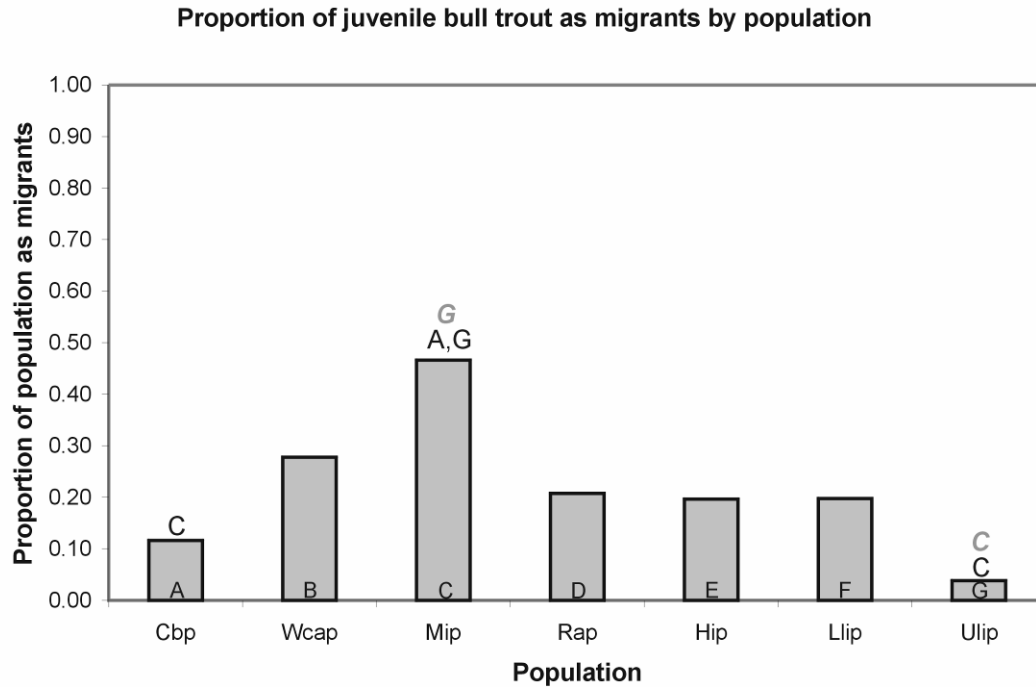**Proportion of juvenile bull trout as migrants by population**

Figure 4-1: Proportions of each population as migratory for juvenile fish. Letters above bars represent post-hoc significant differences, in black at a significance level of 0.2 and grey italic at a significance level of 0.05.

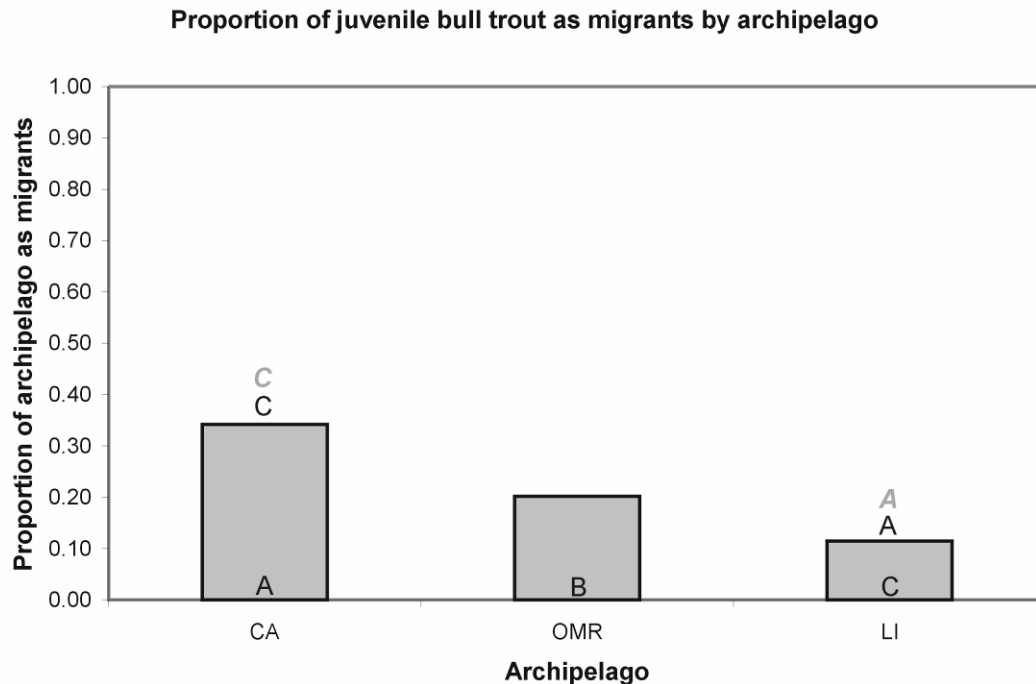**Proportion of juvenile bull trout as migrants by archipelago**

Figure 4-2: Proportions of each archipelago as migratory for juvenile fish. Letters above bars represent post-hoc significant differences, in black at a significance level of 0.2 and grey italic at a significance level of 0.05.

