GENOMIC PERSPECTIVES FOR CONSERVATION AND MANAGEMENT OF ATLANTIC COD IN COASTAL LABRADOR

by

Marion Sinclair-Waters

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

at

Dalhousie University
Halifax, Nova Scotia
April 2017

Table of Contents

List of Tables

List of Figures

Abstract

This thesis examines genomic differentiation between the Gilbert Bay and offshore Atlantic Cod populations in Newfoundland and Labrador. First, I determine whether the Gilbert Bay cod population qualifies as a conservation unit. I found evidence of strong genetic divergence between the Gilbert Bay and offshore samples in both neutral and adaptive regions of the genome. These findings suggest that the Gilbert Bay population contributes to the intraspecific diversity of Atlantic Cod and that it warrants consideration as a conservation unit. Second, I develop genomic tools for management and conservation of the Gilbert Bay population. A panel of 23 single nucleotide polymorphisms (SNPs) was developed for identifying Gilbert Bay cod in mixed-stock fisheries. In addition, I estimated the effective population size of the Gilbert Bay population using thousands of genome-wide SNPs. This research demonstrates the power of genomic-based approaches for management and conservation of an exploited marine species.

List of Abbreviations and Symbols Used

AF	Allele frequency
bp	Base pair
CI	Confidence Interval
CIGENE	Centre for Integrative Genetics
COSEWIC	Committee on the Status of Endangered Wildlife in Canada
cM	Centimorgan
DAPC	Discriminant analysis of principal components
DU	Designatable unit
F_{ST}	Measure of genetic differentiation
GO	Gene ontology
GRRF	Guided regularized random forest
H_O	Observed heterozygosity
ID	Identification
K	Number of populations
km	Kilometre
LD	Linkage disequilibrium
LG	Linkage group
m	Slope
mm	Millimetre
Mbp	Megabase pair
MCMC	Markov chain Monte Carlo
MPA	Marine protected area
N_C	Census population size
N_e	Effective population size
$\widehat{N}_e$	Naïve estimate of effective population size
NAFO	Northwest Atlantic Fisheries Organization
ng	Nanogram
PCA	Principal component analysis
r^2	Correlation coefficient
SD	Standard deviation
SNP	Single nucleotide polymorphism
µL	Microlitre

Acknowledgements

I have many people to thank for helping me with this research. First and foremost, I thank my supervisors Dr. Paul Bentzen and Dr. Ian Bradbury for sharing their knowledge and providing guidance throughout this project. Also, thank you for giving me many opportunities to grow as a scientist and a positive introduction to academia. I thank Dr. Christophe Herbinger, Dr. Jeffrey Hutchings and Dr. Daniel Ruzzante for their valuable contributions as members of my ATC and thesis committees.

I thank Dr. Corey Morris for providing feedback, coordinating sample collection and sharing knowledge of Gilbert Bay cod. I thank Fisheries and Oceans Canada and the NunatuKavut Community Council for providing samples. I also extend my thanks to members of the Gilbert Bay MPA Advisory Committee, Dr. John Green, Curtis Pennell and George Rowe for collecting samples and sharing knowledge of the Gilbert Bay area. I also thank Lorraine Hamilton for assisting with DNA extractions and Dr. Robin Waples for providing consultation on estimating effective population size.

I thank all the members of the Bentzen, Bradbury and Ruzzante labs for many valuable discussions, analytical assistance, and help in the lab. I extend my thanks to Ian Paterson who taught me the lab skills needed for this project.

To all my friends in the Dalhousie Biology Department, thank you for making my time here enjoyable. Finally, I thank my mother, father and sister for never doubting me and providing endless support during my master's.

The Natural Sciences and Engineering Research Council of Canada (NSERC), a Nova Scotia Graduate Scholarship and the Genomics Research and Development Initiative of Fisheries and Oceans Canada provided the funding for this research.

Chapter 1 – Introduction

1.1 Introduction

Intraspecific complexity and diversity have been recognized as critically important for the sustainability and persistence of marine fishes (Hilborn et al. 2003; Schindler et al. 2010). Knowledge of the spatial scale of population structuring is therefore fundamental for the effective management and conservation of a species (Hutchings and Reynolds 2004; Conover et al. 2006). Determining conservation and management units for exploited marine species can become contentious without a thorough understanding of a species' genetic structure. Additionally, harvesting in areas where multiple populations intermingle can be detrimental when the populations contributing to the fishery are not known and differ in conservation status. Advances in genomic technologies; however, offer new and more robust methods for identifying conservation and management units and for assigning individuals to these units (Funk et al. 2012).

Atlantic Cod (*Gadus morhua*), a commercially exploited species for hundreds of years (Kurlansky 1997), is a benthopelagic fish that exists throughout the North Atlantic Ocean and is characterized by its high dispersal ability, high fecundity, and large population sizes (COSEWIC 2010). Largely due to its high economic value and ecological importance, the Atlantic Cod is one of the most extensively studied marine fishes. Despite this, knowledge of fine-scale population structuring is still largely lacking for Atlantic Cod in the northwest Atlantic, which limits our ability to effectively manage cod stocks and implement conservation strategies.

In Canadian waters, there are currently six populations of Atlantic Cod identified by COSEWIC (2010). Tagging and oceanographic studies, however, have identified discrete non-migratory stocks and have recognized instances where multiple spawning locations are considered as a single management unit (Green and Wroblewski 2000; Robichaud and Rose 2004; Fox et al. 2008). Such findings highlight the need for an improved understanding of fine-scale genetic structuring and suggest that components contributing to genetic diversity among Atlantic Cod are yet to be described.

A non-migratory and genetically distinct population of Atlantic Cod exists in Gilbert Bay, Labrador, Canada (Green and Wroblewski 2000; Ruzzante et al. 2000a; Beacham et al. 2002; Hardie et al. 2006; Bradbury et al. 2010, 2013). Gilbert Bay is a narrow inlet with an area of ~60km^2 located in a remote region on the southeast coast of Labrador and was designated a marine protected area (MPA) in 2005 under Canada's *Ocean Act* to protect the resident population of Atlantic Cod and its habitat (DFO 2007). The Gilbert Bay population is currently managed as part of a large management unit (NAFO 2J3KL) that encompasses the Northern cod complex and belongs to the Newfoundland and Labrador Designatable Unit defined by COSEWIC (COSEWIC 2010). Trends reflecting the abundance of Gilbert Bay cod suggest the population has declined considerably since the MPA was created (Morris and Green 2014; Morris and Green in-press). A factor contributing to these declines is that some Gilbert Bay cod move outside MPA boundaries during the summer months and are harvested in fisheries that occur in areas adjacent to the MPA (Morris et al. 2003).

The overall objective of this thesis is to develop genomic tools that directly inform the management and conservation of the Gilbert Bay cod population. Specifically,

I will use genomic tools to explore the degree of genomic differentiation between the Gilbert Bay and offshore populations, to monitor the effective population size (N_e) of the Gilbert Bay population, and to identify Gilbert Bay cod in fisheries that occur outside MPA boundaries. Additionally, I use a genomic framework to examine how the Gilbert Bay population may be adapted to the environmental conditions unique to Gilbert Bay and whether the population warrants consideration as its own conservation unit.

1.2 Thesis structure

This thesis is structured into two data chapters (Chapters 2 and 3). Both data chapters provide a genomic perspective for management and conservation of Atlantic Cod in the Gilbert Bay MPA.

Chapter 2 examines the genetic differentiation observed between Gilbert Bay cod and offshore cod in Newfoundland and Labrador at both neutral and adaptive regions that span the entire genome. I also evaluate the potential for local adaptation of Gilbert Bay cod using a genomic approach that incorporates information from an annotated Atlantic Cod genome. These findings are taken into consideration to determine whether the Gilbert Bay cod population warrants designation as its own conservation unit.

Chapter 3 is aimed at developing genomic tools for management and conservation of Gilbert Bay cod. I develop a cost-effective and informative SNP-panel that can be used to distinguish Gilbert Bay cod from offshore cod in mixed-stock fisheries that occur adjacent to the Gilbert Bay MPA boundaries. Using this new SNP-panel, I look at the proportion of Gilbert Bay cod in fishery samples collected at sites outside MPA boundaries. I also develop a framework for estimating N_e of the Gilbert Bay cod population with a large SNP dataset.

In Chapter 4, I will discuss the general implications of the findings presented in this thesis and potential directions for future research.

Chapter 2 – Genome-wide SNP analysis reveals fine-scale population structure and defines conservation units in coastal Labrador Atlantic Cod

2.1 Abstract

The identification of intraspecific diversity is central to the management and conservation of exploited species, but knowledge of diversity and fine-scale population structure is currently lacking for many marine species. Recent advances in methods for genomic analysis allow genome-wide surveys of intraspecific diversity and offer new opportunities for exploring spatial structure. Here, I analyzed genome-wide polymorphisms to measure population differentiation and help define conservation units of Atlantic Cod in coastal Labrador. A total of 141 individuals were collected from offshore sites and from a coastal site within Gilbert Bay, Labrador. All individuals were genotyped for ~11k SNPs using a cod-specific SNP array. Both discriminant analysis of principal components and Bayesian clustering reveal strong genetic differentiation between offshore and Gilbert Bay cod. In total, 106 highly divergent loci were identified using outlier tests and an F_{ST} threshold. Two chromosomal inversions were detected on linkage groups 1 and 12, which coincide with inversions previously found in Atlantic Cod. Gene annotations of highly divergent SNPs demonstrate that local adaptation may play a role in the divergence between the offshore and Gilbert Bay populations. This work strongly supports the designation of the Gilbert Bay population as its own conservation unit and demonstrates the power of using a genomic approach for defining conservation units.

2.2 Introduction

Describing intraspecific genetic diversity is an important component of effective management and conservation of exploited marine species. Such diversity, particularly when associated with fitness-related traits, is expected to promote sustainability and persistence of marine species (Hilborn et al. 2003; Schindler et al. 2010). Despite the importance of recognizing patterns of intraspecific diversity, knowledge of fine-scale population structuring is often lacking. Due to a recent increase in available genomic tools for non-model organisms, we now have an opportunity to define population structure in marine species with unprecedented resolution (Allendorf et al. 2010; Lamichhaney et al. 2012; Bourret et al. 2013a; Bradbury et al. 2013; Hemmer-Hansen et al. 2014; Milano et al. 2014; Benestan et al. 2015; Van Wyngaarden et al. 2016). In addition, advances in genomic technology improve our ability to characterize the evolutionary processes influencing intraspecific diversity (Bradbury et al. 2010; Funk et al. 2012; Bourret et al. 2013b; Hess et al. 2013; Moore et al. 2014; Aykanat et al. 2015; Lemay and Russello 2015). Combining knowledge of population structure and the evolutionary significance of such structure allows us to delineate conservation units, which are generally defined as populations that are reproductively isolated and represent a significant evolutionary component of the species (Waples 1991, 1995; USFWS 1996; COSEWIC 2015). Considering units below the species-level helps to develop management and conservation strategies that will maximize the evolutionary potential of marine species in the face of environmental change.

2.2.1 Defining conservation units

To define species-specific conservation units, a comprehensive examination of both neutral and adaptive regions of the genome is required. Examining neutral differentiation provides insight into how populations are shaped by neutral processes such as historical isolation (produces molecular differences) and gene flow (maintains molecular similarities). The designation of a conservation unit is generally only considered when neutral genetic differentiation exists. On the other hand, it is also important to examine adaptive differentiation when defining conservation units because adaptive processes such as divergent selection can also play a role in shaping populations (Funk et al. 2012). Genome-scale data provide knowledge of neutral genetic differentiation and can also be used to identify loci or genomic regions that are potentially under selection. Furthermore, annotated genomes, available for an increasing number of marine species (e.g. Star et al. 2011; McGowen et al. 2012; Martinez Barrio et al. 2016; Lien et al. 2016), allow us to examine the functional significance of these candidate loci, which begins to reveal how populations may be adapted to their local environments. Such local adaptations, which can contribute to future evolutionary potential of a species, may warrant the designation of a conservation unit (Waples 1991; Crandall et al. 2000). Due to evolutionary inferences based on both neutral and adaptive divergence that genome-wide approaches offer (Kelley et al. 2016), the use of genomics for delineating conservation units could make a valuable contribution to the management and conservation of marine species.

2.2.2 Patterns of population structure in Atlantic Cod of the Northwest Atlantic

Atlantic Cod, a commercially exploited species for hundreds of years (Kurlansky 1997), is a benthopelagic fish that inhabits the continental shelves throughout the temperate regions of the North Atlantic Ocean. It is characterized by its high dispersal potential, high fecundity, and large population sizes (COSEWIC 2010). In the northwest Atlantic, studies using microsatellite loci revealed weak, yet significant, population structure over large spatial scales (Bentzen et al. 1996; Ruzzante et al. 1996; Beacham et al. 2002; O'Leary et al. 2007). This pattern has been supported by more recent studies using SNPs that include adaptive loci in analyses of population structure (Bradbury et al. 2010, 2013). These SNP-based approaches also indicate that adaptive variation in Atlantic Cod is not randomly distributed across the genome. Large genomic islands of divergence, in at least some cases associated with chromosomal inversions (Berg et al. 2016; Kirubakaran et al. 2016; Sodeland et al. 2016), show adaptive clines separating cod into southern and northern groups on both sides of the Atlantic Ocean (Bradbury et al. 2013). This evidence of broad-scale population structure has been incorporated into the designation of conservation units and status assessments in the Northwest Atlantic (COSEWIC 2010). Evidence of fine-scale population structure, however, is largely lacking and therefore the consideration of conservation units at small spatial scales has been limited.

2.2.3 A candidate population: The Gilbert Bay Atlantic Cod population

One exception to this pattern of low spatial differentiation, however, occurs in southern Labrador. Here, a coastal population is found in Gilbert Bay, a small (~60 km^2) semi-enclosed inlet. The Gilbert Bay population is a non-migratory (Green and

Wroblewski 2000) and genetically distinct population of Atlantic Cod (Ruzzante et al. 2000a; Beacham et al. 2002; Hardie et al. 2006; Bradbury et al. 2010, 2013). To enable conservation of Gilbert Bay cod and their habitat, Gilbert Bay was designated a MPA under Canada's *Ocean Act* in 2005. Fishing pressure, however, continued to affect the Gilbert Bay population because the home range of Gilbert Bay cod includes an area (~270 km^2) that is outside MPA boundaries. The movement of Gilbert Bay cod outside the MPA occurs during late summer (Morris et al. 2014), which can overlap with the timing of commercial fishing in areas adjacent to the MPA. Despite efforts to protect the Gilbert Bay cod population, decreases in abundance and recruitment have continued (Morris and Green 2014; Morris and Green 2017 in-press). Given its unique characteristics and declining state, consideration of the Gilbert Bay cod population as its own conservation unit is warranted.

2.2.4 Aims of study

In this study, I examine the genomic differentiation between Gilbert Bay cod and offshore cod in coastal Labrador using data from thousands of SNPs distributed across all chromosomes. I combine genomic data for Atlantic Cod with linkage map information to evaluate whether the Gilbert Bay population qualifies as a separate conservation unit. I describe the nature and extent of genetic differentiation through the application of outlier tests, chromosomal inversion detection and gene annotations. Here, I demonstrate the value of considering genomic data when defining conservation units in economically important marine species.

2.3 Methods

2.3.1 Sample collection, DNA extraction and genotyping

Atlantic Cod were sampled from four offshore sites located in marine waters off of Newfoundland and Labrador and one coastal site located in the most inner region of Gilbert Bay, a narrow inlet on the southeast coast of Labrador. From 2010 to 2015, 78 individuals were collected from the offshore sites over four sampling periods that occurred on 5 December 2010, 7 June 2011, 8 October 2015 and 10 October 2015. Another 63 individuals were collected from the coastal site during two sampling periods: 9 September 2012, and 1-10 June 2015 (Figure 1, Table 1). Fin clips were collected and immediately preserved in 95% ethanol. Genomic DNA was extracted from a 2 mm^3 tissue sample taken from the preserved fin clips over two separate batches. The first batch was extracted using a Qiagen DNeasy 96 blood and tissue kit following protocol described by manufacturer and the second batch of DNA was extracted using a standard phenol-chloroform procedure (Sambrook et al. 1987). All DNA was quantified using QuantIT PicoGreen (Life Technologies) and was normalized to a final concentration of 50ng/μL prior to genotyping. Genotyping of all individuals took place at the Centre for Integrative Genetics (CIGENE), Norwegian University of Life Sciences in As, Norway using a cod-derived Illumina SNP-chip. The SNP-chip was developed using genomes of seven Atlantic Cod from across the Northeast Atlantic that were sequenced using a shotgun approach (Kent et al. in prep.). Reads were aligned to the gadMor1 reference genome (Star et al. 2011) and 2,877,794 putative SNPs were identified. Of these, 10,913 SNPs were chosen for the array based on their physical distribution and functional associations. This includes 260 SNPs from previous studies (Hubert et al. 2010; Moen et

al. 2008), 672 SNPs in close proximity to candidate genes, and 1595 non-synonymous coding SNPs. On average, the array includes 409 SNPs per chromosome (Kent et al. in prep.).

2.3.2 Quality control filters and population genetic statistics

To ensure optimal data quality, I filtered out markers and individuals that failed particular quality thresholds. Any locus that did not classify as a bi-allelic SNP based on a clustering pattern determined using a sample of more than 5000 individuals (Kent et al. in prep.), was removed prior to the following filtering steps. PLINK, a tool set for manipulating and analyzing large genomic datasets, was used for the following filtering procedures (Purcell et al. 2007). Any individual with low genotyping success (less than 85% complete) was removed from the dataset. SNPs with a minor allele frequency lower than 0.01 were removed. SNPs were also filtered to ensure missing data at any given loci were not greater than 15%. The order of filtered SNPs on chromosomes was determined using the linkage map for Atlantic Cod (S. Lien, Centre for Integrative Genetics, Ås, Norway, personal communication).

Global and per locus pairwise genetic differentiation, Weir and Cockerham's F_{ST} (1984), between the two putative populations, Gilbert Bay and offshore, were calculated using the *R* package *diveRsity* (Keenan et al. 2013). Observed heterozygosity (H_o) for each locus was also calculated using *diveRsity*. For both the offshore group and coastal group, allele frequencies were calculated per locus using the *genepop_allelefreq* function in the *R* package *genepopedit* (Stanley et al. 2016).

2.3.3 Analysis of population structure

Population structure was examined and visualized using two methods. First, the Bayesian clustering program STRUCTURE 2.3.4 was used to determine population structure with the filtered SNP dataset (Pritchard et al. 2000). For each possible value of K (1-4), three replicate runs consisting of a burn-in period of 100,000 and run length of 500,000 iterations were performed. An approach implemented in CLUMPAK was used to visualize and determine the best number of populations (or clusters (K)) for the samples (Kopelman et al. 2015). The likelihood of each value of K was estimated using the DeltaK statistic of Evanno et al. (2005) to determine the number of populations present. The second method used to examine population structure was a discriminant analysis of principal components (DAPC) as implemented in the *R* package *adegenet* (Jombart et al. 2010). A DAPC uses no prior knowledge of population structure and estimates population structure by maximizing differences between clusters while minimizing differences within clusters. The optimal number of populations was determined using the function *find.clusters,* which uses a Bayesian information criterion method.

2.3.4 Identifying highly divergent loci

Loci that displayed elevated divergence between the two putative populations, Gilbert Bay and offshore, were identified. I used two different approaches for identifying highly divergent loci. First, I selected any SNP that had a value of F_{ST} greater than 0.3 as calculated using *diveRsity.* Second, I used three genome scan methods for detecting SNPs with greater than expected levels of divergence (i.e. outlier loci). To reduce the possibility of false positives, only the SNPs that were identified as outliers using all three

methods were considered true outliers. Loci that were not identified as outliers by any of the three methods were classified as neutral in further analyses.

The first method for outlier detection was a Bayesian approach as implemented in the program *BayeScan v.2.1* (Foll and Gaggiotti 2008). Following a burn-in period of 50,000, 100,000 iterations were run. The prior odds were set at 100, which makes the neutral model 100 times more likely than the model that includes selection. Any SNP with a posterior probability over 0.95 was considered an outlier. The second method used the *Fdist* approach (Beaumont and Nichols 1996) as implemented in the program *Arlequin v3.5*. In total, 200,000 coalescent simulations were performed and loci with a *P*-value less than or equal to 0.01 were considered outliers. The final method was based on a principal component analysis (PCA) as implemented in the *R* package *pcadapt* (Luu et al. 2017). The statistical method used in *pcadapt* identifies loci that are significantly associated with population structure as outliers. I used the default method that computed Mahalanobis distance and set K=1 to reflect population structure. Any SNP with a *P*-value less than 0.05 was considered an outlier.