157

a)

	Cbp	Wcap	Mip	Rap	Hip	Llip	Ulip
				population			
mean (km)	41.2	46.11667	74.35789	34.325	32.9	47.46667	10.07
s.e. (km)	4.1	0.360015	8.654851	2.225	0	18.28233	3.57
n*	2	6	19	4	2	12	10

b)

	CA	OMR	LI
		archipelago	
mean (km)	65.62593	33.85	30.46818
s.e. (km)	6.600217	1.438923	10.70011
n*	27	6	22

c)

	total
	drainage
mean (km)	48.09636
s.e. (km)	2.982166
n*	55

* numbers of fish were rounded to the nearest whole number for corrected numbers of self assigned and migratory individuals.

Table 4-4: Mean distance (+/- s.e.) moved of juvenile migrant fish by a) population, b) archipelago and c) drainage.

a)

	Cbp	Wcap	Mip	Rap	Hip	Llip	Ulip
				population			
mean (km)	59.48636	114.65	54.69535	35.72	33.7	23.06667	54
s.e. (km)	6.663795	14.34861	3.412515	1.32	2.247221	7.614796	n/a
n*	22	4	43	5	16	6	1

b)

	CA	OMR	LI
		archipelago	
mean (km)	59.69855	34.18095	27.48571
s.e. (km)	3.49075	1.733694	7.806787
n*	69	21	7

c)

	total
	drainage
mean (km)	51.84948
s.e. (km)	2.856038
n*	97

* numbers of fish were rounded to the nearest whole number for corrected numbers of self assigned individuals.

Table 4-5: Mean distance (+/- s.e.) moved of adult and subadult migrant fish by a) population, b) archipelago and c) drainage.

a)

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28907.589	6	4817.932	3.258	.009
Within Groups	70982.110	48	1478.794		
Total	99889.699	54			

Multiple Comparisons

(I) Population	(J) Population	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Cb	Wca	-4.9167	4.1158	1.000	-994.657	984.824
	Mi	-33.1579	9.5769	.076	-68.387	2.071
	Ra	6.8750	4.6648	1.000	-156.996	170.746
	Hi	8.3000	4.1000	.999	-1061.619	1078.219
	Lli	-6.2667	18.7364	1.000	-78.136	65.603
	Uli	31.1300	5.4364	.216	-23.203	85.463
Wca	Cb	4.9167	4.1158	1.000	-984.824	994.657
	Mi	-28.2412	8.6623	.087	-58.726	2.243
	Ra	11.7917	2.2539	.223	-8.267	31.850
	Hi	13.2167[*]	.3600	.000	11.188	15.245
	Lli	-1.3500	18.2859	1.000	-72.810	70.110
	Uli	36.0467[*]	3.5881	.000	21.216	50.877
Mi	Cb	33.1579	9.5769	.076	-2.071	68.387
	Wca	28.2412	8.6623	.087	-2.243	58.726
	Ra	40.0329[*]	8.9363	.005	9.045	71.020
	Hi	41.4579[*]	8.6549	.003	10.983	71.933
	Lli	26.8912	20.2275	.991	-45.775	99.557

	Uli	64.2879*	9.3622	.000	32.476	96.100
Ra	Cb	-6.8750	4.6648	1.000	-170.746	156.996
	Wca	-11.7917	2.2539	.223	-31.850	8.267
	Mi	-40.0329*	8.9363	.005	-71.020	-9.045
	Hi	1.4250	2.2250	1.000	-19.811	22.661
	Lli	-13.1417	18.4172	1.000	-84.570	58.287
	Uli	24.2550*	4.2066	.002	8.167	40.343
Hi	Cb	-8.3000	4.1000	.999	-1078.219	1061.619
	Wca	-13.2167*	.3600	.000	-15.245	-11.188
	Mi	-41.4579*	8.6549	.003	-71.933	-10.983
	Ra	-1.4250	2.2250	1.000	-22.661	19.811
	Lli	-14.5667	18.2823	1.000	-86.028	56.894
	Uli	22.8300*	3.5700	.003	7.971	37.689
Lli	Cb	6.2667	18.7364	1.000	-65.603	78.136
	Wca	1.3500	18.2859	1.000	-70.110	72.810
	Mi	-26.8912	20.2275	.991	-99.557	45.775
	Ra	13.1417	18.4172	1.000	-58.287	84.570
	Hi	14.5667	18.2823	1.000	-56.894	86.028
	Uli	37.3967	18.6276	.772	-34.012	108.805
Uli	Cb	-31.1300	5.4364	.216	-85.463	23.203
	Wca	-36.0467*	3.5881	.000	-50.877	-21.216
	Mi	-64.2879*	9.3622	.000	-96.100	-32.476
	Ra	-24.2550*	4.2066	.002	-40.343	-8.167
	Hi	-22.8300*	3.5700	.003	-37.689	-7.971
	Lli	-37.3967	18.6276	.772	-108.805	34.012

*. The mean difference is significant at the 0.05 level.

b)

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16351.025	2	8175.512	5.089	.010
Within Groups	83538.675	52	1606.513		
Total	99889.699	54			

Multiple Comparisons

(I) Archipelago	(J) Archipelago	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CA	OMR	31.7759*	6.7552	.000	14.631	48.921
	LI	35.1577*	12.5720	.025	3.671	66.645
OMR	CA	-31.7759*	6.7552	.000	-48.921	-14.631
	LI	3.3818	10.7964	.986	-24.539	31.302
LI	CA	-35.1577*	12.5720	.025	-66.645	-3.671
	OMR	-3.3818	10.7964	.986	-31.302	24.539

*. The mean difference is significant at the 0.05 level.