2.3.5 Inversion detection

The *R* package *inveRsion* (Cáceres et al. 2012) was used to detect large chromosomal inversions. Only SNPs with known order in the genome based on the linkage map were included in the analysis. A block size of 3 SNPs was used to determine haplotypes for each candidate breakpoint. Next, a sliding window size of 100 SNPs was used to scan each chromosome and a Bayesian information criterion threshold (*thbic)* of 0 was used when identifying inverted regions of the genome. To further investigate the possibility of chromosomal inversions, patterns of linkage disequilibrium (LD) within

each linkage group (LG) were examined. High levels of LD between loci pairs in inverted regions would be expected due to a reduced rate of recombination (Feder et al. 2014). The degree of LD between pairs of SNPs on each LG was calculated for both the offshore and Gilbert Bay groups using PLINK (Purcell et al. 2007).

2.3.6 SNP annotation and enrichment analysis

Gene annotations for the outlier SNPs were obtained from Berg et al. (2016). Gene annotations for SNPs that did not overlap with the Berg et al. (2016) study were obtained from *Ensembl*. A 201 base pair (bp) sequence surrounding the SNP was aligned to the gadMor1 genome using BLAT and parameters that required a 201 bp alignment length and >95% identify. Gene ontologies (GOs) associated with outlier SNPs located within genes or less than 5000 bp from closest gene were retrieved using *g:Profiler* (Reimand et al. 2016). Functional enrichment analysis was performed using *g:Profiler* to determine whether an over-representation of a biological process, molecular function or cellular component existed. The Benjamini-Hochberg False Discovery Rate was used to correct *P*-values for multiple testing and only *P*-values < 0.05 were considered significant.

2.4 Results

2.4.1 Quality control filters and population genetic statistics

Genotypes for 141 individuals were obtained. Of the 10,913 genotyped loci, 8581 were categorized as bi-allelic SNPs and included in the analyses for this study. The overall genotyping rate was 0.98 and no individuals were filtered out due to low genotyping success (<0.85). The filtering steps that focused on SNP quality removed 257 SNPs with low genotyping success (<0.85) and an additional 793 SNPs with a minor

allele frequency lower than 0.01. In total, 7,531 SNPs were selected from the original 10,913 genotyped SNPs for further analyses. The filtered dataset consisted of 7,318 SNPs that could be ordered according to the linkage map. There was an average of 318 SNPs per chromosome with a minimum of 224 SNPs (LG19) and a maximum of 405 SNPs (LG4) per chromosome (Table S1).

Per locus F_{ST} between Gilbert Bay and offshore ranged between 0 and 0.65 for the filtered dataset (Figure 2). Average F_{ST} per chromosome ranged from 0.04 (LG17) to 0.11 (LG1) (Table S1, Figure S1). Global F_{ST} calculated between groups, offshore and Gilbert Bay was 0.06. Average H_o across all filtered loci was 0.31 and 0.27 for the offshore group and Gilbert Bay group, respectively (Figure 2). Private alleles were not found in either the offshore or Gilbert Bay group. Alleles that were fixed or close to fixation (>0.95) were more abundant in the Gilbert Bay group than the offshore group (Figure 2).

2.4.2 Analysis of population structure

Strong genetic structuring between Gilbert Bay and the offshore population was shown by both the STRUCTURE and DAPC analyses using all 7,531 filtered SNPs (Figure 3). The most likely value of K calculated by CLUMPAK was two (Figure S2, Figure S3). Very little evidence of admixture was observed with >80% of all individual genotypes associated with one of the two populations. For the DAPC analysis, the lowest BIC value (986.07), and therefore the most probable number of populations, corresponded to K=2. A total of 40 PCA axes and one discriminate function were retained for the analysis of population structure.

2.4.3 Identifying highly divergent loci

The three methods of outlier detection identified a total of 249 outlier loci, of which 35 were detected by all methods. Global F_{ST} of the 35 consensus outlier loci was 0.47, whereas global F_{ST} of neutral loci (7282 SNPs) was 0.05. Using *Arlequin,* 241 SNPs were detected as outliers (P-value ≤ 0.01). *pcadapt* detected 60 outlier SNPs (P-value <0.05), of which 59 were also identified using *Arlequin*. *BayeScan* detected 37 outliers (posterior probability > 0.95), all of which were detected using *Arlequin* (Figure 4). In total, 106 loci were above the F_{ST} threshold of 0.3 and therefore flagged as highly divergent for further analyses. All 35 consensus outlier loci identified by all three outlier detection methods had F_{ST} values greater than 0.3. Over half (51%) of the 106 flagged SNPs were found on LG1. Furthermore, 34 of 35 consensus outlier loci were located on LG1 (Figure 5, Table S2).

2.4.4 Inversion detection and linkage disequilibrium patterns

Using the *R* package *inveRsion*, analyses of the 7,318 SNPs with a known order in the genome revealed two candidate chromosomal inversions. The first inverted region occurs on LG1 and spans 18.19 Mbp. The inverted allele is observed in the offshore population only. Of the 106 highly divergent SNPs, 48 were located within this inverted region on LG1. The second inverted region was detected on LG12 and spans 10.97 Mbp. The inverted allele is observed in both populations, however, more frequently in the offshore population (Table 2). The chromosomal inversions found here are supported by patterns of LD observed within each linkage group. The r^2 value indicating the level of LD between each pair of SNPs was obtained for each linkage group. Regions displaying high r^2 values are observed on LG1 in the offshore population and on LG12 in both

populations (Figure 6). No other linkage groups show LD patterns suggestive of chromosomal inversions (Figure S4).

2.4.5 SNP annotation and enrichment analysis

Annotations for 86 of the 106 highly divergent SNPs were retrieved. Of these, 74 SNPs were located within the coding region of a gene or less than 5000 bp away from the closest gene (Table S2). Furthermore, 20 SNPs are known nonsynonymous mutations. The enrichment analysis did not identify a biological process, molecular function or cellular component that was significantly (P-value < 0.05) over-represented among these gene-associated loci (Table S3).

2.5 Discussion

Here, I used a genomic approach to provide insight into the evolution of two populations of Atlantic Cod in coastal Labrador. Through analysis of 7531 SNPs distributed across the genome, I show evidence of strong genomic differentiation between the Gilbert Bay and offshore cod population. Differentiation was observed at both neutral and outlier loci suggesting that both adaptive and neutral evolutionary processes contribute to the divergence of these two populations. Tests aimed at detecting chromosomal inversions found inverted regions on LG1 and LG12. The inversion on LG1 includes many potentially adaptive loci and may help facilitate adaptive divergence. The degree of both adaptive and neutral differentiation revealed here suggests that the Gilbert Bay and offshore populations should be considered as separate conservation units.

2.5.1 Gilbert Bay population as its own conservation unit

I evaluated whether the Gilbert Bay population is its own conservation unit by considering both adaptive and neutral information through the analysis of genome-wide

SNPs. The global F_{ST} for neutral loci observed between the Gilbert Bay and offshore cod populations is the highest of any population comparison in Canadian waters, excluding the land-locked Arctic Lake populations (Bentzen et al. 1996; Ruzzante et al. 1997, 1998, 2000; Lage et al. 2004; Hardie et al. 2006). Genetic diversity is not substantially lower in the small Gilbert Bay population than it is in the larger offshore population. Therefore, the high degree of neutral differentiation can be attributed to low levels of gene flow and suggests that the two populations are reproductively isolated. This is supported by the fact that Gilbert Bay cod spawn later than offshore cod (May 1966). Additionally, shallow sills and a low-surface salinity help retain cod eggs and early stage larvae in the bay (Morris and Green 2002).

In addition to showing high neutral differentiation, a large proportion (70%) of loci displaying high levels of divergence between the Gilbert Bay and offshore population were located within or in close proximity to genes. These genes associated with elevated differentiation between the two populations indicate potential for local adaptations and also serve as a good starting point for characterizing the genetic basis of such adaptations (Lotterhos and Whitlock 2015; Hoban et al. 2016). The combination of neutral and adaptive divergence observed in the Gilbert Bay population supports the designation of this population as its own conservation unit. Currently, Atlantic Cod in Newfoundland and Labrador are managed as a single conservation unit (COSEWIC 2010) and therefore these findings should be considered in future conservation efforts.

2.5.2 Potential targets of selection

A set of loci displaying high divergence between the Gilbert Bay and offshore populations was identified using outlier detection tests and an inclusion threshold based

on F_{ST}. An F_{ST} threshold was used due to the limitations associated with using outlier tests on only two populations and when neutral differentiation is high. When using *BayeScan* for outlier detection, at least six populations are recommended for detecting divergent selection and at least 10 for detecting balancing selection. Power for detecting outliers was low in scenarios where only two populations were considered (Foll and Gaggiotti 2008). In addition, F_{ST} based outlier tests (e.g. *Arlequin* and *BayeScan*) identify loci that are distinguishable from patterns of neutral differentiation (Beaumont and Nichols 1996; Foll and Gaggiotti 2008) and are therefore problematic when neutral differentiation is high. In such cases, the proportion of false negatives likely increases because the variance of the distribution for neutral loci is large and thus only loci with extremely high F_{ST} values will fall in the tail of the distribution and be identified as an outlier (Hoban et al. 2016).

Methods for identifying adaptive loci have received substantial criticism (Narum and Hess 2011; Vilas et al. 2012; Lotterhos and Whitlock 2014, 2015; Hoban et al. 2016) and thus the interpretation of highly divergent loci and their role in local adaptation should be made with caution. Nevertheless, I demonstrate how outlier tests and F_{ST} thresholds remain useful tools for identifying potential targets of selection when loci with remarkably high levels of differentiation exist and gene annotations are available. Most of the highly divergent SNPs identified in this study were located within or close to genes, including some nonsynonymous substitutions. Furthermore, we do not observe a substantial drop in H_O for the Gilbert Bay population, which is characteristic of populations that have experienced a bottleneck event (Nei et al. 1975; Chakraborty and Nei 1977). This provides further evidence that the genetic differences observed between

Gilbert Bay cod and offshore cod have more to do with selective pressures than a bottleneck event.

The most seemingly informative gene annotations were associated with SNPs found within the inverted region on LG1. One SNP (Gdist:446136_208) is located in a gene that encodes a cold shock domain-containing protein E1 (CSDE1), which is an RNA-binding protein involved in regulating transcription. Cold shock proteins are thought to improve cell survival at lower than optimal growth temperatures (Obokata et al. 1991). A nonsynonymous SNP (NS:129362_255) in LG1 is located in a gene that encodes a solute carrier family membrane transport protein (SLC35C2). A nonsynonymous SNP (NS:51949_161) on LG7 is in the *SLC7A1* gene that also encodes a membrane transport protein, but from a different solute carrier family. In Threespine Sticklebacks (*Gasterosteus aculeatus*), variation at genes encoding solute carrier proteins has been associated with annual salinity variation in the Baltic Sea (Guo et al. 2015) and promoting divergence between freshwater and marine habitats in Alaska (Hohenlohe et al. 2010). Because Gilbert Bay cod spend their lifetime in or near the bay (Morris et al. 2014), they would likely experience different salinities and temperatures than offshore cod. Given these differences, the *CSDE1*, *SLC35C2* and *SLC7A1* loci may play key roles in promoting adaptive differentiation between the Gilbert Bay and offshore populations.

Although the gene-annotation results are plausible, we cannot conclude that these genes are the true targets of selection. It is possible that they are merely linked to another locus that is the actual target of selection. The likelihood of identifying the true target of selection is particularly difficult within chromosomal inversions where LD is strong (Hoffmann et al. 2004), causing loci to act in unison with other loci in the region.

Identifying true targets of selection would require follow-up research such as genome resequencing and an investigation of the functional effects of variation in gene regions (Stinchcombe and Hoekstra 2007; Barrett and Hoekstra 2011; Savolainen et al. 2013). Additionally, the approach used in this study is not aimed at identifying targets of selection characterized by polygenic adaptation (i.e. subtle allelic changes at many loci) (Pritchard and Di Rienzo 2010) and may limit our ability to fully understand the role selection plays in driving divergence between the Gilbert Bay and offshore populations. Nevertheless, these findings provide a valuable starting point for more detailed work on identifying targets of selection.

2.5.3 Chromosomal inversions and population differentiation

Large chromosomal inversions have been previously reported in Atlantic Cod and their role in promoting divergence between populations has been recognized (Berg et al. 2016; Kirubakaran et al. 2016; Sodeland et al. 2016). The two inversions detected here, on LG1 and LG12, have been previously found in populations of Atlantic Cod in the northeast Atlantic. The inversion on LG1 was associated with divergence between the migratory North East Arctic cod and non-migratory Norwegian coastal cod ecotypes, where the inverted allele is found predominantly in the migratory ecotype (Berg et al. 2016). A similar pattern is seen in this study where the inverted allele is present in the migratory offshore population, but is absent in the non-migratory Gilbert Bay population. This suggests that the inversion may contain genes associated with migratory behaviours. Chromosomal inversions can play a key role in adaptive divergence by suppressing recombination between co-adaptive genes and thus ensuring the co-inheritance of multiple favourable alleles (Hoffmann and Rieseberg 2008; Feder et al. 2014).

The genomic locations of the inversions found in this study were consistent with previous work, however, the size of the inversions differed by ~0.79 Mbp (LG1) and ~2.03 Mbp (LG12) (LG1: Kirubakaran et al. 2016; LG12: Sodeland et al. 2016). The detection of candidate breakpoints reflects the density of SNP coverage available. Here, SNP array data was used and therefore the proposed breakpoints are approximate and reflect only SNPs that are polymorphic in this study. Determining more precise locations of breakpoints would require a higher-density SNP-array or full genome sequencing.

2.5.4 Conclusion

This study provides evidence for a new conservation unit of Atlantic Cod in coastal Labrador, and in a more general context it demonstrates the power of using a genomic approach for delineating conservation units in marine species. Genome-scale data provide an opportunity to define population structure with unprecedented resolution over fine spatial scales. The inclusion of non-neutral loci allows us to examine patterns of adaptive differentiation that help illustrate how particular populations may contribute to the evolutionary potential of a species. Furthermore, the analysis of thousands of markers distributed across all chromosomes can provide insight into the mechanisms that promote and maintain adaptive differentiation. In conclusion, the use of genomics improves our ability to delineate conservation units and thus helps inform management and conservation efforts for marine species.

Table 1. Site locations, date of collection and sample size for all Atlantic Cod samples included in Chapter 2.

	Location	Latitude	Longitude	Sample ID	Sampling time	Sample size
Offshore	3L	48.287	-49.247	N00	07-Jun-11	21
	3K	50.322	-50.762	T00	05-Dec-10	20
	2J	52.482	-53.810	2JA	10-Oct-15	17
	2J	52.633	-54.817	2JB	08-Oct-15	20
Gilbert Bay	Shinney's	52.585	-56.033	SMP	09-Sep-12	20
	Shinney's	52.585	-56.033	GBJ	09-Sep-12	19
	Shinney's	52.585	-56.033	GBM	1-Jun to 10-Jun-15	24

Table 2. Chromosomal inversions detected using *inveRsion* and their corresponding linkage group, breakpoints and genotype frequencies for each population. The left (min.) and right (max.) breakpoints are reported. Genotypes are as listed: AA=homozygous for the reference orientation, AB=heterozygous for the inversion, BB=homozygous for the inversion.

		Inversion frequencies					
	Inversion breakpoints	Offshore			Gilbert Bay		
LG	(bp position)	AA	AB	BB	AA	AB	BB
1	9,334,923-27,525,445	0.23	0.41	0.36	1.00	0	0
12	2,734,575-13,700,971	0.68	0.32	0	0.89	0.11	0

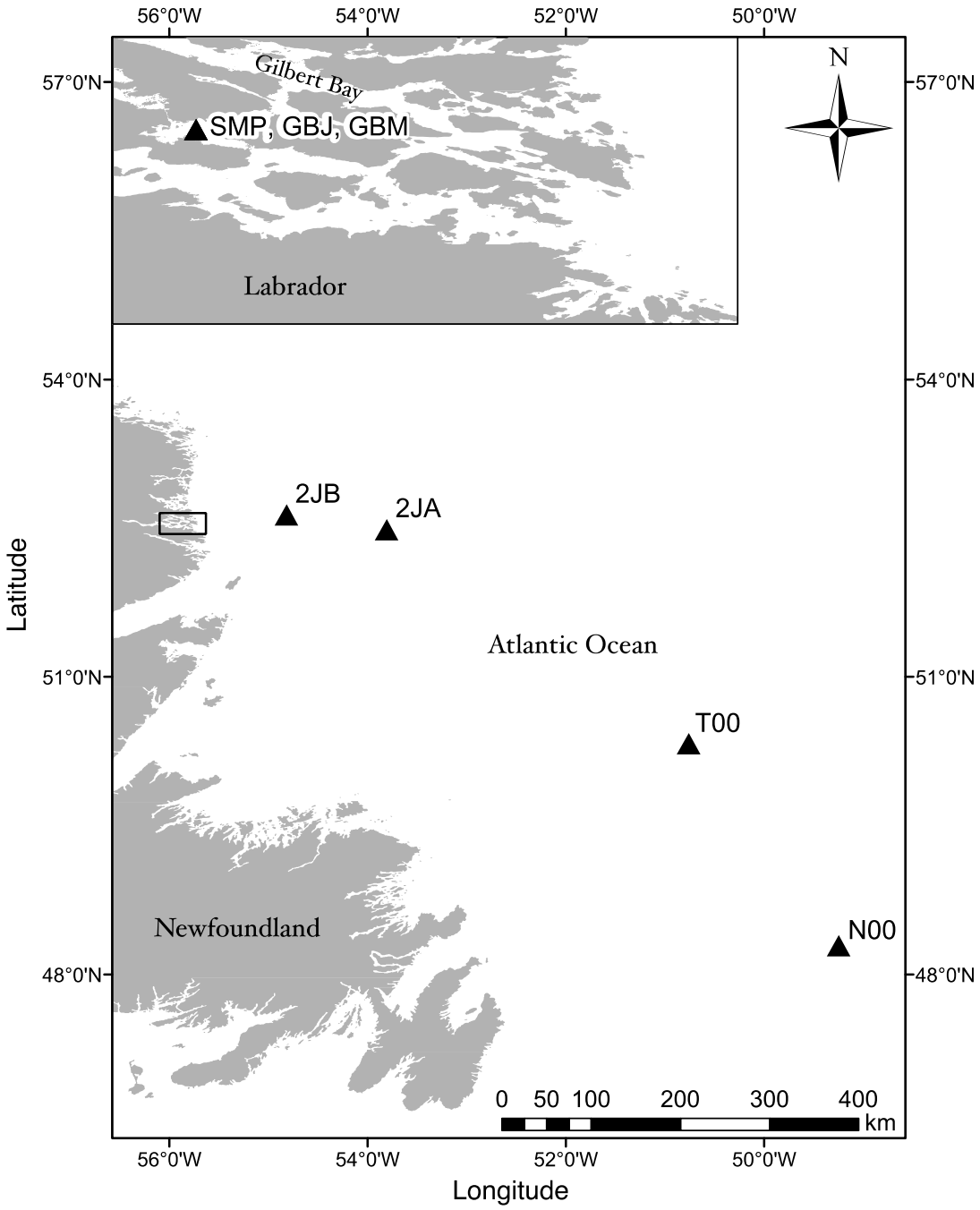


Figure 1. Map of sampling locations in Newfoundland and Labrador with fine-scale map of Gilbert Bay area. Additional information on each sample is provided in Table 1.