Table 4-6: ANOVA and post-hoc comparisons of mean juvenile migrant movement between a) populations and b) archipelagos. Significant p-values at <0.05 are highlighted.

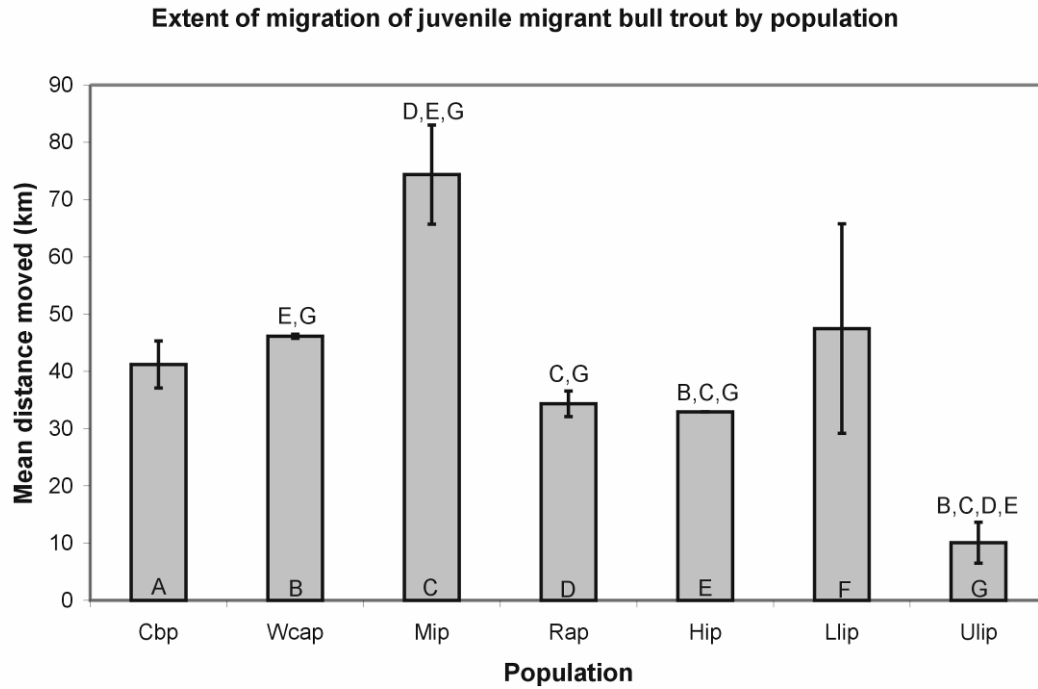**Extent of migration of juvenile migrant bull trout by population**

Figure 4-3: Mean scale of movement of juvenile migrant fish for each population in the study area +/- the standard error. Letters above bars represent post-hoc significant differences.

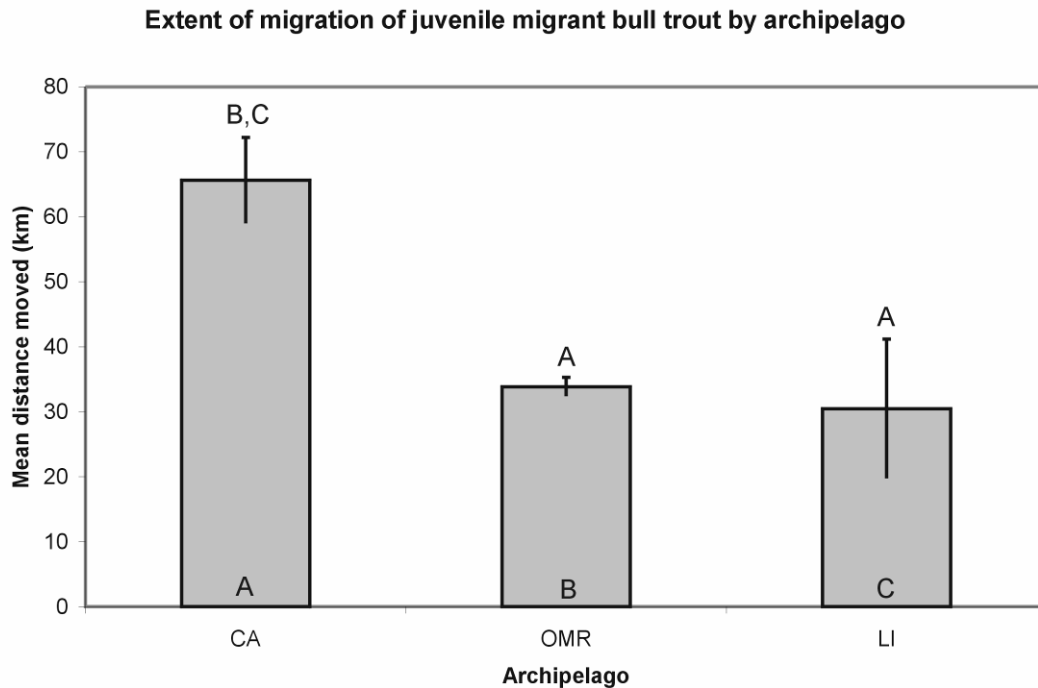**Extent of migration of juvenile migrant bull trout by archipelago**