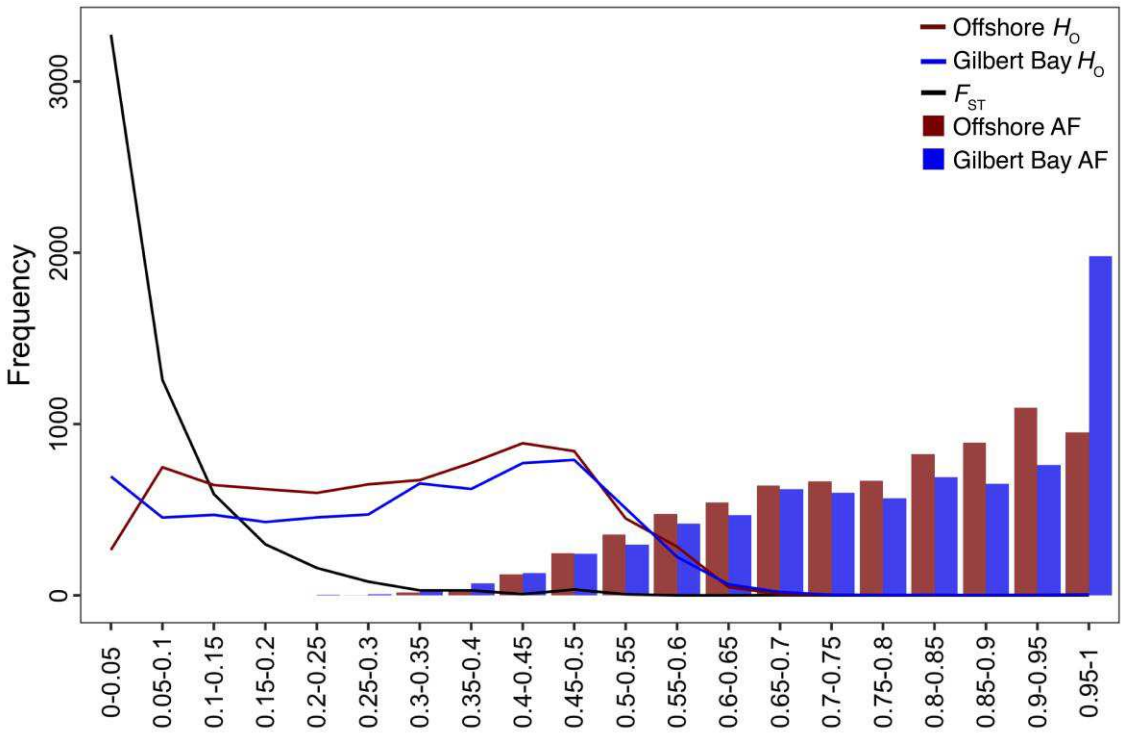


Figure 2. Allele frequency (AF), observed heterozygosity (H_o), and F_{ST} distribution of filtered loci (7531 SNPs). Allele frequencies and H_o calculated for each putative population: offshore allele frequencies (red bar), Gilbert Bay allele frequencies (blue bar), offshore H_o (red line) and Gilbert Bay H_o (blue line). Black line represents distribution of pairwise F_{ST} per locus.

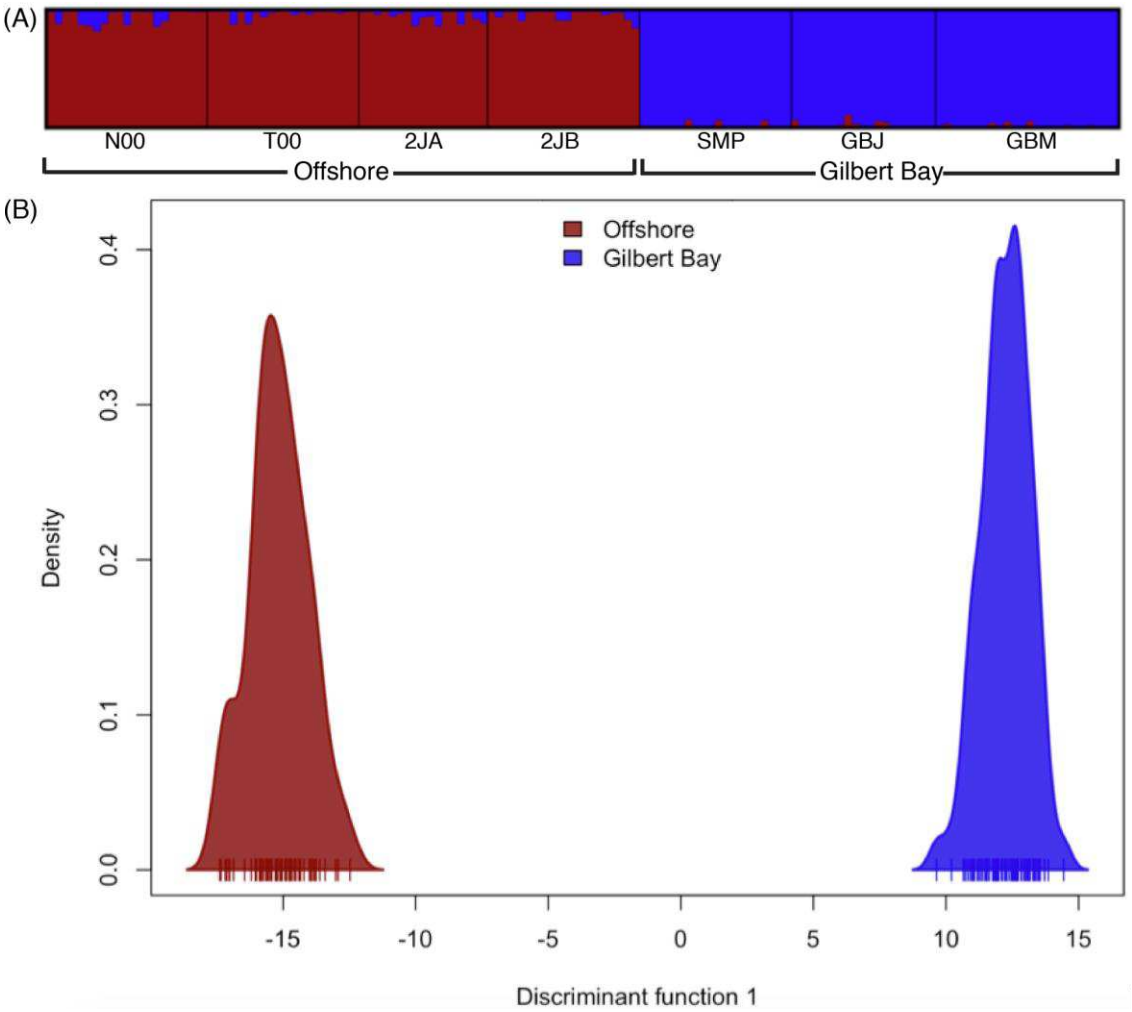


Figure 3. Analysis of population structure using filtered dataset (7531 SNPs). (A) Plot of individual admixture determined by STRUCTURE analysis for K=2. (B) Plot of discriminant function from the DAPC based on two clusters.

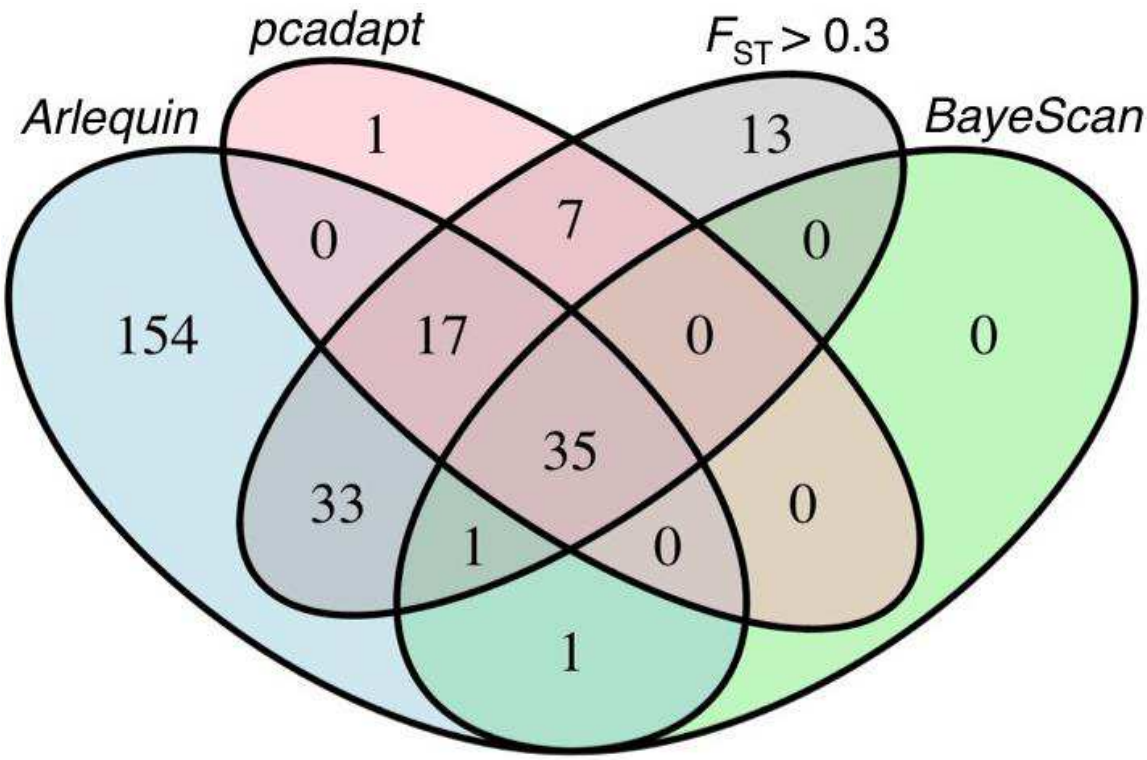


Figure 4. Venn diagram representing the overlap observed among methods used for selecting highly divergent SNPs: *Arlequin* (blue), *pcadapt* (red), F_{ST} threshold (grey) and *BayeScan* (green). Numbers indicate the number of SNPs identified by the corresponding method of selection.

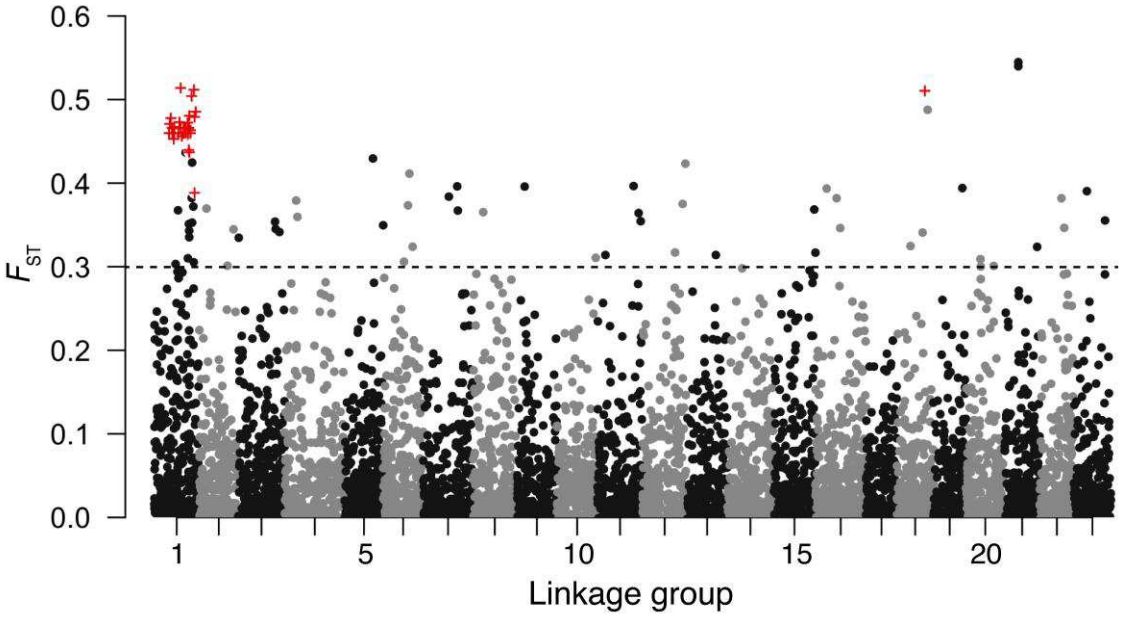


Figure 5. Manhattan plot of genetic differentiation showing per locus F_{ST} variation across each linkage group. The red crosses represent outlier loci identified by all three methods of outlier detection. Dashed line marks the F_{ST} threshold of 0.3.

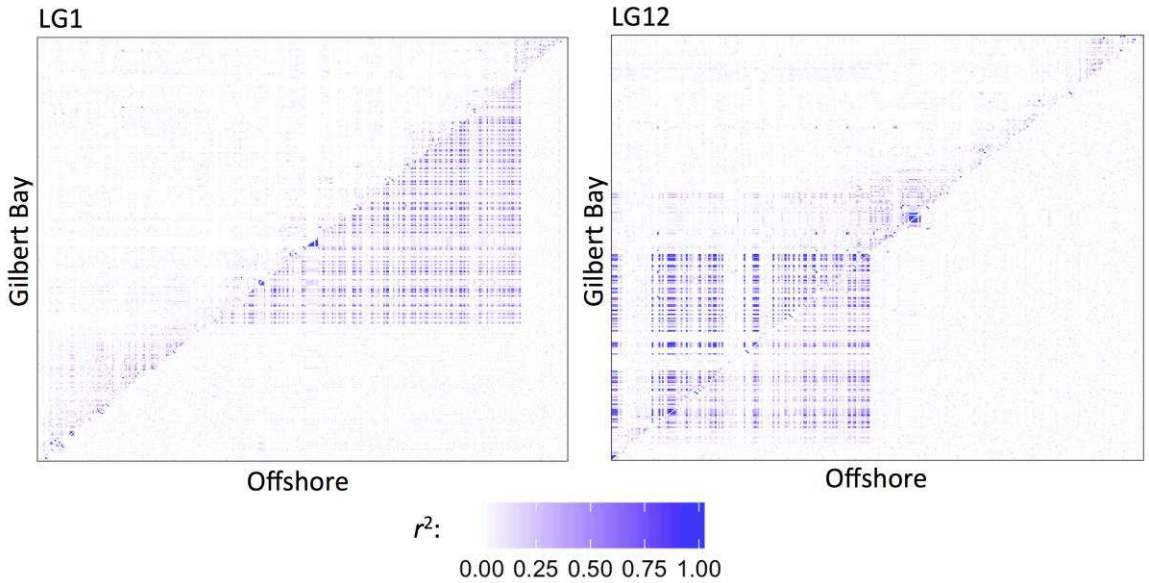


Figure 6. Pattern of pairwise LD, measured as r^2, within LG1 and LG12 for each population: Gilbert Bay (above diagonal) and offshore (below diagonal).

Chapter 3 – Genomic tools for management and conservation of Atlantic Cod in a coastal marine protected area

3.1 Abstract

Well-designed and managed MPAs can serve as effective tools for management and conservation. The Gilbert Bay MPA in coastal Labrador was created to protect a genetically distinct population of Atlantic Cod, however, decreases in abundance continue to occur potentially due to exploitation outside the MPA. I developed a SNP panel to identify Gilbert Bay cod in areas outside MPA boundaries where mixing with offshore cod occurs. A total of 365 individuals from Gilbert Bay, the surrounding areas, and offshore were genotyped for 10,913 genome-SNPs. Using F_{ST} rankings and guided regularized random forest, I selected 23 SNPs that obtain 100% accuracy in individual assignment and accurately estimate mixture proportions of Gilbert Bay cod in fishery samples from sites outside MPA boundaries. On average, fishery samples comprised of 17.3% Gilbert Bay cod. Estimates of N_e for the Gilbert Bay population ranged from 139 to 1256. These findings demonstrate the power of using genomic approaches for management of an exploited marine species and enhancing the design of MPAs.

3.2 Introduction

Marine protected areas and marine reserves are considered valuable tools for marine conservation and resource management (Gaines et al. 2010). They can play an important role in protecting marine species and habitat, conserving biodiversity, maintaining ecosystem function and increasing resilience to environmental changes (Allison et al. 1998; Lubchenco et al. 2003; Salomon et al. 2006). There is evidence that

MPAs can help ecosystems recover and are associated with more favourable fishing conditions outside MPA boundaries (Pauly et al. 2002; Lester et al. 2009; Fenberg et al. 2012). However, there are many instances where MPAs have failed to achieve their specific goals. The protection provided by an MPA depends heavily on the size and spatial arrangement of the MPA, and the dispersal potential of the targeted species (Shanks et al. 2003; Kininmonth et al. 2011; Moffitt et al. 2011). In addition, factors such as illegal or unsustainable harvesting, and emigration of target species or populations outside MPA boundaries must be avoided or limited for an MPA to provide maximum protection (Babcock et al. 2010; Edgar 2011; Edgar et al. 2014).

The effectiveness of MPAs for the protection of highly mobile species such as many marine fishes can be compromised (Gaines et al. 2010; Laurel and Bradbury 2006). In such cases, understanding distribution and dispersal patterns of the target populations is of key importance. To date, acoustic tagging has been a common tool used to track movements in space and time of highly mobile fish species (Nielsen et al. 2009) and can also help inform management and design of MPAs (Pittman et al. 2014; Morris et al. 2014) though sample sizes are often small. Genetic methods of population identification comprise another, potentially powerful tool for monitoring boundaries of populations and movement of individuals in and around MPAs. Genetic data can also contribute to other measures of the effectiveness of MPAs. For example, higher genetic diversity within MPAs, compared to non-protected areas, could be evidence of larger effective population sizes and therefore, in many cases, considered an indication of an effective MPA (Syms and Carr 2001; Munguía-Vega et al. 2015). Additionally, genetic patterns of isolation can be indicative of the amount of larval dispersal (Palumbi 2003; Kinlan and Gaines 2003)

and genetic connectivity between MPAs and surrounding non-protected areas (Green et al. 2015; Pujolar et al. 2015; Calò et al. 2016). DNA parentage analysis can also be used to estimate connectivity in MPA networks by tracking individuals (larvae, juveniles, adults) dispersing between MPAs (Planes et al. 2009). However, to date relatively few studies have used genetic approaches to evaluate the effectiveness of MPAs, despite the decreasing costs, relative ease of application, and broad utility of genetic methods (Allendorf et al. 2010).

In 2005, Gilbert Bay, a relatively enclosed embayment with an area of 60 km^2 in southern Labrador, Canada, was designated an MPA under Canada's *Ocean Act* to protect a resident population of Atlantic Cod. Previous microsatellite and SNP studies showed that the resident Gilbert Bay cod population is genetically distinguishable from other populations of cod (Ruzzante et al. 2000a; Beacham et al. 2002; Bradbury et al. 2013). Currently, the Gilbert Bay population is managed as part of the northern cod stock complex (NAFO subdivisions 2J3KL) (COSEWIC 2010). Gilbert Bay cod are thought to have existed at a relatively high density after the northern cod collapse in the 1990s (Morris and Green 2002), while other northern cod stocks had declined by as much as 99% since the early 1960s and remained at these low levels (COSEWIC 2010). In more recent years, however, the Gilbert Bay cod population has decreased in abundance despite the creation of the Gilbert Bay MPA. Biomass and research catch rates are estimated to have decreased by 83% and 54%, respectively, since 2005 (Morris and Green 2014). Acoustic tracking showed that some Gilbert Bay cod migrate to waters outside the MPA during the summer months, and that the home range of Gilbert Bay cod includes ~270km^2 outside the MPA limits (Morris et al. 2003, 2014). A reduced research

catch rate of Gilbert Bay cod was strongly correlated with commercial fishing in areas adjacent to the MPA, which suggests that fisheries outside the MPA boundaries may be contributing to the observed population decline (Morris and Green 2014). Given the current declining state of the Gilbert Bay cod population, harvesting of mature individuals, which are targeted in fisheries, is of concern because they are essential for population growth (Hutchings 1999; Kjesbu et al. 1996). Given these circumstances, there is a need for tools that will enable identification of Gilbert Bay cod wherever they occur, including in fisheries adjacent to the MPA.

Here, I use genome-wide SNP data to develop genomic tools to aid the management of the Gilbert Bay cod population. The first objective is to identify SNPs that enable accurate assignment of Gilbert Bay and offshore individuals to their respective populations of origin. Second, I use these SNPs to estimate the proportion of Gilbert Bay cod relative to offshore cod that are present in harvests in unprotected waters surrounding the MPA. The final objective is to use the complete, genome-wide SNP data to estimate N_e of the Gilbert Bay cod population over a 17-year period. Understanding the spatial and temporal dynamics of exploitation of Gilbert Bay cod and determining N_e will directly inform management and conservation of the Gilbert Bay Atlantic Cod population and improve the effectiveness of the Gilbert Bay MPA.

3.3 Methods

3.3.1 Sample collection

A total of 126 individuals were collected to represent each of the baseline groups, Gilbert Bay and offshore. Collection of Gilbert Bay cod occurred at The Shinney's, an important overwintering and spawning site located in one of the most inner regions of the

bay (Figure 1). Samples were collected over two different periods: 9 September 2012, and 1-10 June 2015 (Table 3). Collection for the offshore group occurred at four different locations in marine waters off the coast of Newfoundland and Labrador (Figure 1). The four sampling events occurred on 5 December 2010, 7 June 2011, 8 October 2015 and 10 October 2015 (Table 3). The remaining samples were collected during the mixed-stock period during late summer to represent fishery samples. It is during this period that Gilbert Bay cod are presumed to mix with offshore cod that migrate inshore to feed. Fishery samples (i.e. potential mixed aggregation samples) were collected from various locations in unprotected waters at the mouth of Gilbert Bay (Figure 1). These samples consisted of 240 individuals that were collected over a six-year period (September 2009 – August 2015) (Table 3).

3.3.2 DNA extraction and genotyping

Fin clips ($2mm^3$) were collected from live individuals and were immediately preserved in 95% ethanol. Genomic DNA was extracted using a Qiagen DNeasy 96 blood and tissue kit following protocol described by the manufacturer or using a standard phenol-chloroform procedure (Sambrook et al. 1987). All DNA was quantified using QuantIT PicoGreen (Life Technologies) and was normalized to a final concentration of 50ng/µL.

Individuals were genotyped for 10,913 SNPs using an Illumina (San Diego, USA) array developed by a Norwegian consortium consisting of four research organisations (CEES, CIGENE, NOFIMA, and Havforskningsinstituttet). The custom array was constructed using genomes of seven Atlantic Cod across the Northeast Atlantic. Genomes were sequenced using a shotgun approach and by aligning reads to the garMor1 reference

genome (Star et al. 2011). A total of 2,877,794 putative SNPs were identified, from which 10,913 SNPs were selected for the assay based on their physical distribution and functional associations. This includes 260 SNPs from previous studies (Hubert et al. 2010; Moen et al. 2008), 672 SNPs in close proximity to candidate genes, and 1595 non-synonymous coding SNPs. On average, the array includes 409 SNPs per chromosome (Kent et al. in prep.). Genotyping of individuals for this study was carried out at CIGENE, Norwegian University of Life Sciences in As, Norway.

3.3.3 Quality control filters and population genetic statistics

Of the 10,913 genotyped SNPs, only SNPs that met specific criteria were selected for further analysis. Any loci that did not meet the criteria for bi-allelic SNP according to their clustering pattern using a large sample set of more than 5000 individuals (Kent et al. in prep.), were removed prior to the filtering steps for this study. PLINK was used for the following filtering procedures (Purcell et al. 2007). To begin, any individual with low genotyping (less than 85% complete) was removed from the dataset. For the next set of filtering steps, only baseline individuals were considered. To ensure that SNPs did not result due to genotyping error, any SNP with a minor allele frequency less than 0.01 was removed. SNPs were also filtered to ensure missing data at any given loci was not greater than 15%.

Observed heterozygosity (H_o) for each locus was calculated with *Arlequin v3.5.* A measure of pairwise population differentiation, Weir and Cockerham's F_{ST} (1984), and its statistical significance were also estimated using *Arlequin v3.5* (Excoffier and Lischer 2010).

3.3.4 Linkage disequilibrium detection

Following filtering, LD among loci was detected using PLINK with an r^2 threshold greater than 0.4 (Purcell et al. 2007). Next, a panel of independent SNPs (i.e. showing no evidence of LD) with the highest measures of pairwise F_{ST} was selected. The function *genepop_toploci()* in the *R* package *genepopedit* (Stanley et al. 2016) was used to retain the locus with the highest F_{ST} from each group of loci in LD. Only SNPs with an F_{ST} greater than 0 were considered for inclusion in the panel.

3.3.5 Determining population of origin

The population of origin was assessed for each individual using the Bayesian method implemented in STRUCTURE 2.3.4 (Pritchard et al. 2000). The set of 7,568 filtered SNPs (including loci in LD) was used for the analysis of population structure. Three replicate runs were conducted for each value of K (i.e. number of populations) from 1 to 4. Each replicate consisted of a burn-in period of 100,000 followed by a run length of 500,000 iterations. An approach implemented in CLUMPAK was used to determine the best value of K for the baseline samples. Calculated likelihood values and the DeltaK statistic were used to identify the value of K that captured the most structure (Evanno et al. 2005).

3.3.6 Baseline assignment

To assess the accuracy of baseline assignment, an approach developed by Anderson et al. (2008) as implemented in the *R* package *assigner* was used (Gosselin et al. 2016). Baseline populations were defined based on results from STRUCTURE. Three methods for selecting and ranking highly informative SNPs for baseline assignment were used. First, SNPs from the complete panel (not filtered for LD) were ranked according to

F_{ST}'s and used for baseline assignment. Second, baseline assignment was repeated using top-ranked SNPs showing no evidence of LD. The third method used to select and rank the most informative SNPs for baseline assignment was guided regularized random forest (GRRF). Random forest is a machine learning technique that uses a series of decision trees for classification. GRRF uses the importance scores from a prior random forest run to guide the process of selecting loci in the regularized random forest. The *R* package *RRF* with a gamma of 0.5 was used to select and rank loci (Deng 2013). Following ranking and selection, individuals were assigned to the baseline using a classic leave-one-out method (Anderson et al. 2008). Assignment accuracy was determined by assessing whether the holdout individual assigned correctly using the reference sample, which was comprised of all other baseline individuals. The assignment accuracies of different subsets of top ranked loci (top 1 to top 25) according to the F_{ST} methods and GRRF were compared.

3.3.7 Analysis of fishery samples

Based on the results of the baseline assignment test, the minimum number of SNPs needed to obtain 100% accuracy was determined and used as a standard number of loci to compare methods of selection. The standard number of loci was a subset from the complete dataset using all three methods of selection. The three resulting panels of reduced size and the complete panel of filtered loci (7,568 SNPs) were used to estimate mixture proportions of fishery samples and to assign individuals from the fishery samples to one of the baseline populations. Mixture proportions were calculated using a Bayesian approach implemented in *gsi_sim* (Anderson et al. 2008). The burn-in period consisted of 5000 steps and was followed by 25000 Markov Chain Monte Carlo (MCMC) replicates.

The results of the subset panels and the complete panel were compared to evaluate the performance of the selected panels. A linear regression between the proportions of Gilbert Bay cod determined by each selected panel and the complete panel was fitted in order to test whether the proportion estimates differed between panels.

3.3.8 Estimating effective population size

A total of 121 individuals were collected from Gilbert Bay spawning grounds (Table 3) over a 17-year period and used for calculations of N_e. N_e is the number of individuals in a idealized population that would experience genetic drift at the same rate as a real population. Because loci under selection can bias estimates of N_e, a panel of only neutral loci was identified. Any locus marked as an F_{ST} outlier or located within a known chromosomal inversion was removed prior to estimating N_e. To minimize the possibility of false negatives, two methods of outlier detection were used to identify potential SNPs under selection. These included a Bayesian approach as implemented in the program *BayeScan v.2.1* (Foll and Gaggiotti 2008) and the *Fdist* approach (Beaumont and Nichols 1996) as implemented in the program *Arlequin v3.5*. *BayeScan* estimates population specific F_{ST} coefficients and based on the posterior distributions identifies SNPs under selection if they fall above a specified threshold. Here, 100,000 iterations were run with prior odds set at 100 and SNPs with a posterior probability over 0.95 were considered as under selection. *Arlequin* uses measures of F_{ST} and the corresponding *P*-value for each locus to determine which loci are outliers. A total of 200,000 coalescent simulations were performed and loci with a *P*-value less than or equal to 0.01 were considered outliers. To detect large chromosomal inversions and SNPs located within

them, the *R* package *inveRsion* was used. A block size of 3 and *thbic* threshold of 0 was used to identify inverted genomic regions.

Once the panel of neutral loci was identified, N_e was calculated using 1) the LD method and 2) Jorde and Ryman's temporal method as implemented in the software *NeEstimator v2* (Do et al. 2014). Generally, estimates calculated using the LD method assume that all loci in a SNP data set are unlinked (Hollenbeck et al. 2016). Consequently, downward biases in N_e estimates can occur due to unusually high coefficients of correlation caused by physical linkage and not true LD if no information on chromosome location is used (Waples et al. 2016). To account for these biases, naïve N_e estimates ($\widehat{N}_e$) were adjusted using the following formula by Waples et al. (2016):

$$\frac{\widehat{N}_e}{N_e} = -0.910 + 0.219 \times \ln(cM)$$

This formula uses the total length of the genome to estimate the degree that $\widehat{N}_e$ has been downwardly bias due to physical linkage and generates an adjusted estimate of N_e that accounts for this bias.

The total length of the genome (cM) is specified in centimorgans. In this case, the female linkage map was used to determine total genome length (S. Lien, Centre for Integrative Genetics, Ås, Norway, personal communication). Here, $\propto = 0.05$ and the 0.025 and 0.975 points of the chi square distribution of *J* were used, where *J* equals the degrees of freedom, to calculate confidence intervals (C.I.'s) for bias-corrected N_e as outlined in Waples (2006).

3.4 Results

3.4.1 Quality control filters, population genetic statistics and linkage disequilibrium detection

Genome-wide SNP data were obtained from 365 individuals; however, one individual was discarded during the filtering process due to low genotyping. Of the 10,913 genotyped loci, 8581 were categorized as bi-allelic SNPs and included in the analyses for this study. Filtration steps flagged 801 SNPs with a minor allele frequency less than 0.01 and 258 SNPs with missing genotypes in more than 85% of individuals. Flagged SNPs were removed and the resulting filtered dataset consisted of 7,568 SNPs. Estimates of H_o and F_{ST} for filtered loci ranged from 0 to 0.67 and from 0 to 0.54, respectively (Figure 7). After the detection and elimination of SNPs in high linkage disequilibrium (r^2>0.4), a panel of 5,025 loci showing no evidence of LD and with the highest F_{ST}'s was selected for baseline assignment.

3.4.2 Determining population of origin

The population of origin for all 365 individuals was determined using the Bayesian method implemented in STRUCTURE 2.3.4 (Figure S5). The filtered panel of 7,568 was analyzed to determine population structure and to calculate the fraction of each individual's genome derived from each population. The best number of populations (K) estimated by CLUMPAK was two (Figure S6). All individuals were assigned to one of the two populations determined by CLUMPAK with a probability greater than 80% (Table S4).

3.4.3 Baseline assignment

Gilbert Bay and offshore individuals were assigned to one of the baseline populations using an approach developed by Anderson et al. (2008). Accuracy of assignment using 1-25 of the top-ranked SNPs was calculated for each of the three panels of loci: all SNPs ranked according to F_{ST}, SNPs showing no evidence of LD ranked according to F_{ST}, and SNPs ranked via GRRF. Removing SNPs in LD and using GRRF for selection significantly increased the accuracy of baseline assignment with the *top 25* loci (Figure 8). A minimum of 23 top-ranked SNPs from the panel of loci showing no evidence of LD was required to obtain 100% assignment accuracy for individual assignment to both baseline populations. Whereas, the 23 top-ranked SNPs from the complete panel and from GRRF yielded 88.5% and 99% accuracy overall, respectively (Table S5). These 23-SNP panels were selected for analysis of mixed fishery samples.

3.4.4 Analysis of fishery samples

Over the 6-year sampling period (2009-2015), mixing of Gilbert Bay cod and offshore cod was observed at all sites sampled outside the Gilbert Bay MPA except the most southern site, Spear Point (Figure 9). The highest proportion of Gilbert Bay cod (0.65) was observed at William's Harbour Run in 2011. The proportions of Gilbert Bay cod at Spear Point in 2009 and at Kelly's Island in 2015 were 0. Otherwise, the proportional contribution of Gilbert Bay cod to fishery samples was 0.09-0.26 for all sites over the 6-year sampling period (Figure 10, Table S6).

Fishery samples were analyzed using the three selected 23-SNP panels and the complete panel (7,568 SNPs) to estimate mixture proportions. Bayesian and maximum likelihood estimates of mixture proportions were similar and therefore, only Bayesian

estimates are considered here. The mixture proportion estimates calculated for each fishery sample with the 23-SNP panel showing no evidence of LD and complete panel were similar (Table S6). A linear regression relating estimates of this 23-SNP panel to estimates of the complete panel indicated no difference between the two panels, indicated by a high correlation value ($r = 1$) (Figure 11). In some cases, the GRRF 23-SNP panel overestimated the proportion of Gilbert Bay cod in a sample (Table S6), however, a linear regression analysis indicated only a small overall difference between panels, as indicated by a high correlation value ($r = 0.99$) (Figure 11). The 23-SNP panel that included SNPs in LD was unable to accurately estimate mixture proportions of fishery samples where Gilbert Bay cod were found (Table S6). Consequently, significant differences were observed between estimates of this panel and the complete panel, as indicated by a low correlation value ($r = 0.11$) (Figure 11).

3.4.5 Estimating effective population size

Both LD and temporal methods yielded N_e estimates under 1300 (Figure 12, Table S7). To create a neutral panel for estimating N_e, a total of 529 SNPs identified as potentially under selection or located within a chromosomal inversion were removed from the complete dataset. The female linkage map used to calculate the bias correction for LD estimates has 23 chromosomes and a total length of 1681.7 cM. Waples et al.'s (2016) method for bias correction using total genomic length suggests the estimates are biased 28.3% down from the true value of N_e. Bias-corrected values of N_e ranged between 719 and 1048 over the 17-year period (1998-2015). No overall decrease in N_e was observed over this period and the width of the 95% C.I.'s ranged between 21.9 and 92.5. Estimates using the LD method for years with a larger number of samples (e.g. 1998 and

2012) had tighter C.I.'s (Figure 12). Temporal estimates ranged between 139 and 1256, excluding negative estimates (Table S7). When using the temporal method, larger estimates of N_e were observed when samples were separated by a greater number of years (Figure 12).

3.5 Discussion

Using the Gilbert Bay MPA as a study site, I demonstrate the power of genomic tools for identifying exploitation outside MPA boundaries and estimating N_e, thus providing valuable information for the effective design and management of an MPA. Using a panel of 7568 SNPs, I found that the Gilbert Bay population contributed an average of 17.3% (range: 0-65%) to the fishery samples collected in waters adjacent to the MPA. In addition, N_e is estimated to be fewer than 1300 individuals and therefore the observed mixing of Gilbert Bay cod with other cod stocks in areas where harvesting occurs is a cause for concern. This has clear implications for the design of the MPA with respect to size and placement of boundaries. These findings suggest that the current MPA boundaries and existing fishery regulations are insufficient to protect commercial sized Gilbert Bay cod from harvesting in areas adjacent the MPA.

I identified a panel of 23 SNP markers that can accurately distinguish Gilbert Bay cod and offshore cod. This reduced panel can serve as a cost-effective tool to enhance management and improve the design of the Gilbert Bay MPA. In addition, finite and stable estimates of effective population size were calculated using a genomic approach. This provides useful information on the status of the Gilbert Bay cod population, but also serves as a framework for future work on estimating N_e with thousands of genome-wide SNP markers.

When Gilbert Bay was designated an MPA in 2005, the primary conservation goal of the Gilbert Bay MPA was to protect the genetically distinct Atlantic Cod population that overwinters in the bay (DFO 2007). Using a genome-wide panel of SNP markers, this study confirms that strong population divergence of the Gilbert Bay cod population exists. Shallow sills at the entrances of the bay may help retain eggs that are most abundant at depths of 4-7 m (Morris and Green 2002) and thus help maintain genetic differentiation between Gilbert Bay cod and offshore cod. In addition, the spawning times of Gilbert Bay cod and offshore cod do not coincide, which likely serves as a mechanism to reproductively isolate Gilbert Bay cod from other populations. Most spawning along the Labrador coast occurs during March and April (May 1966), but Gilbert Bay cod spawn for about three weeks during mid-May (Morris and Green 2002). The mechanisms that drive population divergence of Gilbert Bay cod are not fully understood and further insight into potential drivers such as pre- and post- zygotic barriers or adaptation to biotic and abiotic factors is required.

Designing effective MPAs requires spatial information on distribution and adult movements of the focal species or population (Apostolaki et al. 2002; Botsford et al. 2003; Costello et al. 2010). I used a genomic approach to investigate the occurrence of Gilbert Bay cod outside the MPA boundaries during late summer/early autumn. This study provides genetic evidence that Gilbert Bay cod occur outside MPA boundaries. Substantial mixing of Gilbert Bay cod and offshore cod was observed at coastal sites sampled east of the MPA boundaries during August and September. The highest proportion of Gilbert Bay cod was observed at William's Harbour Run, where 65.0% of collected individuals were from the Gilbert Bay population. At the more northern

entrance to Gilbert Bay, fishery samples consisted of 26.1, 9.5%, 13.8%, and 6.5% Gilbert Bay cod at the Bull, Hare Island, Kelly's Island, and Red Island, respectively. This strongly suggests that the protection provided by the current MPA is inadequate, and consequently, Gilbert Bay cod are being accidently harvested in the Northern cod Stewardship fishery outside the MPA. The Gilbert Bay cod population is a relatively small population, and therefore the protection provided by the MPA is important for its conservation. Over-exploitation of smaller and less productive populations, such as the Gilbert Bay Atlantic Cod population, is of concern because it can contribute to loss of overall biodiversity (Ruzzante et al. 2000b), which is associated with decreased sustainability and stability of fisheries due to portfolio effects (Hilborn et al. 2003; Schindler et al. 2010).

Although the findings of this study provide clear evidence of mixing of Gilbert Bay cod in areas outside the MPA boundaries, our understanding of the spatial and temporal characteristics of this mixing is limited. A greater number of samples collected over a longer time period in late summer/early fall would provide greater resolution of how the proportion of Gilbert Bay cod in areas outside MPA boundaries changes over this period. In addition, a greater number of fishery samples from varying locations would provide a better understanding of areas where mixing of offshore and Gilbert Bay cod occurs. In particular, the distribution of fishery samples was sparse at the more southern entrance to Gilbert Bay.

Recent advances in genomic techniques have allowed SNPs to be screened with greater ease and at lower costs (Allendorf et al. 2010). Consequently the application of genomics in fisheries management, such as genetic stock identification using SNP-panels,

is becoming more common (e.g. Larson et al. 2014; Bradbury et al. 2015, 2016). A genomic approach to genetic stock identification is highly informative, however, it is not always cost-effective. Therefore, the identification of highly informative SNPs is crucial for optimizing panel performance, while limiting panel size. Here, I developed a reduced SNP panel of 23 SNPs that can accurately perform individual assignment and determine mixture proportions of Gilbert Bay cod and offshore cod in coastal Labrador. I used three methods for selecting informative loci: ranking all loci, ranking only unlinked loci, and GRRF. Removing loci in LD prior to ranking resulted in a clear improvement in panel performance. Groups of linked loci display similar patterns in allele frequencies and therefore contribute repetitive information for genetic stock identification. Retaining only the highest ranked SNP per group of linked loci avoids the inclusion of loci with redundant information in the selected SNP-panel and is therefore an improved method for selecting SNP panels for genetic stock identification. A recent study aimed at assigning Atlantic Salmon (*Salmo salar*) to their river of origin found that GRRF performed the best for selecting informative SNPs (Sylvester et al. in prep.). Here, GRRF was an effective method for selection, but performed slightly worse than ranking SNPs showing no evidence of LD according to F_{ST}.