Figure 4-4: Mean scale of movement of juvenile migrant fish for each archipelago in the study area +/- the standard error. Letters above bars represent post-hoc significant differences.

a)

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28948.873	5	5789.775	11.086	.000
Within Groups	47003.976	90	522.266		
Total	75952.850	95			

Multiple Comparisons[1]

(I) Population	(J) Population	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Cb	Wca	-55.1636	15.8205	.279	-145.775	35.448
	Mi	4.7910	7.4867	1.000	-18.867	28.449
	Ra	23.7664[*]	6.7933	.029	1.537	45.995
	Hi	25.7864[*]	7.0325	.017	3.093	48.480
	Lli	36.4197[*]	10.1189	.044	.662	72.178
Wca	Cb	55.1636	15.8205	.279	-35.448	145.775
	Mi	59.9547	14.7488	.281	-49.771	169.681
	Ra	78.9300	14.4092	.159	-40.987	198.847
	Hi	80.9500	14.5235	.140	-35.231	197.131
	Lli	91.5833[*]	16.2440	.043	3.097	180.070
Mi	Cb	-4.7910	7.4867	1.000	-28.449	18.867
	Wca	-59.9547	14.7488	.281	-169.681	49.771
	Ra	18.9753[*]	3.6589	.000	7.663	30.287
	Hi	20.9953[*]	4.0860	.000	8.506	33.485
	Lli	31.6287	8.3445	.093	-4.162	67.419
Ra	Cb	-23.7664[*]	6.7933	.029	-45.995	-1.537
	Wca	-78.9300	14.4092	.159	-198.847	40.987
	Mi	-18.9753[*]	3.6589	.000	-30.287	-7.663
	Hi	2.0200	2.6062	1.000	-6.709	10.749
	Lli	12.6533	7.7284	.926	-26.181	51.488
Hi	Cb	-25.7864[*]	7.0325	.017	-48.480	-3.093

	Wca	-80.9500	14.5235	.140	-197.131	35.231
	Mi	-20.9953[*]	4.0860	.000	-33.485	-8.506
	Ra	-2.0200	2.6062	1.000	-10.749	6.709
	Lli	10.6333	7.9395	.980	-26.845	48.111
Lli	Cb	-36.4197[*]	10.1189	.044	-72.178	-.662
	Wca	-91.5833[*]	16.2440	.043	-180.070	-3.097
	Mi	-31.6287	8.3445	.093	-67.419	4.162
	Ra	-12.6533	7.7284	.926	-51.488	26.181
	Hi	-10.6333	7.9395	.980	-48.111	26.845

1. ANOVA and post-hoc were performed without the Upper Livingstone population, as only a single migrant was observed in this population
*. The mean difference is significant at the 0.05 level.

b)

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15777.314	2	7888.657	12.192	.000
Within Groups	60175.536	93	647.049		
Total	75952.850	95			

Multiple Comparisons

(I) Archipelago	(J) Archipelago	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CA	OMR	25.5176[*]	3.8976	.000	16.030	35.005
	LI	36.6319[*]	8.3768	.009	10.845	62.419
OMR	CA	-25.5176[*]	3.8976	.000	-35.005	-16.030
	LI	11.1143	7.8097	.504	-15.242	37.470
LI	CA	-36.6319[*]	8.3768	.009	-62.419	-10.845
	OMR	-11.1143	7.8097	.504	-37.470	15.242

*. The mean difference is significant at the 0.05 level.