Despite a decrease in biomass since the MPA designation, N_e estimates of the Gilbert Bay cod population do not show a decline over the same period. Therefore it seems that census population size (N_C) has declined, while N_e has not. This suggests the N_e/N_C ratio for the Gilbert Bay cod population has increased. This may be driven by a decrease in variance in reproductive success at a low density, or small population size, due to reduced competition. This means that some individuals that would not have

otherwise reproduced do so at low densities. This phenomenon is called "genetic compensation" and can explain an increase in N_e/N_C when population size decreases (Kuparinen et al. 2016). The observation of larger N_e/N_C ratios when population size is reduced is not unprecedented and similar trends are reported in studies by Palstra and Ruzzante (2008), Beebee (2009), and Saarinen et al. (2010). Additionally, genetic variability and N_e of an Atlantic Cod population in the southern Gulf of St. Lawrence did not decline despite a substantial population decline (Therkildsen et al. 2010). In contrast, analyses of N_e in a large population of New Zealand snapper (*Pagrus auratus*) did detect a reduction in N_e when census size was reduced due to fishing (Hauser et al. 2002). In the face of population declines, small populations may be more resilient to changes in effective population size than large populations due to genetic compensation. However, changes in life history traits such as fecundity and early-life mortality can also affect N_e (Nunney 1996; Waples 2002; Hedrick 2005).

Unlike previous studies that aim to estimate N_e of different Atlantic Cod populations using microsatellites (Therkildsen et al. 2010; Poulsen et al. 2006; Hutchinson et al. 2003), this work uses thousands of markers to calculate N_e. This novel approach gave estimates of N_e with relatively narrow confidence intervals that are unprecedented in other studies on Atlantic Cod. When N_e is large, it is difficult to calculate estimates with finite bounds and extensive sampling is required (Ovenden et al. 2007; Palstra and Ruzzante 2008). In this study, only two estimates using the temporal method had a range that included infinity, which is a common issue when estimating N_e of marine species with large population sizes. This uncertainty arises when the signal-to-noise ratio becomes smaller due to drift having only a slight effect in large populations

(Waples 1989). The precision of all other estimates using the temporal method and LD method was high. It should be noted, however, that precision of N_e estimated using the LD method is overestimated when a large number of loci are used. For example, Waples et al. (2016) showed that when N_e is 200 and 4096 loci are used, the 95% C.I.'s should be 20 times wider than what is estimated under naïve assumptions. This occurs because the effective degrees of freedom are substantially fewer than the nominal degrees of freedom used to calculate N_e (Waples et al. 2016). The precision of this study's N_e estimates could be improved by: 1) increasing the number of individuals sampled at each time; 2) increasing the number of time points; 3) increasing the number of generations between samples; or 4) increasing the number of unlinked loci (Waples 1989; Wang 2001; Palstra and Ruzzante 2008).

Both the temporal and LD method for estimating N_e are based on models that assume discrete generations, no mutation, no selection, and no migration. The assumption of discrete generations is violated here because Atlantic Cod have overlapping generations. To reduce any bias induced by this violation, Jorde and Ryman (1995) suggest that samples be separated according to age class and N_e estimated be based on comparisons among consecutive year classes. Here, I separated samples according to their date of collection. As a result, samples collected in different years may include individuals from the same cohort. In such a case, a greater change in allele frequencies would be observed (Ryman 1997), thus causing a downward bias in N_e estimates. A bias in N_e may also arise when migration is high. This can cause changes to allele frequencies in the recipient population so that they approach those of donor populations. Therefore, estimates of N_e reflect that of the entire metapopulation, opposed to the study population

(Palstra and Ruzzante 2011; Waples and England 2011; Gomez-Uchida et al. 2013). A bias in N_e can also occur when migration is low and ignored. In such cases, a downward bias can occur due to LD resulting from mixture (Palstra and Ruzzante 2011; Waples and England 2011). In this study, however, no significant admixture was observed suggesting an absence of recent immigrants. In addition, physical attributes of Gilbert Bay and biological characteristics of Gilbert Bay cod both suggest that immigration cannot occur and therefore we suspect any bias in N_e of the Gilbert Bay population due to migration is minimal. Additionally, loci under selection and loci linked to them can influence N_e estimates (Nunney and Elam 1994). Multiple methods of detecting selection were used in this study to reduce the number of false negatives and thus ensuring only neutral loci were used for estimating N_e.

3.5.1 Conclusion

In conclusion, this study shows significant contributions of Gilbert Bay cod to fisheries outside the Gilbert Bay MPA boundaries, thus demonstrating that the current design of the Gilbert Bay MPA needs improvement. Consequently, the MPA provides insufficient protection for the resident population of Atlantic Cod. The 23-SNP panel developed here can serve as a cost-effective and accurate tool to further investigate and monitor contributions of Gilbert Bay cod to coastal fisheries. Despite exploitation outside MPA boundaries, N_e has remained fairly stable. This can be attributed to genetic compensation that maintains a relatively constant N_e when N_C is reduced. In addition, I show that using a genomic approach for estimating N_e is well suited for achieving N_e estimates with finite bounds. The implementation of genomic tools in future work and the

findings of this study will help inform management of the Gilbert Bay cod population and improve the design of the Gilbert Bay MPA.

Table 3. Site locations, coordinates, sample ID, date of collection and sample size for all collected samples of Atlantic Cod used for Chapter 3.

	Location	Latitude	Longitude	Sample ID	Sampling time	Sample size
Baseline samples						
Offshore	3L	48.287	-49.247	N00	07-Jun-11	16
	3K	50.322	-50.762	T00	05-Dec-10	15
	2J	52.482	-53.810	2JA	10-Oct-15	16
	2J	52.633	-54.817	2JB	08-Oct-15	16
Gilbert Bay	Shinney's	52.585	-56.033	SMP	09-Sep-12	20
	Shinney's	52.585	-56.033	GBJ	09-Sep-12	19
	Shinney's	52.585	-56.033	GBM	1-Jun to 10-Jun-15	24
Fishery samples						
	Hare Island	52.588	-55.724	CCF11	07-Sep-11	8
	Kelly's Island	52.593	-55.766	GBK	09-Sep-12	41
	Kelly's Island	52.593	-55.766	MFF/MFI	06-Aug-13	48
	Kelly's Island	52.593	-55.766	MFE	1-Aug to 5-Aug-14	24
	Kelly's Island	52.593	-55.766	MFG	1-Aug to 5-Aug-15	24
	Red Island	52.583	-55.709	CCF09	18-Sep-09	31
	Spear Point	52.45	-55.631	SP	29-Jul-09	20
	The Bull	52.606	-55.748	MFH	03-Aug-13	23
	William's Harbour Run	52.533	-55.733	SR	15-Sep-11	20
Temporal samples						
1998	Shinney's	52.585	-56.033	GBA	Aug-98	39
2004	Fox Cove	52.568	-55.802	GBE	Sep-04	19
2012	Shinney's	52.585	-56.033	GBJ/SMP	09-Sep-12	39
2015	Shinney's	52.585	-56.033	GBM	1-Jun to 10-Jun-15	24

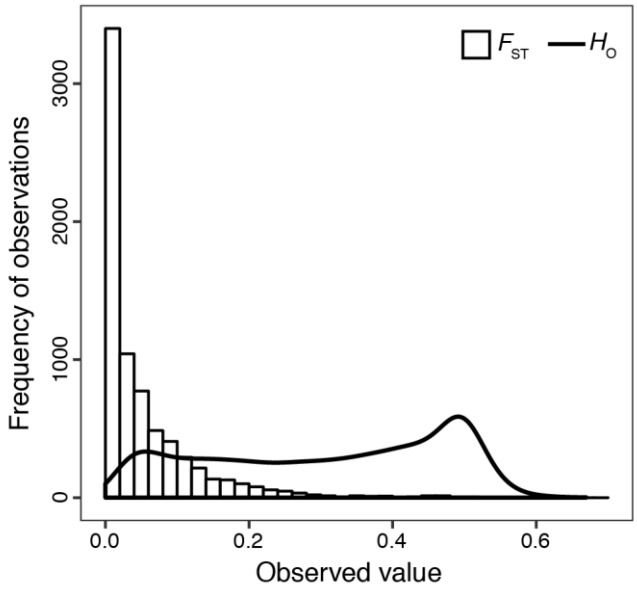


Figure 7. Frequency distribution of pairwise F_{ST} (vertical bars) and observed heterozygosity (H_O) of all samples combined (solid line) for each locus in the filtered dataset (7568 SNPs).

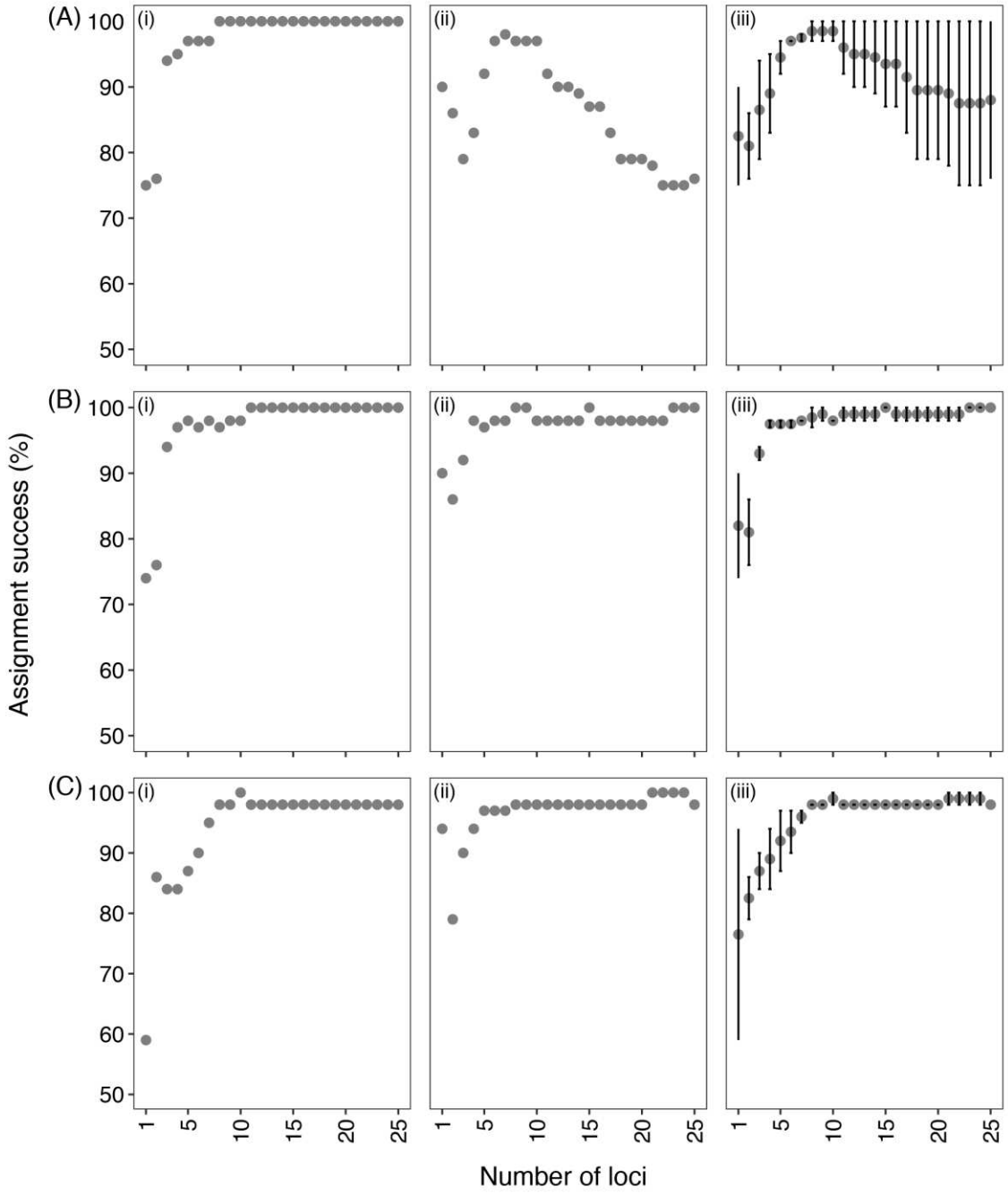


Figure 8. Baseline assignment accuracy determined by *assigner* for panels chosen by each of the three selection methods: (A) F_{ST} ranking of all loci (B) F_{ST} ranking of SNPs showing no evidence of LD and (C) GRRF ranking. Each box corresponds to a baseline group or the overall mean for both baseline groups: (i) Gilbert Bay, (ii) Offshore, (iii) overall.

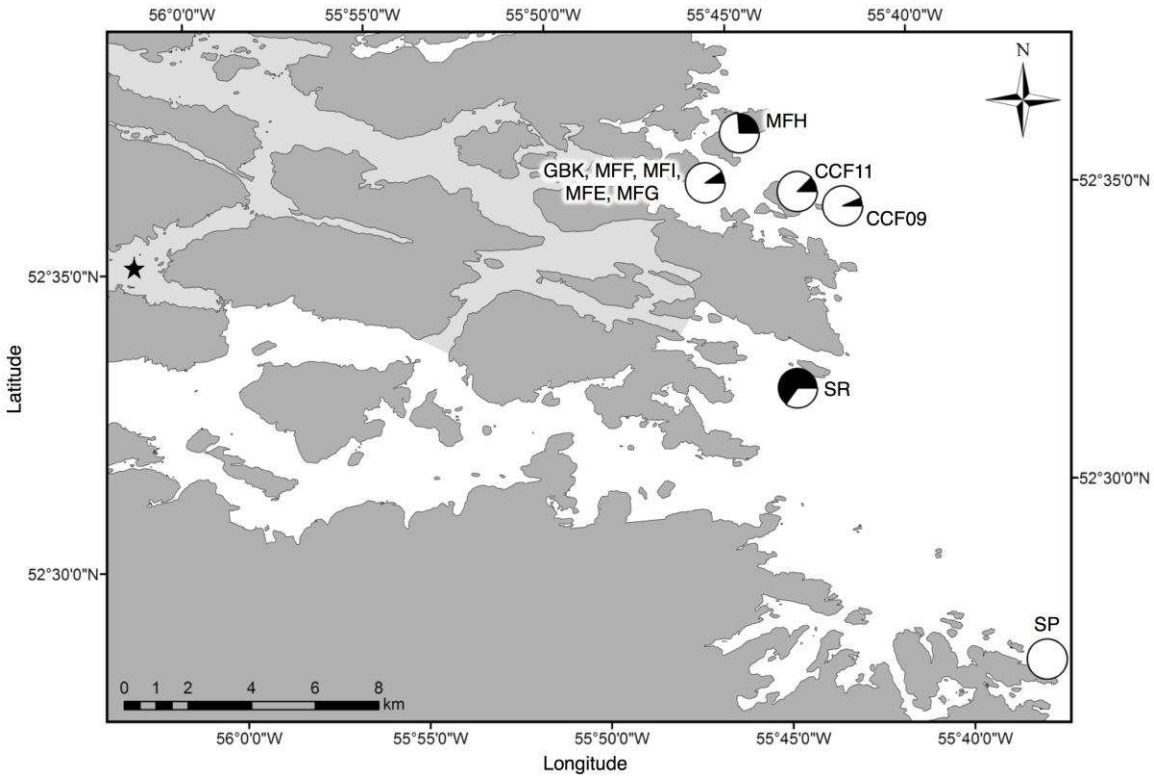


Figure 9. Map of the Gilbert Bay MPA and surrounding waters. The light grey shading marks the area protected by the MPA designation. Pie charts show proportional contributions of Gilbert Bay cod (black) and offshore cod (white) to sites where fishery samples were collected as determined by individual assignment in STRUCTURE. Proportions of individuals from the Gilbert Bay population at each sampling location are listed in Table S6. Data from different collection dates is combined for each site. The star marks the location of the Shinney's, an important spawning site for Gilbert Bay cod.

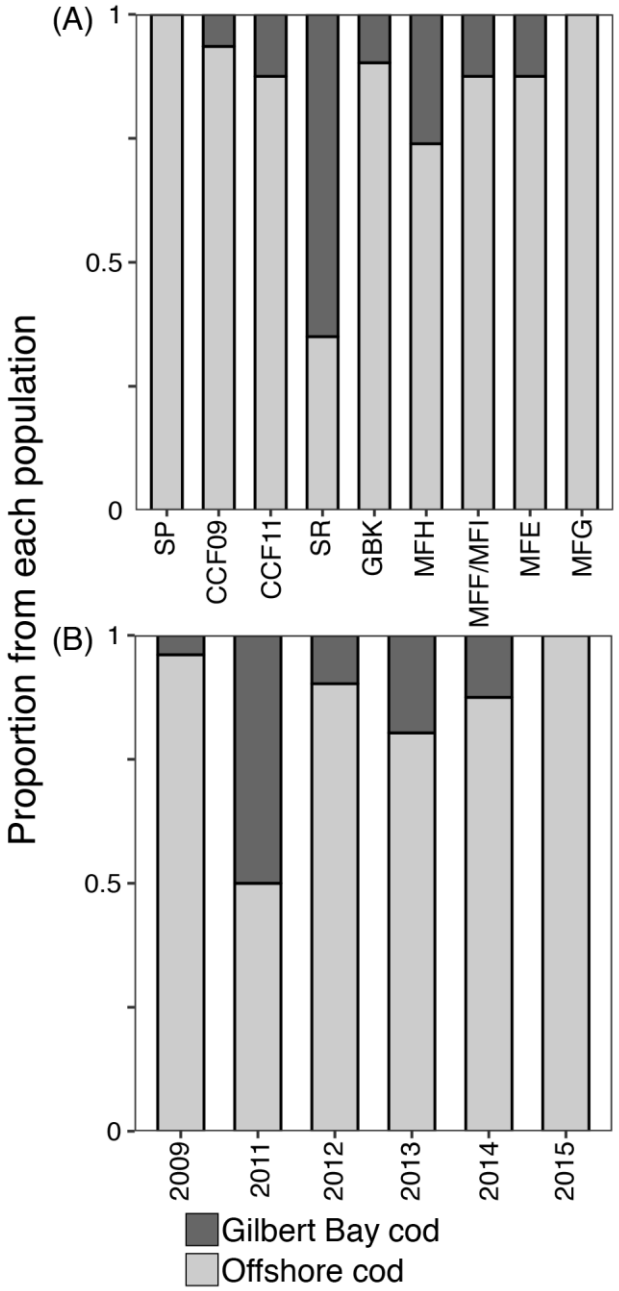


Figure 10. Changes in the proportional contribution of Gilbert Bay cod to fishery samples overtime. Each bar indicates the proportion of each cod population based on the STRUCTURE analysis used to determine population of origin for each individual. (A) Contributions observed at each site over time. (B) Each collection year is summarized by combining all data for that year.

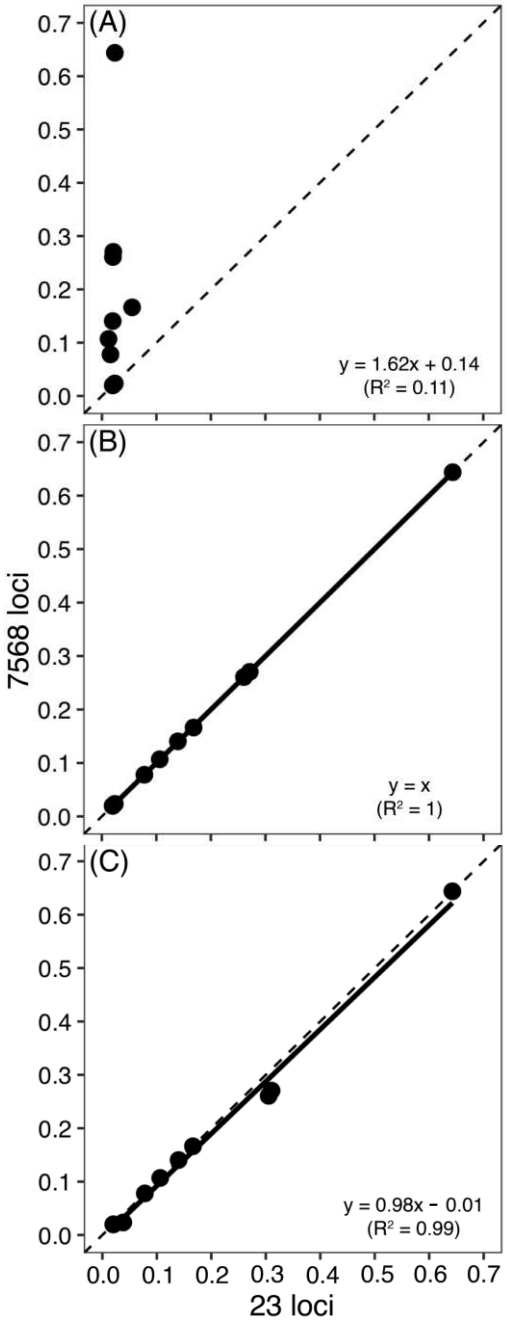


Figure 11. Proportions of Gilbert Bay cod in fishery samples calculated in *gsi_sim* using the complete panel (7568 SNPs) are plotted against mixture proportions calculated using the top 23 SNPs chosen by each selection method: (A) F_{ST} ranking of all loci (B) F_{ST} ranking of SNPs showing no evidence of LD and (C) GRRF ranking. The dashed line (m=1) indicates where the mixture proportions determined by each panel are equal. Linear regression parameters and coefficients for comparisons between each of the 23-SNP panels and complete panel are listed.

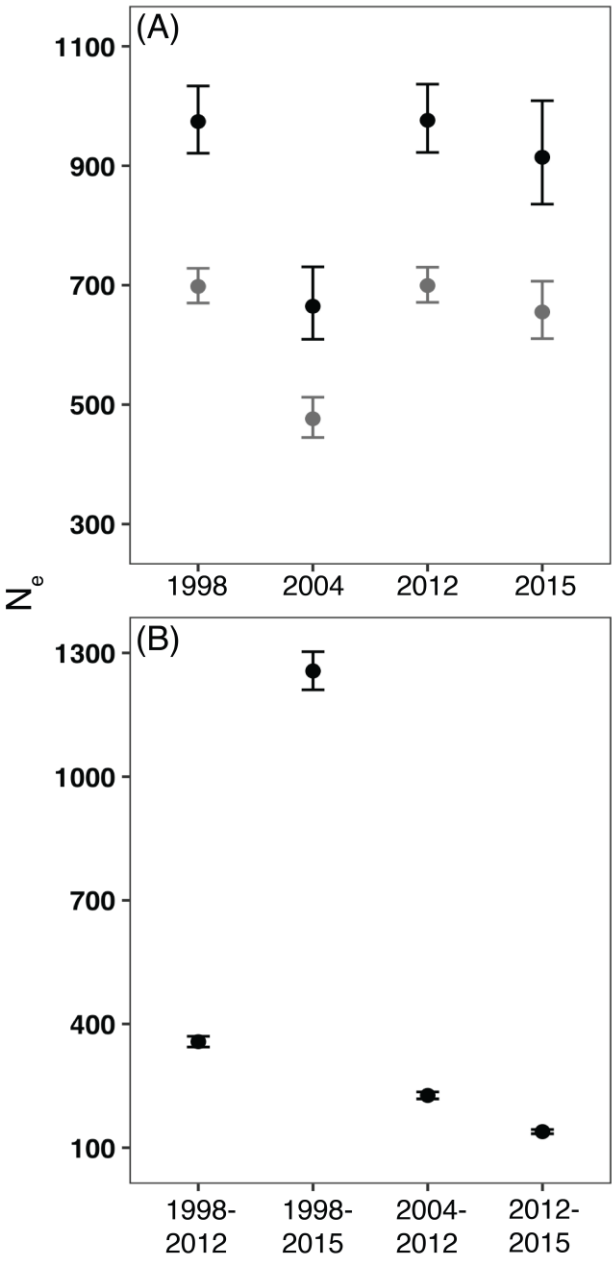


Figure 12. N_e estimates and 95% C.I.'s calculated using: (A) LD method and (B) Jorde and Ryman's temporal method (Table S7). LD method yielded two estimates: naïve estimates (grey) and bias-corrected estimates (black). The single time point or pair of time points analysed for LD estimates and temporal estimates, respectively, are indicated on the x-axis. Indefinite (negative) estimates are not shown.

Chapter 4 – Conclusion

4.1 Summary

Recent increases in available genomic resources for many exploited marine species have created new opportunities for developing effective conservation and management plans (Allendorf et al. 2010; Hoban et al. 2016). For example, genomic analyses provide increased resolution for defining population structure of marine species (Allendorf et al. 2010; Lamichhaney et al. 2012; Bourret et al. 2013a; Bradbury et al. 2013; Hemmer-Hansen et al. 2014; Milano et al. 2014; Benestan et al. 2015; Van Wyngaarden et al. 2016). A thorough understanding of the genetic structure of marine species across their geographic range is essential for developing effective conservation and management strategies. Additionally, the analysis of thousands of markers distributed across the entire genome provides insight into both adaptive and neutral processes that are important to consider when implementing management and conservation strategies (Funk et al. 2012). In this thesis, I demonstrate the power of using a genomic approach for delineating conservation units, determining the composition of mixed-stock fisheries and monitoring N_e.

The findings presented in this thesis are a result of analyzing a genomic dataset consisting of over 8,000 SNPs distributed across all 23 chromosomes of the Atlantic Cod genome. In Chapter 2, I used a genomic framework to determine that Gilbert Bay warrants status as its own conservation unit. The Gilbert Bay population displays strong genetic divergence from the offshore population at both neutral and adaptive regions of the genome. The evidence shown here suggests that Gilbert Bay provides an important component of the genetic diversity found in this species and likely contributes to the

evolutionary potential of Atlantic Cod in the face of future environmental changes. This illustrates the importance of conserving small populations like the Gilbert Bay cod population even though their relative contribution to fisheries is usually small.

In Chapter 3, I developed genomic tools to help improve the design of the Gilbert Bay MPA and inform management of the Gilbert Bay cod population. To develop a panel of SNPs to distinguish Gilbert Bay cod from offshore cod that was both effective and cost-efficient, the most informative markers were selected from a genomic dataset. The developed SNP-panel provided genetic evidence that Gilbert Bay cod are harvested outside MPA boundaries. These findings show that the current MPA boundaries are insufficient for protecting the Gilbert Bay cod from harvesting and provide insight for how the design of the MPA can be improved. In addition, a method for estimating N_e using thousands of SNP markers showed that N_e is small and provides a framework for monitoring future trends in N_e. The results of this thesis demonstrate how the implementation of genomic tools can provide valuable information for the design of MPAs and management of exploited marine fishes.

4.2 Implications

The findings presented in this thesis have direct implications for the management and conservation of the Gilbert Bay Atlantic Cod population. In Canada, conservation units are defined by COSEWIC and are called Designatable Units (DUs). In Canada, there are currently six DUs of Atlantic Cod. The Gilbert Bay cod belong to the Newfoundland and Labrador DU, however, the research presented here provides evidence that the Gilbert Bay population warrants consideration as its own DU. COSEWIC states that a population may be recognized as a DU is it both “discrete” and

"evolutionarily significant relative to other populations" (COSEWIC 2015). The strong genetic differentiation at neutral loci between the Gilbert Bay and offshore populations demonstrates the discreteness of the Gilbert Bay population. Additionally, Gilbert Bay cod are likely adapted to the local conditions in Gilbert Bay as indicated by the number of outlier loci located within or close to genes. This demonstrates that the Gilbert Bay population is an evolutionarily significant component of Atlantic Cod. Since the Gilbert Bay population meets the criteria outlined by COSEWIC, changes to how Atlantic Cod are managed in coastal Labrador may be required.

The presence of Gilbert Bay cod in fishery samples documented in this thesis will be important to consider in future management plans concerning fisheries in areas adjacent to the Gilbert Bay MPA. Changes to the timing of the fishery or to the areas targeted by fishing are actions that could help to reduce accidental harvesting of Gilbert Bay cod outside MPA boundaries. Given the declines in the abundance of Gilbert Bay cod (Morris et al. 2014) and increases in individual quotas for the cod fishery in the Gilbert Bay area (Morris and Green in-press), precautionary measures that minimize exploitation of Gilbert Bay cod will likely become increasingly important for the conservation of the Gilbert Bay population.

References

Allendorf, FW, Hohenlohe, PA, and Luikart, G. 2010. Genomics and the Future of Conservation Genetics. *Nat. Rev. Genet.* **11** (10): 697–709. doi:10.1038/nrg2844.

Allison, GW, Lubchenco, J, and Carr, MH. 1998. Marine Reserves Are Necessary but Not Sufficient for Marine Conservation. *Ecol. Appl.* **8** (1): S79–92. doi:10.1890/1051-0761(1998)8[S79:MRANBN]2.0.CO;2.

Anderson, EC, Waples, RS, and Kalinowski, ST. 2008. An Improved Method for Predicting the Accuracy of Genetic Stock Identification. *Can. J. Fish. Aquat. Sci.* **65** (7): 1475–86. doi:10.1139/F08-049.

Apostolaki, P, Milner-Gulland, EJ, McAllister, MK, and Kirkwood, GP. 2002. Modelling the Effects of Establishing a Marine Reserve for Mobile Fish Species. *Can. J. Fish. Aquat. Sci.* **59** (3): 405–15. doi:10.1139/f02-018.

Aykanat, T, Johnston, SE, Orell, P, Niemelä, E, Erkinaro, J, and Primmer, CR. 2015. Low but Significant Genetic Differentiation Underlies Biologically Meaningful Phenotypic Divergence in a Large Atlantic Salmon Population. *Mol. Ecol.* **24** (20): 5158–74. doi:10.1111/mec.13383.

Babcock, RC, Shears, NT, Alcala, AC, Barrett, NS, Edgar, GJ, Lafferty, KD, McClanahan, TR, and Russ, GR. 2010. Decadal Trends in Marine Reserves Reveal Differential Rates of Change in Direct and Indirect Effects. *Proc. Natl. Acad. Sci. U. S. A.* **107** (43): 18256–61. doi:10.1073/pnas.0908012107.

Barrett, RDH, and Hoekstra, HE. 2011. Molecular Spandrels: Tests of Adaptation at the Genetic Level. *Nat. Rev. Genet.* **12** (11): 767–80. doi:10.1038/nrg3015.

Beacham, TD, Brattey, J, Miller, KM, Khai, DL, and Withler, RE. 2002. Multiple Stock Structure of Atlantic Cod (*Gadus Morhua*) off Newfoundland and Labrador Determined from Genetic Variation. *ICES J. Mar. Sci.* **59** (4): 650–65. doi:10.1006/jmsc.2002.1253.

Beaumont, MA, and Nichols, RA. 1996. Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proc. Biol. Sci.* **263** (1377): 1619–26.

Beebee, TJC. 2009. A Comparison of Single-Sample Effective Size Estimators Using Empirical Toad (*Bufo Calamita*) Population Data: Genetic Compensation and Population Size-Genetic Diversity Correlations. *Mol. Ecol.* **18** (23): 4790–97. doi:10.1111/j.1365-294X.2009.04398.x.

Benestan, L, Gosselin, T, Perrier, C, Sainte-Marie, B, Rochette, R, and Bernatchez, L. 2015. RAD-Genotyping Reveals Fine-Scale Genetic Structuring and Provides Powerful Population Assignment in a Widely Distributed Marine Species; the American Lobster (*Homarus Americanus*). *Mol. Ecol.* **24**: 3299–3315. doi:10.1111/mec.13245.

Bentzen, P, Taggart, CT, Ruzzante, DE, and Cook, D. 1996. Microsatellite Polymorphism and the Population Structure of Atlantic Cod (*Gadus Morhua*) in the Northwest Atlantic. *Can. J. Fish. Aquat. Sci.* **4**: 2706–21.

Berg, PR, Star, B, Pampoulie, C, Sodeland, M, and Barth, JMI. 2016. Three Chromosomal Rearrangements Promote Genomic Divergence between Migratory and Stationary Ecotypes of Atlantic Cod. *Sci. Rep.* **6** (23246): 1–12. doi:10.1038/srep23246.

Botsford, LW, Micheli, F, and Hastings, A. 2003. Principles for the Design of Marine Reserves. *Ecol. Appl.* **13** (1): S25–31. http://www.jstor.org/stable/3099995.

Bourret, V, Dionne, M, Kent, MP, Lien, S, and Bernatchez, L. 2013a. Landscape Genomics in Atlantic Salmon (*Salmo Salar*): Searching for Gene-Environment Interactions Driving Local Adaptation. *Evolution* **67** (12): 3469–87. doi:10.1111/evo.12139.

Bourret, V, Kent, MP, Primmer, CR, Vasemägi, A, Karlsson, S, Hindar, K, McGinnity, P, Verspoor, E, Bernatchez, L, and Lien, S. 2013b. SNP-Array Reveals Genome-Wide Patterns of Geographical and Potential Adaptive Divergence across the Natural Range of Atlantic Salmon (*Salmo Salar*). *Mol. Ecol.* **22** (3): 532–51. doi:10.1111/mec.12003.

Bradbury, IR, Hamilton, LC, Chaput, G, Robertson, MJ, Goraguer, H, Walsh, A, Morris, V, et al. 2016. Genetic Mixed Stock Analysis of an Interceptory Atlantic Salmon Fishery in the Northwest Atlantic. *Fish. Res.* **174**: 234–44. doi:10.1016/j.fishres.2015.10.009.

Bradbury, IR, Hamilton, LC, Rafferty, S, Meerburg, D, Poole, R, Dempson, JB, Robertson, MJ, et al. 2015. Genetic Evidence of Local Exploitation of Atlantic Salmon in a Coastal Subsistence Fishery in the Northwest Atlantic. *Can. J. Fish. Aquat. Sci.* **71**: 83–95.

Bradbury, IR, Hubert, S, Higgins, B, Borza, T, Bowman, S, Paterson, IG, Snelgrove, PVR, et al. 2010. Parallel Adaptive Evolution of Atlantic Cod on Both Sides of the Atlantic Ocean in Response to Temperature. *Proc. R. Soc. B* **277** (1701): 3725–34. doi:10.1098/rspb.2010.0985.

Bradbury, IR, Hubert, S, Higgins, B, Bowman, S, Borza, T, Paterson, IG, Snelgrove, PVR, et al. 2013. Genomic Islands of Divergence and Their Consequences for the Resolution of Spatial Structure in an Exploited Marine Fish. *Evol. Appl.* **6** (3): 450–61. doi:10.1111/eva.12026.

Cáceres, A, Sindi, SS, Raphael, BJ, Cáceres, M, and González, JR. 2012. Identification of Polymorphic Inversions from Genotypes. *BMC Bioinformatics* **13** (1). BioMed Central Ltd: 28. doi:10.1186/1471-2105-13-28.

Calò, A, Muñoz, I, Pérez-Ruzafa, Á, Vergara-Chen, C, and García-Charton, JA. 2016. Spatial Genetic Structure in the Saddled Sea Bream (*Oblada Melanura* [Linnaeus, 1758]) Suggests Multi-Scaled Patterns of Connectivity between Protected and Unprotected Areas in the Western Mediterranean Sea. *Fish. Res.* **176**: 30–38. doi:10.1016/j.fishres.2015.12.001.

Chakraborty, R, and Nei, M. 1977. Bottleneck Effects on Average Heterozygosity and Genetic Distance with the Stepwise Mutation Model. *Evolution* **31** (2): 347–56. http://www.jstor.org/stable/2407757.

Conover, DO, Clarke, LM, Munch, SB, and Wagner, GN. 2006. Spatial and Temporal Scales of Adaptive Divergence in Marine Fishes and the Implications for Conservation. *J. Fish Biol.* **69** (SUPPL. C): 21–47. doi:10.1111/j.1095-8649.2006.01274.x.

COSEWIC. 2010. Assessment and Status Report on the Atlantic Cod (*Gadus Morhua*) in Canada.

COSEWIC. 2015. Guidelines for Recognizing Designatable Units. http://www.cosepac.gc.ca/default.asp?lang=En&n=DD31EAEE-1.

Costello, C, Rassweiler, A, Siegel, D, Leo, G De, Micheli, F, and Rosenberg, A. 2010. The Value of Spatial Information in MPA Network Design. *Proc. Natl. Acad. Sci.* **107** (43): 18294–99. doi:10.1073/pnas.0908057107.

Crandall, KA, Bininda-emonds, ORP, Mace, GM, and Wayne, RK. 2000. Considereing Evolurionary Processes in Conservation Biology. *Trends Ecol. Evol.* **15** (7): 290–95. doi:10.1016/S0169-5347(00)01876-0.

Deng, H. 2013. Guided Random Forest in the RRF Package. *arXiv:1306.0237v3*, 1–2. http://arxiv.org/abs/1306.0237.

DFO. 2007. Gilbert Bay Marine Protected Area Management Plan.

Do, C, Waples, RS, Peel, D, Macbeth, GM, Tillett, BJ, and Ovenden, JR. 2014. NeEstimator v2: Re-Implementation of Software for the Estimation of Contemporary Effective Population Size (N_e) from Genetic Data. *Mol. Ecol. Resour.* **14** (1): 209–14. doi:10.1111/1755-0998.12157.

Edgar, GJ. 2011. Does the Global Network of Marine Protected Areas Provide an Adequate Safety Net for Marine Biodiversity? *Aquat. Conserv. Mar. Freshw. Ecosyst.* **21** (4). John Wiley & Sons, Ltd: 313–16. doi:10.1002/aqc.1187.

Edgar, GJ, Stuart-Smith, RD, Willis, TJ, Kininmonth, S, Baker, SC, Banks, S, Barrett, NS, et al. 2014. Global Conservation Outcomes Depend on Marine Protected Areas with Five Key Features. *Nature* **506** (7487)]: 216–20. doi:10.1038/nature13022.

Evanno, G, Regnaut, S, and Goudet, J. 2005. Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study. *Mol. Ecol.* **14** (8): 2611–20. doi:10.1111/j.1365-294X.2005.02553.x.

Excoffier, L, and Lischer, HEL. 2010. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Mol. Ecol. Resour.* **10** (3): 564–67. doi:10.1111/j.1755-0998.2010.02847.x.

Feder, JL, Nosil, P, and Flaxman, SM. 2014. Assessing When Chromosomal Rearrangements Affect the Dynamics of Speciation: Implications from Computer Simulations. *Front. Genet.* **5** (295): 1–14. doi:10.3389/fgene.2014.00295.

Fenberg, PB, Caselle, JE, Claudet, J, Clemence, M, Gaines, SD, García-Charton, JA, Gonçalves, EJ, et al. 2012. The Science of European Marine Reserves: Status, Efficacy, and Future Needs. *Mar. Policy* **36** (5): 1012–21. doi:10.1016/j.marpol.2012.02.021.

Foll, M, and Gaggiotti, O. 2008. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* **180** (2): 977–93. doi:10.1534/genetics.108.092221.

Fox, CJ, Taylor, M, Dickey-Collas, M, Fossum, P, Kraus, G, Rohlf, N, Munk, P, et al. 2008. Mapping the Spawning Grounds of North Sea Cod (*Gadus Morhua*) by Direct and Indirect Means. *Proc. R. Soc. B* **275** (1642): 1543–48. doi:10.1098/rspb.2008.0201.

Funk, WC, McKay, JK, Hohenlohe, PA, and Allendorf, FW. 2012. Harnessing Genomics for Delineating Conservation Units. *Trends Ecol. Evol.* **27** (9). Elsevier Ltd: 489–96. doi:10.1016/j.tree.2012.05.012.

Gaines, SD, White, C, Carr, MH, and Palumbi, SR. 2010. Designing Marine Reserve Networks for Both Conservation and Fisheries Management. *Proc. Natl. Acad. Sci.* **107** (43): 18286–93. doi:10.1073/pnas.0906473107.

Gomez-Uchida, D, Palstra, FP, Knight, TW, and Ruzzante, DE. 2013. Contemporary Effective Population and Metapopulation Size (N_e and Meta-N_e): Comparison among Three Salmonids Inhabiting a Fragmented System and Differing in Gene Flow and Its Asymmetries. *Ecol. Evol.* **3** (3): 569–80. doi:10.1002/ece3.485.

Gosselin, T, Anderson, EC, and Bradbury, IR. 2016. Assigner: Assignment Analysis with GBS/RAD Data Using R. doi:10.5281/zenodo.46723.

Green, AL, Maypa, AP, Almany, GR, Rhodes, KL, Weeks, R, Abesamis, RA, Gleason, MG, Mumby, PJ, and White, AT. 2015. Larval Dispersal and Movement Patterns of Coral Reef Fishes, and Implications for Marine Reserve Network Design. *Biol. Rev.* **90** (4): 1215–47. doi:10.1111/brv.12155.

Green, JM, and Wroblewski, JS. 2000. Movement Patterns of Atlantic Cod in Gilbert Bay, Labrador: Evidence for Bay Residency and Spawning Site Fidelity. *J. Mar. Biol. Assoc. UK* **80** (6): 1077–85. doi:10.1017/S0025315400003143.