Table 4-7: ANOVA and post-hoc comparisons of mean adult and subadult movement between a) populations and b) archipelagos. Significant p-values at <0.05 are highlighted.
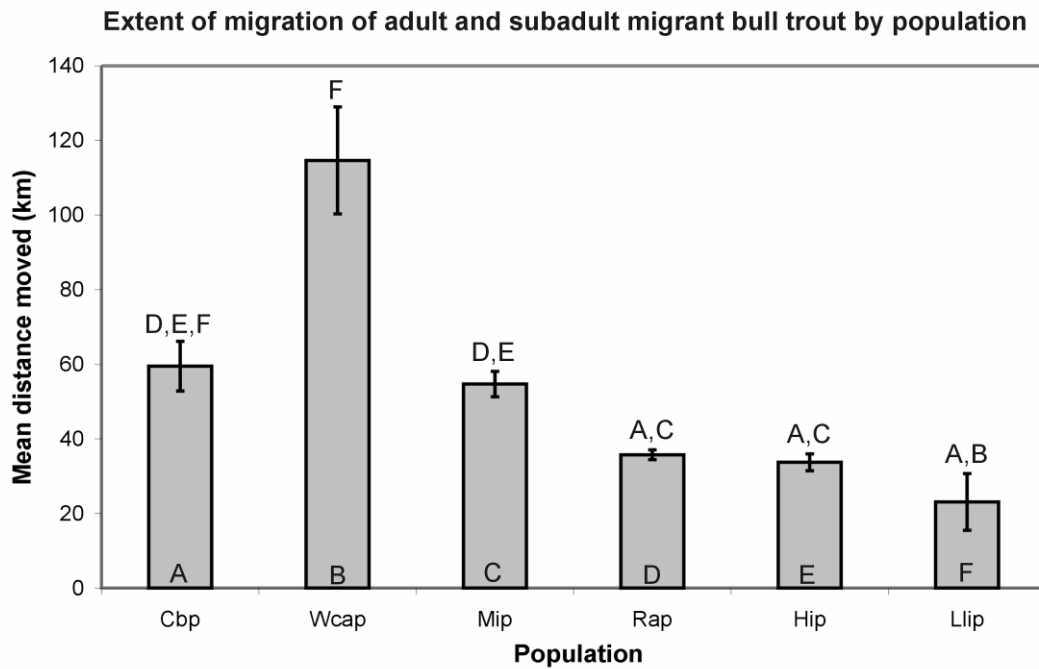
Figure 4-5: Mean scale of movement of adult and subadult migrant fish for each population in the study area +/- the standard error. Letters above bars represent post-hoc significant differences. Note Uli_p was not included in statistical analysis as only as single migrant was observed in this population.
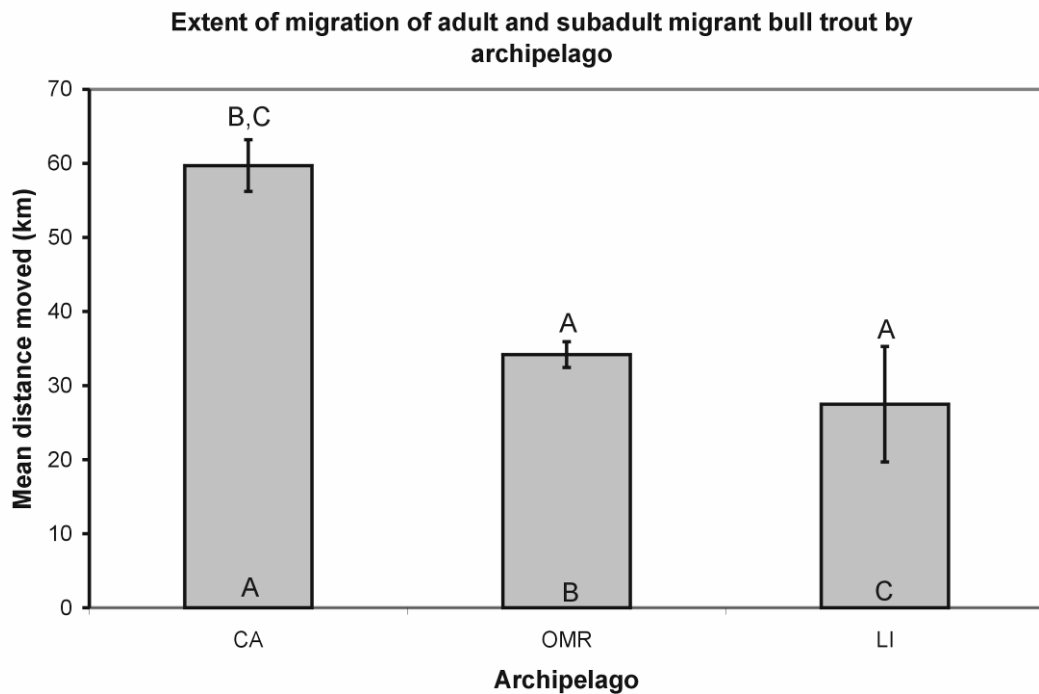

Figure 4-6: Mean scale of movement of adult and subadult migrant fish for each archipelago in the study area +/- the standard error. Letters above bars represent post-hoc significant differences.

a)

	population						
	Cbp	Wcap	Mip	Rap	Hip	Llip	Ulip
mean (km)	6.90	12.77	35.61	7.13	12.04	22.84	3.66
n	17.35	20.77	42.13	16.82	19.29	26.83	29.32

b)

	archipelago		
	CA	OMR	LI
mean (km)	23.49	9.75	12.83
n	80.25	36.11	56.15

c)

	drainage
	total
mean (km)	17.14318
n	172.52

Table 4-8: Mean scale of movement of juvenile fish by a) population, b) archipelago and c) drainage.