Guo, B, DeFaveri, J, Sotelo, G, Nair, A, and Merilä, J. 2015. Population Genomic Evidence for Adaptive Differentiation in Baltic Sea Three-Spined Sticklebacks. *BMC Biol.* **13** (1):1–18. doi:10.1186/s12915-015-0130-8.

Hardie, DC, Gillett, RM, and Hutchings, JA. 2006. The Effects of Isolation and Colonization History on the Genetic Structure of Marine-Relict Populations of Atlantic Cod (*Gadus Morhua*) in the Canadian Arctic. *Can. J. Fish. Aquat. Sci.* **63** (8): 1830–39. doi:10.1139/f06-085.

Hauser, L, Adcock, GJ, Smith, PJ, Ramiréz, JHB, and Carvalho, GR. 2002. Loss of Microsatellite Diversity and Low Effective Population Size in an Overexploited Population of New Zealand Snapper (*Pagrus Auratus*). *Proc. Natl. Acad. Sci.* **99** (18): 11742–47. doi:10.1073/pnas.172242899.

Hedrick, P. 2005. Large Variance in Reproductive Success and the N_e/N Ratio. *Evolution* **59** (7): 1596–99. doi:10.1038/438753a.

Hemmer-Hansen, J, Therkildsen, NO, Meldrup, D, and Nielsen, EE. 2014. Conserving Marine Biodiversity: Insights from Life-History Trait Candidate Genes in Atlantic Cod (*Gadus Morhua*). *Conserv. Genet.* **15** (1): 213–28. doi:10.1007/s10592-013-0532-5.

Hess, JE, Campbell, NR, Close, DA, Docker, MF, and Narum, SR. 2013. Population Genomics of Pacific Lamprey: Adaptive Variation in a Highly Dispersive Species. *Mol. Ecol.* **22** (11): 2898–2916. doi:10.1111/mec.12150.

Hilborn, R, Quinn, TP, Schindler, DE, and Rogers, DE. 2003. Biocomplexity and Fisheries Sustainability. *Proc. Natl. Acad. Sci.* **100** (11): 6564–68. doi:10.1073/pnas.1037274100.

Hoban, S, Kelley, JL, Lotterhos, KE, Antolin, MF, Bradburd, G, Lowry, DB, Poss, ML, Reed, LK, Storfer, A, and Whitlock, MC. 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am. Nat.* **188** (4): 379–97. doi:10.1086/688018.

Hoffmann, AA, and Rieseberg, LH. 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.* **39**: 21–42. doi:10.1146/annurev.ecolsys.39.110707.173532.

Hoffmann, AA, Sgrò, CM, and Weeks, AR. 2004. Chromosomal Inversion Polymorphisms and Adaptation. *Trends Ecol. Evol.* **19** (9): 482–88. doi:10.1016/j.tree.2004.06.013.

Hohenlohe, PA, Bassham, S, Etter, PD, Stiffler, N, Johnson, EA, and William, A. 2010. Population Genomics of Parallel Adaptation in Threespine Stickleback Using Sequenced RAD Tags *PLoS Genet.* **6** (2). doi:10.1371/journal.pgen.1000862.

Hollenbeck, CM, Portnoy, DS, and Gold, JR. 2016. A Method for Detecting Recent Changes in Contemporary Effective Population Size from Linkage Disequilibrium at Linked and Unlinked Loci. *Heredity* **117** (4): 207–16. doi:10.1038/hdy.2016.30.

Hubert, S, Higgins, B, Borza, T, and Bowman, S. 2010. Development of a SNP Resource and a Genetic Linkage Map for Atlantic Cod (*Gadus Morhua*). *BMC Genomics* **11** (191): 1–14. doi:10.1186/1471-2164-11-191.

Hutchings, JA. 1999. Influence of Growth and Survival Costs of Reproduction on Atlantic Cod, *Gadus Morhua*, Population Growth Rate. *Can. J. Fish. Aquat. Sci.* **56**: 1612–23. doi:10.1139/cjfas-56-9-1612.

Hutchings, JA, and Reynolds, JD. 2004. Marine Fish Population Collapses: Consequences for Recovery and Extinction Risk. *Bioscience* **54** (4): 297–309.

Hutchinson, WF, Oosterhout, C van, Rogers, SI, and Carvalho, GR. 2003. Temporal Analysis of Archived Samples Indicates Marked Genetic Changes in Declining North Sea Cod (*Gadus Morhua*). *Proc. R. Soc. B* **270** (1529): 2125–32. doi:10.1098/rspb.2003.2493.

Jombart, T, Devillard, S, Balloux, F, Falush, D, Stephens, M, Pritchard, J, Pritchard, J, et al. 2010. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. *BMC Genet.* **11** (1): 94. doi:10.1186/1471-2156-11-94.

Jorde, PE, and Ryman, N. 1995. Temporal Allele Frequency Change and Estimation of Effective Size in Populations with Overlapping Generations. *Genetics* **139** (2): 1077–90.

Keenan, K, Mcginnity, P, Cross, TF, Crozier, WW, and Prodöhl, PA. 2013. DiveRsity: An R Package for the Estimation and Exploration of Population Genetics Parameters and Their Associated Errors. *Methods Ecol. Evol.* **4** (8): 782–88. doi:10.1111/2041-210X.12067.

Kelley, JL, Brown, AP, Therkildsen, NO, and Foote, AD. 2016. The Life Aquatic: Advances in Marine Vertebrate Genomics. *Nat. Rev. Genet.* **17** (9): 523–34. doi:10.1038/nrg.2016.66.

Kininmonth, S, Beger, M, Bode, M, Peterson, E, Adams, VM, Dorfman, D, Brumbaugh, DR, and Possingham, HP. 2011. Dispersal Connectivity and Reserve Selection for Marine Conservation. *Ecol. Modell.* **222** (7): 1272–82. doi:10.1016/j.ecolmodel.2011.01.012.

Kinlan, BP, and Gaines, SD. 2003. Propagule Dispersal in Marine and Terrestrial Environments: A Community Perspective. *Ecology* **84** (8). Ecological Society of America: 2007–20. http://www.jstor.org/stable/3450025.

Kirubakaran, TG, Grove, H, Kent, MP, Sandve, SR, Baranski, M, Nome, T, Rosa, MC De, et al. 2016. Two Adjacent Inversions Maintain Genomic Differentiation between Migratory and Stationary Ecotypes of Atlantic Cod. *Mol. Ecol.* **25**: 2130–43. doi:10.1111/mec.13592.

Kjesbu, OS, Solemdal, P, Bratland, P, and Fonn, M. 1996. Variation in Annual Egg Production in Individual Captive Atlantic Cod (*Gadus Morhua*). *Can. J. Fish. Aquat. Sci.* **53** (3): 610–20. doi:10.1139/cjfas-53-3-610.

Kopelman, NM, Mayzel, J, Jakobsson, M, Rosenberg, NA, and Mayrose, I. 2015. Clumpak: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences across K. *Mol. Ecol. Resour.* doi:10.1111/1755-0998.12387.

Kuparinen, A, Hutchings, JA, and Waples, RS. 2016. Harvest-Induced Evolution and Effective Population Size. *Evol. Appl.* **9** (5): 658–72. doi:10.1111/eva.12373.

Kurlansky, M. 1997. *Cod: A Biography of the Fish That Changed the World*. New York: Walker and Company.

Lage, C, Kuhn, K, and Kornfield, I. 2004. Genetic Differentiation among Atlantic Cod (*Gadus Morhua*) from Browns Bank, Georges Bank, and Natucket Shoals. *Fish. Bull.* **102** (2): 289–97.

Lamichhaney, S, Barrio, AM, Rafati, N, Sundstrom, G, Rubin, C-J, Gilbert, ER, Berglund, J, et al. 2012. Population-Scale Sequencing Reveals Genetic Differentiation due to Local Adaptation in Atlantic Herring. *Proc. Natl. Acad. Sci.* **109** (47): 19345–50. doi:10.1073/pnas.1216128109.

Larson, WA, Seeb, JE, Pascal, CE, Templin, WD, and Seeb, LW. 2014. SNPs Identified through Genotyping-by-Sequencing Improve Genetic Stock Identification of Chinook Salmon (*Oncorhynchus Tshawytscha*) from Western Alaska. *Can. J. Fish. Aquat. Sci.* **71** (5): 698–708. doi:dx.doi.org/10.1139/cjfas-2013-0502.

Laurel, BJ, and Bradbury, IR. 2006. "Big" Concerns with High Latitude Marine Protected Areas (MPAs): Trends in Connectivity and MPA Size. *Can. J. Fish. Aquat. Sci.* **63** (12): 2603–7. doi:10.1139/f06-151.

Lemay, MA, and Russello, MA. 2015. Genetic Evidence for Ecological Divergence in Kokanee Salmon. *Mol. Ecol.* **24** (4): 798–811. doi:10.1111/mec.13066.

Lester, SE, Halpern, BS, Grorud-Colvert, K, Lubchenco, J, Ruttenberg, BI, Gaines, SD, Airame, S, and Warner, RR. 2009. Biological Effects within No-Take Marine Reserves: A Global Synthesis. *Mar. Ecol. Prog. Ser.* **384**: 33–46. doi:10.3354/meps08029.

Lien, S, Koop, BF, Sandve, SR, Miller, JR, Kent, MP, Nome, T, Hvidsten, TR, et al. 2016. The Atlantic Salmon Genome Provides Insights into Rediploidization. *Nature* **533** (7602). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 200–205. http://dx.doi.org/10.1038/nature17164.

Lotterhos, KE, and Whitlock, MC. 2014. Evaluation of Demographic History and Neutral Parameterization on the Performance of FST Outlier Tests. *Mol. Ecol.* **23** (9): 2178–92. doi:10.1111/mec.12725.

Lotterhos, KE, and Whitlock, MC. 2015. The Relative Power of Genome Scans to Detect Local Adaptation Depends on Sampling Design and Statistical Method. *Mol. Ecol.*, no. 336: 1031–46. doi:10.1111/mec.13100.

Lubchenco, J, Palumbi, SR, Gaines, SD, and Andelman, S. 2003. Plugging a Hole in the Ocean : The Emerging Science of Marine Reserves. *Ecol. Appl.* **13** (1): 3–7.

Luu, K, Bazin, E, and Blum, MGB. 2017. Pcadapt : An R Package to Perform Genome Scans for Selection Based on Principal Component Analysis. *Mol. Ecol. Resour.* **17**: 67–77. doi:10.1111/1755-0998.12592.

Martinez Barrio, A, Lamichhaney, S, Fan, G, Rafati, N, Pettersson, M, Zhang, H, Dainat, J, et al. 2016. The Genetic Basis for Ecological Adaptation of the Atlantic Herring Revealed by Genome Sequencing. Edited by Magnus Nordborg. *Elife* **5** (May). eLife Sciences Publications, Ltd: e12081. doi:10.7554/eLife.12081.

May, AW. 1966. Biology and Fishery of Atlantic Cod (*Gadus Morhua Morhua* L.) from Labrador. Memorial University, Newfoundland.

McGowen, MR, Grossman, LI, and Wildman, DE. 2012. Dolphin Genome Provides Evidence for Adaptive Evolution of Nervous System Genes and a Molecular Rate Slowdown. *Proc. R. Soc. - Biol. Sci.* **279** (1743): 3643–51. doi:10.1098/rspb.2012.0869.

Milano, I, Babbucci, M, Cariani, A, Atanassova, M, Bekkevold, D, Carvalho, GR, Espiñeira, M, et al. 2014. Outlier SNP Markers Reveal Fine-Scale Genetic Structuring across European Hake Populations (*Merluccius Merluccius*). *Mol. Ecol.* **23** (1): 118–35. doi:10.1111/mec.12568.

Moen, T, Hayes, B, Nilsen, F, Delghandi, M, Fjalestad, KT, Fevolden, S-E, Berg, PR, and Lien, S. 2008. Identification and Characterisation of Novel SNP Markers in Atlantic Cod: Evidence for Directional Selection. *BMC Genet.* **9** (18): 1–9. doi:10.1186/1471-2156-9-18.

Moffitt, EA, Wilson White, J, and Botsford, LW. 2011. The Utility and Limitations of Size and Spacing Guidelines for Designing Marine Protected Area (MPA) Networks. *Biol. Conserv.* **144** (1): 306–18. doi:10.1016/j.biocon.2010.09.008.

Moore, JS, Bourret, V, Dionne, M, Bradbury, I, O'Reilly, P, Kent, M, Chaput, G, and Bernatchez, L. 2014. Conservation Genomics of Anadromous Atlantic Salmon across Its North American Range: Outlier Loci Identify the Same Patterns of Population Structure as Neutral Loci. *Mol. Ecol.* **23**: 5680–97. doi:10.1111/mec.12972.

Morris, C, and Green, J. 2002. Biological Characteristics of a Resident Population of Atlantic Cod (*Gadus Morhua* L.) in Southern Labrador. *ICES J. Mar. Sci.* **59** (4): 666–78. doi:10.1006/jmsc.2002.1228.

Morris, CJ, and Green, JM. 2014. MPA Regulations Should Incorporate Adaptive management—The Case of Gilbert Bay Labrador Atlantic Cod (*Gadus Morhua*). *Mar. Policy* **49**: 20–28. doi:10.1016/j.marpol.2014.03.025.

Morris, CJ, Green, JM, and Simms, JM. 2003. Abundance of Resident Atlantic Cod in Gilbert Bay, Labrador, Based on Mark-Recapture, Sampling Cat per Unit Effort and Commercial Tag Return Data Collected from 1998 to 2002. *Can. Sci. Advis. Secr.*

Morris, CJ, Green, JM, Snelgrove, PVR, Pennell, CJ, and Ollerhead, LM. 2014. Temporal and Spatial Migration of Atlantic Cod (*Gadus Morhua*) inside and Outside a Marine Protected Area and Evidence for the Role of Prior Experience in Homing. *Can. J. Fish. Aquat. Sci.* **71** (11): 1704–12. http://www.nrcresearchpress.com/doi/abs/10.1139/cjfas-2014-0036.

Munguía-Vega, A, Sáenz-Arroyo, A, Greenley, AP, Espinoza-Montes, JA, Palumbi, SR, Rossetto, M, and Micheli, F. 2015. Marine Reserves Help Preserve Genetic Diversity after Impacts Derived from Climate Variability: Lessons from the Pink Abalone in Baja California. *Glob. Ecol. Conserv.* **4**: 264–76. doi:10.1016/j.gecco.2015.07.005.

Narum, SR, and Hess, JE. 2011. Comparison of F_{ST} Outlier Tests for SNP Loci under Selection. *Mol. Ecol. Resour.* **11** (Suppl. 1): 184–94. doi:10.1111/j.1755-0998.2011.02987.x.

Nei, M, Maruyama, T, and Chakraborty, R. 1975. The Bottleneck Effect and Genetic Variability in Populations. *Evolution* **29** (1): 1–10. http://www.jstor.org/stable/2407137.

Nielsen, JL, Arrizabalaga, H, Fragoso, NM, Hobday, AJ, Lutcavage, ME, and Sibert, JR. 2009. Tagging and Tracking of Marine Animals with Electronics Devices. *Rev. Methods Technol. Fish Biol. Fish.* doi:10.1007/s13398-014-0173-7.2.

Nunney, L. 1996. The Influence of Variation in Female Fecundity on Effective Population Size. *Biol. J. Linn. Soc.* doi:10.1006/bijl.1996.0072.

Nunney, L, and Elam, DR. 1994. Estimating the Effective Population Size of Conserved Populations. *Conserv. Biol.* **8** (1): 175–84. doi:10.1046/j.1523-1739.1994.08010175.x.

O'Leary, DB, Coughlan, J, Dillane, E, McCarthy, TV, and Cross, TF. 2007. Microsatellite Variation in Cod *Gadus Morhua* throughout Its Geographic Range. *J. Fish Biol.* **70** (Suppl. C): 310–35. doi:10.1111/j.1095-8649.2007.01451.x.

Obokata, J, Ohme, M, and Hayashida, N. 1991. Nucleotide Sequence of a cDNA Clone Encoding a Putative Glycine-Rich Protein of 19.7 kDa in Nicotiana Sylvestris. *Plant Mol. Biol.* **17** (4): 953–55. doi:10.1007/BF00037080.

Ovenden, JR, Peel, D, Street, R, Courtney, AJ, Hoyle, SD, Peel, SL, and Podlich, H. 2007. The Genetic Effective and Adult Census Size of an Australian Population of Tiger Prawns (*Penaeus Esculentus*). *Mol. Ecol.* **16** (1): 127–38. doi:10.1111/j.1365-294X.2006.03132.x.

Palstra, FP, and Ruzzante, DE. 2008. Genetic Estimates of Contemporary Effective Population Size: What Can They Tell Us about the Importance of Genetic Stochasticity for Wild Population Persistence? *Mol. Ecol.* **17** (15): 3428–47. doi:10.1111/j.1365-294X.2008.03842.x.

Palstra, FP, and Ruzzante, DE. 2011. Demographic and Genetic Factors Shaping Contemporary Metapopulation Effective Size and Its Empirical Estimation in Salmonid Fish. *Heredity* **107** (5): 444–55. doi:10.1038/hdy.2011.31.

Palumbi, S. 2003. Population Genetics, Demographic Connectivity, and the Design of Marine Reserves. *Ecol. Appl.* **13** (1): 146–58. http://www.esajournals.org/doi/abs/10.1890/1051-0761(2003)013%255B0146:PGDCAT%255D2.0.CO%253B2.

Pauly, D, Christensen, V, Guenette, S, Pitcher, TJ, Sumaila, UR, Walters, CJ, Watson, R, and Zeller, D. 2002. Towards Sustainability in World Fisheries. *Nature* **418** (6898): 689–95. http://dx.doi.org/10.1038/nature01017.

Pittman, SJ, Monaco, ME, Friedlander, AM, Legare, B, Nemeth, RS, Kendall, MS, Poti, M, Clark, RD, Wedding, LM, and Caldow, C. 2014. Fish with Chips: Tracking Reef Fish Movements to Evaluate Size and Connectivity of Caribbean Marine Protected Areas. *PLoS One* **9** (5). doi:10.1371/journal.pone.0096028.

Planes, S, Jones, GP, and Thorrold, SR. 2009. Larval Dispersal Connects Fish Populations in a Network of Marine Protected Areas. *Proc. Natl. Acad. Sci.* **106** (14): 5693–97. doi:10.1073/pnas.0808007106.

Poulsen, NA, Nielsen, EE, Schierup, MH, Loeschcke, V, and Grønkjaer, P. 2006. Long-Term Stability and Effective Population Size in North Sea and Baltic Sea Cod (*Gadus Morhua*). *Mol. Ecol.* **15** (2): 321–31. doi:10.1111/j.1365-294X.2005.02777.x.

Pritchard, JK, and Rienzo, A Di. 2010. Adaptation - Not by Sweeps Alone. *Nat. Rev. Genet.* **11** (10): 665–67. doi:10.1038/nrg2880.

Pritchard, JK, Stephens, M, and Donnelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155** (2): 945–59. doi:10.1111/j.1471-8286.2007.01758.x.

Pujolar, JM, Jacobsen, MW, Bekkevold, D, Lobón-Cervià, J, Jónsson, B, Bernatchez, L, and Hansen, MM. 2015. Signatures of Natural Selection between Life Cycle Stages Separated by Metamorphosis in European Eel. *BMC Genomics* **16** (1): 600. doi:10.1186/s12864-015-1754-3.

Purcell, S, Neale, B, Todd-Brown, K, Thomas, L, Ferreira, MAR, Bender, D, Maller, J, et al. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**: 559–75. doi:10.1086/519795.

Reimand, J, Arak, T, Adler, P, Kolberg, L, Reisberg, S, Peterson, H, and Vilo, J. 2016. g:Profiler—a Web Server for Functional Interpretation of Gene Lists (2016 Update). *Nucleic Acids Res.*, 1–7. doi:10.1093/nar/gkw199.

Robichaud, D, and Rose, G a. 2004. Migratory Behaviour and Range in Atlantic Cod: Inference from a Century of Tagging. *Fish Fish.* **5** (3): 185–214. doi:10.1111/j.1467-2679.2004.00141.x.

Ruzzante, DE, Taggart, CT, and Cook, D. 1996. Spatial and Temporal Variation in the Genetic Composition of a Larval Cod (*Gadus Morhua*) Aggregation: Cohort Contribution and Genetic Stability. *Can. J. Fish. Aquat. Sci.* **53** (12): 2695–2705. doi:10.1139/f96-235.