Mean scale of movement of juvenile bull trout by population
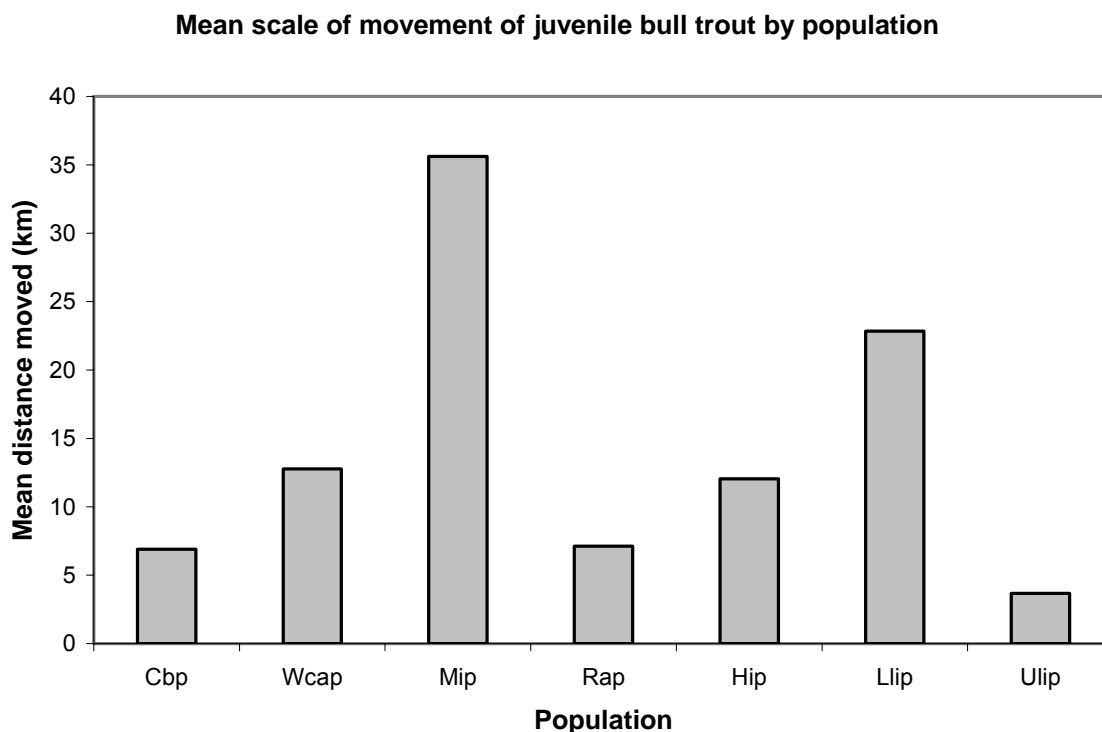

Figure 4-7: Mean scale of movement of juvenile fish for each population in the study area.

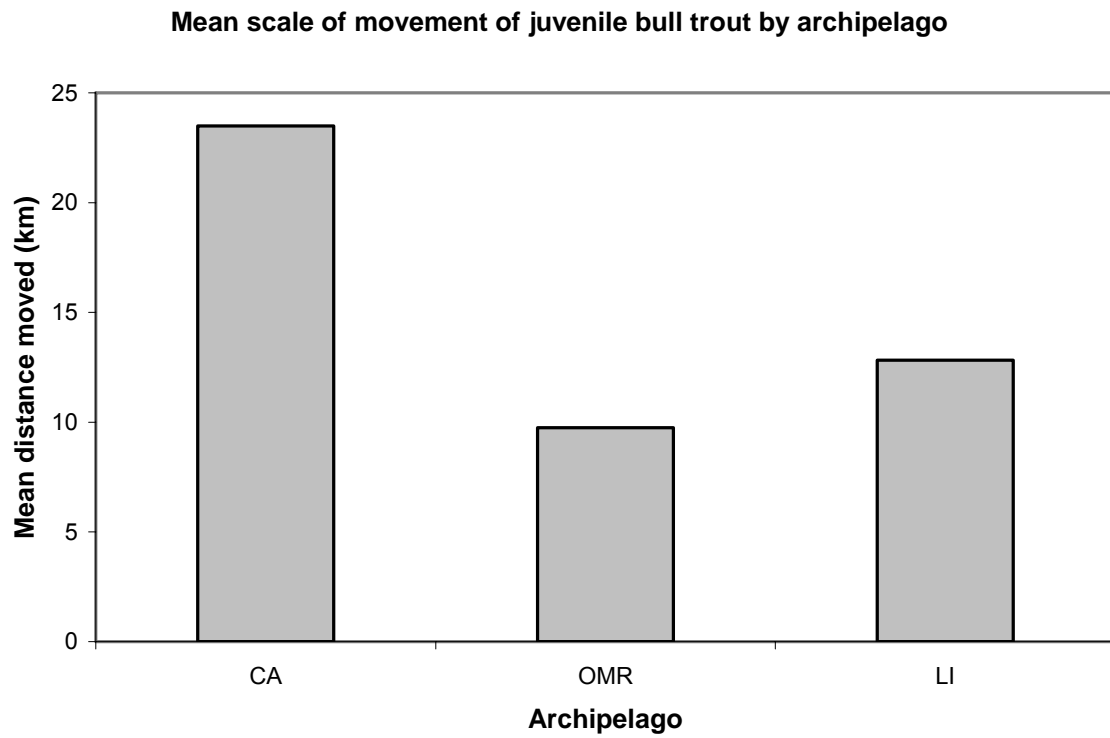**Mean scale of movement of juvenile bull trout by archipelago**

Figure 4-8: Mean scale of movement of juvenile fish for each archipelago in the study area.

Conclusions

Applied conservation efforts are the ultimate goal of theoretical studies in conservation biology. A major theme in this thesis was the exploration of a refined approach to defining populations for such studies, the genetic clustering method. Such an approach defines populations from a genetic rather than a sampling location framework (Waples and Gaggiotti, 1996). This is especially applicable to locally applied conservation studies which tend to define salmonid populations by stream-of-origin (e.g., DeHaan and Arden, 2007; Kassler and Mendel, 2007). Since such studies aim to guide management plans for maintaining genetic variability in threatened or endangered populations of native animals, it is crucial that the methods they use detect true population structure. Genetic clustering methods in this study estimated bull trout population structure as slightly less finely dissected than may be interpreted from traditional F_{ST} based approaches if populations were defined by sampling location. This clustering approach is more likely to detect the true structure because it does not assume *a priori* that discrete sampling sites may be genetically distinguishable, nor does it assume that dispersal is absent between such sites. Furthermore, true population structure in natural populations of organisms, especially those inhabiting river networks, tends to be organized hierarchically (Evanno *et al*., 2005). The clustering method used in this study appears to have merit in revealing differing levels of hierarchical population structure, which may be important considerations when attempting to manage hierarchically structured populations of animals over large spatial scales (Whiteley *et al*., 2006). Since bull trout epitomize such population structure, even at fine spatial scales, such an approach for this species is especially applicable.

When assignment tests from clustering methods are used, inferences of contemporary gene flow and dispersal may be ascertained (Manel *et al*., 2005). Admixed individuals may be identified by these assignment tests, so an investigator may estimate patterns of gene flow between populations based on such measures. Bull trout populations within the study area were found to display such admixture commonly at the finest level of population structure detected, but only infrequently at the coarsest level of population structure. Although such differing rates of gene flow between layers of population structure are to be expected in an hierarchical model (Slatkin and Voelm, 1991), the rates of gene flow between populations within each layer were found to be asymmetrical. These complex patterns of gene flow reveal the arrangement of contemporary connections between populations. Such patterns of connectivity are important in adding to detailed knowledge of local population dynamics for use in conservation efforts (Palstra *et al*., 2007).

Mixed-migrant assignment tests provide an extremely powerful salmonid conservation and management tool, assigning migrant bull trout of unknown origin back to baseline populations. These tests are most successful when populations are defined by clustering methods rather than by *a priori* assuming that each sampling location has its own distinct baseline population. The tandem approach of genetic clustering methods, followed by mixed-migrant assignment analysis may therefore represent a refined approach to performing these tests. The uses of such mixed-migrant assignment tests are many in conservation biology, including assessing population composition of mixed-migrant groups (Koljonen *et al*., 2005), spatio-temporal tendencies of individual migrant populations (Potvin and Bernatchez, 2001), sex-biased dispersal (Hansen, 2001), wildlife

forensics (Primmer *et al*., 2000) or other specialized population questions. Mixed-migrant bull trout groups were found to represent an asymmetrical mosaic of different populations. The populations contributing the most migrants to this group were generally robust. Number of migrants was strongly correlated with intra-population heterozygosity, and may represent a link for predicting effective population sizes from genetic data in future studies.

Variation in contemporary population size, structure and migratory tendencies in salmonid populations can largely be explained by landscape features (Dunham and Rieman, 1999; Castric *et al*., 2001; Costello *et al*., 2003). Work has focused mainly on such landscape features as one-way complete migration barriers (Diener *et al*., 2007), anthropogenic disturbances and occurrence of invasive species (Dunham and Rieman, 1999) and/or physical distances between populations (Castric and Bernatchez, 2003). While all of these factors may contribute to observed population structure in our study, there is cause for additional consideration of spawning ground accessibility. While degrees of accessibility were only qualitatively explored in this study, metrics used to quantify this aspect should be measured in future studies as a possible predictor of size and migratory tendencies in bull trout populations. In most previous studies, populations are considered either isolated (above an impassable migration barrier) or freely accessible to one-way gene flow events (Castric *et al*., 2001; Costello *et al*., 2003). As mentioned in all experimental chapters, many barriers observed in our study area should not be interpreted as complete one-way barriers, but as migration impediments. Such impediments operate on a temporally dynamic degree of navigability as stream flows oscillate and as erosion shapes them over geologic time periods. As such, detailed

descriptive analyses of migration routes, including navigability of specific seasonal barriers, may be warranted in future studies attempting to predict population size, structure and migratory tendencies.

Examination of movement of individual organisms traditionally relied on direct observations; however, assignment tests using highly variable genetic markers have merit as an indirect method of examining such movements (Rannala and Mountain, 1997, Berry et al., 2004). A final advantage of assignment tests used in this study was examined when results were combined with spatial data *a posteriori* in order to quantify the spatial scale of movement of hierarchical populations of bull trout. These quantities may be used for detailed management information of the spatial scale over which populations of bull trout operate, considerations that are especially important when considering the ecological impact of a system's apex predator. Indeed, spatial scale of movement in long-range migrants could be determined for populations of both juvenile and adult bull trout. Juvenile bull trout in particular were found to undergo extreme long-range (>10km) inter-stream migration quite commonly in some populations: movements possibly related to rearing habitat requirements, not life-stage shift (i.e. smolting-like events). This is contrary to long-standing assumptions that rearing juvenile salmonid fish experience restricted movements (Gerking, 1959) and presents further evidence against such a conceptual paradigm (Gowan *et al*., 1994). Evidence thus far for long-range inter-stream movement in juvenile salmonids is sparse (Kennedy *et al*., 2002; Rasmussen *et al*., in press), but represents a large avenue of further research. The underlying mechanisms that drive long-range movements in both juvenile and adult bull trout are poorly understood, but may reflect a conditional dispersal strategy based on habitat or resource availability,

171

or an unconditional strategy to deal with local migratory obstacles (McPeek and Holt, 1992). Nevertheless, scale of migration is variable among populations of bull trout, and exact quantities determined therein may be used for local management purposes.

Genetic clustering methods and assignment tests appear to be powerful, comprehensible tools for use in the field of conservation genetics. The ability not only to reveal hierarchical structure in populations, but to determine patterns of connectivity and scale of migration therein make these tools especially diverse for management and answering a wide realm of ecological questions. Bull trout display a high degree of population structure, and therefore were an excellent model species for exploring the utility of such tools. Results of this study reveal that patterns of population size, structure and migratory trends in this fish are complex, and support the practice of locally applied and hierarchical management practices in order to maintain diversity over large spatial scales.

References

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology, 13*(3), 551-561.

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.

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.

DeHaan, P., & Arden, W. (2007). *Genetic analysis of bull trout in the Jarbidge River watershed, Nevada/Idaho*. Reno, Nevada: U.S. Fish and Wildlife Service.

Deiner, K., Garza, J. C., Coey, R., & Girman, D. J. (2007). Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. *Conservation Genetics, 8*(2), 437-454.

Dunham, J. B., & Rieman, B. E. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications, 9*(2), 642-655.

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

Gerking, S. D. (1959). The restricted movement of fish populations. *Biological Reviews of the Cambridge Philosophical Society, 34*(2), 221-242.

Gowan, C., Young, M. K., Fausch, K. D., & Riley, S. C. (1994). Restricted movement in stream salmonids - a paradigm lost. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(11), 2626-2637.

Hansen, M. M., Kenchington, E., & Nielsen, E. E. (2001). Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries, 2*(2), 93-112.

Kassler, T., & Mendel, G. (2007). *Genetic characterization of bull trout from the Walla Walla River basin*. Olympia, WA: Washington Dept. of Fish and Wildlife.

Kennedy, B. P., Klaue, A., Blum, J. D., Folt, C. L., & Nislow, K. H. (2002). Reconstructing the lives of fish using Sr isotopes in otoliths. *Canadian Journal of Fisheries and Aquatic Sciences, 59*(6), 925-929.

Koljonen, M. L., Pella, J. J., & Masuda, M. (2005). Classical individual assignments versus mixture modeling to estimate stock proportions in Atlantic salmon (*Salmo salar*) catches from DNA microsatellite data. *Canadian Journal of Fisheries and Aquatic Sciences, 62*(9), 2143-2158.

Manel, S., Gaggiotti, O. E., & Waples, R. S. (2005). Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution, 20*(3), 136-142.

McPeek, M. A., & Holt, R. D. (1992). The evolution of dispersal in spatially and temporally varying environments. *American Naturalist, 140*(6), 1010-1027.

Palstra, F. P., O'Connell, M. F., & Ruzzante, D. E. (2007). Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. *Molecular Ecology, 16*(21), 4504-4522.

Potvin, C., & Bernatchez, L. (2001). Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed-stock analyses. *Molecular Ecology, 10*(10), 2375-2388.

Primmer, C. R., Koskinen, M. T., & Piironen, J. (2000). The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. *Proceedings of the Royal Society of London Series B-Biological Sciences, 267*(1453), 1699-1704.

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

Rasmussen, J.B., Trudeau, V., & Morinville, G. (IN PRESS). Estimating the scale of fish feeding movements in rivers using $\delta^{13}C$ signature gradients. *Journal of animal ecology*

Slatkin, M., & Voelm, L. (1991). F-st in a hierarchcial island model. *Genetics, 127*(3), 627-629.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology, 15*(6), 1419-1439.

Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2006). Can common species provide valuable information for conservation? *Molecular Ecology, 15*(10), 2767-2786.