Ruzzante, DE, Taggart, CT, and Cook, D. 1998. A Nuclear DNA Basis for Shelf-and Bank-Scale Population Structure in Northwest Atlantic Cod (*Gadus Morhua*): Labrador to Georges Bank. *Mol. Ecol.* **7** (12): 1663–1680. doi:10.1046/j.1365-294x.1998.00497.x.

Ruzzante, DE, Taggart, CT, Cook, D, and Goddard, SV. 1997. Genetic Differentiation Between Inshore and Offshore Atlantic Cod (*Gadus Morhua*) Off Newfoundland: Microsatellite DNA Variation and Antifreeze Level. *Can. J. Fish. Aquat. Sci.* **54**: 2700–2708. doi:10.1139/cjfas-53-3-634.

Ruzzante, DE, Taggart, CT, Lang, S, and Cook, D. 2000b. Mixed-Stock Analysis of Atlantic Cod near the Gulf of St. Lawrence Based on Microsatellite DNA. *Ecol. Appl.* **10** (4): 1090–1109. http://www.esajournals.org/doi/abs/10.1890/1051-0761(2000)010%255B1090:MSAOAC%255D2.0.CO%253B2.

Ruzzante, DE, Wroblewski, JS, Taggart, CT, Smedbol, RK, Cook, D, and Goddard, S V. 2000a. Bay-Scale Population Structure in Coastal Atlantic Cod in Labrador and Newfoundland, Canada. *J. Fish Biol.* **56**: 431–47. doi:10.1006/jfbi.1999.1168.

Ryman, N. 1997. Minimizing Adverse Effects of Fish Culture: Understanding the Genetics of Populations with Overlapping Generations. *ICES J. Mar. Sci.* **54** (6): 1149–59. doi:10.1016/S1054-3139(97)80021-5.

Saarinen, E V., Austin, JD, and Daniels, JC. 2010. Genetic Estimates of Contemporary Effective Population Size in an Endangered Butterfly Indicate a Possible Role for Genetic Compensation. *Evol. Appl.* **3** (1): 28–39. doi:10.1111/j.1752-4571.2009.00096.x.

Salomon, AK, Ruesink, JL, and DeWreede, RE. 2006. Population Viability, Ecological Processes and Biodiversity: Valuing Sites for Reserve Selection. *Biol. Conserv.* **128** (1): 79–92. doi:10.1016/j.biocon.2005.09.018.

Sambrook, J, Maniatis, T, and Fritsch, EF. 1987. *Molecular Cloning : A Laboratory Manual*. 2nd ed. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press.

Savolainen, O, Lascoux, M, and Merilä, J. 2013. Ecological Genomics of Local Adaptation. *Nat. Rev. Genet.* **14**: 807–20. doi:10.1038/nrg3522.

Schindler, DE, Hilborn, R, Chasco, B, Boatright, CP, Quinn, TP, Rogers, LA, and Webster, MS. 2010. Population Diversity and the Portfolio Effect in an Exploited Species. *Nature* **465** (7298): 609–12. doi:10.1038/nature09060.

Shanks, AL, Grantham, BA, and Carr, MH. 2003. Propagule Dispersal Distance and the Size and Spacing of Marine Reserves. *Ecol. Appl.* **13** (1 Supplement): S159-169. doi:10.1890/1051-0761(2003)013[0159:PDDATS]2.0.CO;2.

Sodeland, M, Jorde, PE, Lien, S, Jentoft, S, Berg, PR, Grove, H, Kent, MP, Arnyasi, M, Olsen, EM, and Knutsen, H. 2016. "Islands of Divergence" in the Atlantic Cod Genome Represent Polymorphic Chromosomal Rearrangements. *Genome Biol. Evol.* **8** (4): 1012–22. doi:10.1093/gbe/evw057.

Stanley, RRR, Jeffery, NW, Wringe, BF, DiBacco, C, and Bradbury, IR. 2016. Genepopedit: A Simple and Flexible Tool for Manipulating Multilocus Molecular Data in R. *Mol. Ecol. Resour.* **17** (1): 12–18. doi:10.1111/1755-0998.12569.

Star, B, Nederbragt, AJ, Jentoft, S, Grimholt, U, Malmstrøm, M, Gregers, TF, Rounge, TB, et al. 2011a. The Genome Sequence of Atlantic Cod Reveals a Unique Immune System. *Nature* **477**: 207–10. doi:10.1038/nature10342.

Stinchcombe, JR, and Hoekstra, HE. 2007. Combining Population Genomics and Quantitative Genetics: Finding the Genes Underlying Ecologically Important Traits. *Heredity* **100** (2): 158–70. doi:10.1038/sj.hdy.6800937.

Syms, C, and Carr, MH. 2001. Marine Protected Areas : Evaluating MPA Effectiveness in an Uncertain World. Santa Cruz. doi:10.2307/2261484.

Therkildsen, NO, Nielsen, EE, Swain, DP, and Pedersen, JS. 2010. Large Effective Population Size and Temporal Genetic Stability in Atlantic Cod (*Gadus Morhua*) in the Southern Gulf of St. Lawrence. *Can. J. Fish. Aquat. Sci.* **67**: 1585–95. doi:10.1139/F10-084.

USFWS. 1996. Policy Regarding the Recognition of Distinct Vertebrate Population Segments Under the Endangered Species Act. *Fed. Regist.* **61** (26): 4722–25.

Vilas, A, Pérez-Figueroa, A, and Caballero, A. 2012. A Simulation Study on the Performance of Differentiation-Based Methods to Detect Selected Loci Using Linked Neutral Markers. *J. Evol. Biol.* **25** (7): 1364–76. doi:10.1111/j.1420-9101.2012.02526.x.

Wang, J. 2001. A Pseudo-Likelihood Method for Estimating Effective Population Size from Temporally Spaced Samples. *Genet. Res.* **78** (3): 243–57. doi:10.1017/S0016672301005286.

Waples, RK, Larson, WA, and Waples, RS. 2016. Estimating Contemporary Effective Population Size in Non-Model Species Using Linkage Disequilibrium across Thousands of Loci. *Heredity* **117**: 233–440. doi:10.1038/hdy.2016.60.

Waples, RS. 1989. A Generalized Approach for Estimating Effective Population Size from Temporal Changes in Allele Frequency. *Genetics* **121** (2): 379–91.

Waples, RS. 1991. Pacific Salmon, Oncorhynchus Spp., and the Definition of 'species' under the Endangered Species Act. *Mar. Fish. Rev.* **53** (3): 11–22.

Waples, RS. 1995. Evolutionarily Significant Units and the Conservation of Biological Diversity under the Endangered Species. *Am. Fish. Soc. Symp.* **17**: 8–27.

Waples, RS. 2002. Evaluating the Effect of Stage-Specific Survivorship on the Ne/N Ratio. *Mol. Ecol.* **11** (6): 1029–37. isi:000175841000007.

Waples, RS. 2006. A Bias Correction for Estimates of Effective Population Size Based on Linkage Disequilibrium at Unlinked Gene Loci. *Conserv. Genet.* **7** (2): 167–84. doi:10.1007/s10592-005-9100-y.

Waples, RS, and England, PR. 2011. Estimating Contemporary Effective Population Size on the Basis of Linkage Disequilibrium in the Face of Migration. *Genetics* **189** (2): 633–44. doi:10.1534/genetics.111.132233.

Weir, BS, and Cockerham, CC. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38** (6): 1358–70. doi:10.2307/2408641.

Wyngaarden, M Van, Snelgrove, PVR, DiBacco, C, Hamilton, LC, Rodríguez-Ezpeleta, N, Jeffery, NW, Stanley, RRE, and Bradbury, IR. 2016. Identifying Patterns of Dispersal, Connectivity, and Selection in the Sea Scallop, *Placopecten Magellanicus,* Using RAD-Seq Derived SNPs. *Evol. Appl.* **10** (1): 102–17. doi:10.1111/eva.12432.

Appendix A – Supplementary Tables

Table S1. Number of ordered SNPs per linkage group, their average F_{ST} and the corresponding standard deviation.

Linkage group	# of SNPs	Average F_{ST}	Standard deviation
1	396	0.1102	0.1427
2	335	0.0446	0.0588
3	372	0.0543	0.0685
4	405	0.0455	0.0629
5	352	0.0443	0.0589
6	314	0.0544	0.0694
7	333	0.0474	0.0648
8	337	0.0484	0.0672
9	297	0.0417	0.0578
10	284	0.0437	0.0530
11	326	0.0467	0.0637
12	321	0.0538	0.0683
13	277	0.0476	0.0606
14	309	0.0496	0.0593
15	310	0.0515	0.0699
16	403	0.0487	0.0626
17	272	0.0371	0.0445
18	287	0.0446	0.0677
19	224	0.0479	0.0587
20	287	0.0453	0.0619
21	322	0.0490	0.0719
22	289	0.0492	0.0666
23	266	0.0408	0.0595

Table S2. Gene annotations with IDs and distance to annotated gene for each highly divergent locus (106 SNPs). SNPs detected as outliers, included based on Fst threshold or located within a detected inversion are marked with an asterisk (*).

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:132312_899	1	Laminin, alpha 3 [Source:HGNC Symbol;Acc:HGNC:6483]	LAMA3	ENSGMOG00000006641	>5000		*	
Gdist:18775_786	1	Novel gene		ENSGMOG00000018058	>5000		*	*
Gdist:207918_923	1	Zinc finger protein 704 [Source:ZFIN;Acc:ZDB-GENE-060503-775]	ZNf704	ENSGMOG00000003488	>5000	*	*	*
Gdist:296871_513	1	Neuropeptides B/W receptor 2 [Source:HGNC Symbol;Acc:HGNC:4530]	NPBWR2	ENSGMOG00000020377	>5000		*	*
Gdist:475298_1154	1			ENSGMOG00000004015	>5000	*	*	*
Gdist:59772_256	1	Cyclin-dependent kinase 16 [Source:HGNC Symbol;Acc:HGNC:8749]	CDK16	ENSGMOG00000004816	>5000		*	*
Gdist:88524_1775	1	elastase 2 [Source:ZFIN;Acc:ZDB-GENE-041117-1]	ELA2	ENSGMOG00000015725	>5000		*	*
Gdist:88581_4571	1	PR domain containing 2, with ZNF domain a [Source:ZFIN;Acc:ZDB-GENE-081104-139]	PRDM2A	ENSGMOG00000015917	4626	*	*	*
Gdist:28852_1491	1	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3bl [Source:ZFIN;Acc:ZDB-GENE-030131-2956]	SEMA3Bl	ENSGMOG00000016820	3481		*	*
Gdist:34799_4530	1	Snail homolog 1b (Drosophila) [Source:ZFIN;Acc:ZDB-GENE-980526-514]	SNAI1B	ENSGMOG00000000127	3270	*	*	*
Gdist:114493_3343	1	Cadherin 4, retinal [Source:ZFIN;Acc:ZDB-GENE-991207-1]	CDH4	ENSGMOG00000002481	2979	*	*	
Gdist:446136_208	1	Cold shock domain containing E1, RNA-binding [Source:ZFIN;Acc:ZDB-GENE-	CSDE1	ENSGMOG00000012731	2846	*	*	*

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
		030131-8623]						
LD:475244_2952	1	Tensin like C1 domain containing phosphatase a [Source:ZFIN;Acc:ZDB-GENE-090312-163]	TENC1A	ENSGMOG00000004659	2638	*	*	*
Gdist:173307_1239	1	si:rp71-17i16.4 [Source:ZFIN;Acc:ZDB-GENE-090312-215]		ENSGMOG00000017780	2475		*	*
LD:475260_465	1	Novel gene		ENSGMOG00000004807	2411		*	*
Gdist:58189_308	1	agrin [Source:HGNC Symbol;Acc:329]	AGRN	ENSGMOG00000015163	2023		*	
Gdist:475055_4232	1	Cytochrome P450, family 4, subfamily V, polypeptide 7 [Source:ZFIN;Acc:ZDB-GENE-061103-88]	CYP4V7	ENSGMOG00000015493	1401	*	*	*
Gdist:71253_609	1	opiate receptor-like 1 [Source:HGNC Symbol;Acc:8155]	OPRL1	ENSGMOG00000013222	1244	*	*	*
LD:475269_1389	1	glutathione synthetase [Source:HGNC Symbol;Acc:4624]	GSS	ENSGMOG00000003980	872		*	*
Gdist:475157_3548	1	Cilia and flagella associated protein 74 [Source:HGNC Symbol;Acc:HGNC:29368]	CFAP74	ENSGMOG00000011643	765	*	*	*
LD:475245_192	1	Transmembrane protein 106C [Source:HGNC Symbol;Acc:HGNC:28775]	TMEM106C	ENSGMOG00000004729	671	*	*	*
Gdist:293954_469	1	Vitamin D (1,25- dihydroxyvitamin D3) receptor [Source:HGNC Symbol;Acc:HGNC:12679]	VDR	ENSGMOG00000012720	521		*	*
Gdist:357995_2161	1	RNA binding motif protein 15 [Source:ZFIN;Acc:ZDB-GENE-041008-192]	RBM15	ENSGMOG00000019941	477	*	*	*

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:114574_6556	1	Protein kinase C binding protein 1, like [Source:ZFIN;Acc:ZDB-GENE-030131-2199]	PRKCBP1L	ENSGMOG00000014051	396	*	*	*
Gdist:114729_8581	1	Glutamate receptor, metabotropic 4 [Source:HGNC Symbol;Acc:HGNC:4596]	GRM4	ENSGMOG00000010065	259	*	*	*
Gdist:206925_768	1	ArfGAP with coiled-coil, ankyrin repeat and PH domains 3a [Source:ZFIN;Acc:ZDB-GENE-030131-1711]	ACAP3A	ENSGMOG00000006685	75		*	*
Gdist:101097_1339	1	phosphofructokinase, muscle [Source:HGNC Symbol;Acc:8877]	PFKM	ENSGMOG00000008238	72	*	*	*
Gdist:326080_1142	1	ATPase, class II, type 9A [Source:HGNC Symbol;Acc:HGNC:13540]	ATP9A	ENSGMOG00000003768	70	*	*	*
Gdist:136060_418	1	Potassium voltage-gated channel, KQT-like subfamily, member 2a [Source:ZFIN;Acc:ZDB-GENE-080220-37]	KCNQ2A	ENSGMOG00000006540	68	*	*	*
GENE:475255_2349	1	Family with sequence similarity 50, member A [Source:ZFIN;Acc:ZDB-GENE-050417-110]	FAM50A	ENSGMOG00000004783	59	*	*	*
Gdist:114351_4596	1	WNK lysine deficient protein kinase 2 [Source:HGNC Symbol;Acc:HGNC:14542]	WNK2	ENSGMOG00000017405	48	*	*	*
Gdist:110622_4217	1	Inositol 1,4,5-triphosphate receptor, type 3 [Source:ZFIN;Acc:ZDB-GENE-070605-1]	ITPR3	ENSGMOG00000012713	34	*	*	*
Gdist:114396_5097	1	Nucleoporin 210 [Source:ZFIN;Acc:ZDB-GENE-050208-132]	NUP210	ENSGMOG00000017069	18	*	*	*
Gdist:474992_2283	1	Xeroderma pigmentosum, complementation group C [Source:ZFIN;Acc:ZDB-GENE-030131-8461]	XPC	ENSGMOG00000019246	17		*	*

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:136085_7350	1	Ral GTPase activating protein, beta subunit (non-catalytic) [Source:ZFIN;Acc:ZDB-GENE-130530-945]	RALGAPB	ENSGMOG00000006453	0	*	*	*
Gdist:167858_229	1	Plexin A3 [Source:ZFIN;Acc:ZDB-GENE-070613-1]	PLXNA3	ENSGMOG00000010058	0	*	*	*
Gdist:326103_183	1	Microtubule associated monooxygenase, calponin and LIM domain containing 1	MICAL1	ENSGMOG00000013037	0	*	*	*
Gdist:358014_6981	1	Potassium voltage-gated channel, Shaw-related subfamily, member 4 [Source:ZFIN;Acc:ZDB-GENE-060503-773]	KCNC4	ENSGMOG00000001282	0	*	*	*
Gdist:475221_2078	1	Keratin 18 [Source:ZFIN;Acc:ZDB-GENE-030411-6]	KRT18	ENSGMOG00000000117	0		*	*
Gdist:475340_466	1	Vacuolar protein sorting 13D (yeast) [Source:ZFIN;Acc:ZDB-GENE-050912-1]	VPS13D	ENSGMOG00000008937	0	*	*	*
Gdist:88548_7921	1	elastase 2 [Source:ZFIN;Acc:ZDB-GENE-041117-1]	ELA2	ENSGMOG00000015725	0	*	*	*
NS:100935_2238	1	Peroxisome biogenesis factor 10 [Source:ZFIN;Acc:ZDB-GENE-041010-71]	PEX10	ENSGMOG00000001654	0	*	*	*
NS:101105_7972	1	Structural maintenance of chromosomes 5 [Source:ZFIN;Acc:ZDB-GENE-061013-288]	SMC5	ENSGMOG00000008193	0	*	*	*
NS:101105_8442	1	Structural maintenance of chromosomes 5 [Source:ZFIN;Acc:ZDB-GENE-061013-288]	SMC5	ENSGMOG00000008193	0	*	*	*
NS:114346_2416	1	Novel gene		ENSGMOG00000017574	0	*	*	*
NS:114623_127	1	glycerol-3-phosphate dehydrogenase 1 (soluble) [Source:HGNC Symbol;Acc:4455]	GPD1A	ENSGMOG00000009648	0		*	

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
NS:129362_255	1	solute carrier family 35, member C2 [Source:ZFIN;Acc:ZDB-GENE-030131-2202]	SLC35C2	ENSGMOG00000017176	0		*	*
NS:132298_4379	1	phosphoinositide-3-kinase, catalytic, gamma polypeptide [Source:ZFIN;Acc:ZDB-GENE-040426-2532]	PIK3CG	ENSGMOG00000000049	0		*	*
NS:326148_6716	1	arginine-glutamic acid dipeptide (RE) repeats a [Source:ZFIN;Acc:ZDB-GENE-060718-1]	REREA	ENSGMOG00000003783	0	*	*	*
NS:357968_4575	1	thymopoietin [Source:HGNC Symbol;Acc:11875]	TMPOB	ENSGMOG00000019696	0	*	*	*
NS:475081_867	1	nephronophthisis 4 [Source:HGNC Symbol;Acc:19104]	NPHP4	ENSGMOG00000015728	0		*	*
NS:58174_843	1	agrin [Source:HGNC Symbol;Acc:329]	AGRN	ENSGMOG00000015163	0	*	*	*
GENE:276602_1629	2	SRY (sex determining region Y)-box 10 [Source:HGNC Symbol;Acc:11190]	SOX10	ENSGMOG00000009956	1080		*	
Gdist:172424_7303	2	Ring finger protein 213 [Source:HGNC Symbol;Acc:HGNC:14539]	RNF213	ENSGMOG00000005695	0		*	
NS:549045_140	2	gypsy retrotransposon integrase 1 [Source:HGNC Symbol;Acc:25959]	GIN1	ENSGMOG00000013636	0		*	
Gdist:201955_983	3	protocadherin-10-like isoform		ENSGMOG00000013454	>5000		*	
Gdist:392165_819	3			ENSGMOG00000014327	>5000		*	
Gdist:355802_143	3			ENSGMOG00000000574	1082		*	
NS:172052_3413	3	Lipoic acid synthetase [Source:ZFIN;Acc:ZDB-GENE-040426-1528]	LIAS	ENSGMOG00000001015	0		*	

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:344968_120	4	myb sant-like dna-binding domain-containing protein 2-like		ENSGMOG00000001589	238		*	
Gdist:90457_1467	4	inhibin beta b chain-like		ENSGMOG00000015845	0		*	
NS:100187_1282	4	prostaglandin E receptor 2 (subtype EP2), 53kDa [Source:HGNC Symbol;Acc:9594]		ENSGMOG00000016940	0		*	
Gdist:243631_4426	5			ENSGMOG00000000442	1738		*	
Gdist:95514_681	5						*	
Gdist:317541_988	6	T-box 1 [Source:HGNC Symbol;Acc:11592]	TBX1	ENSGMOG00000018613	>5000		*	
Gdist:339093_1304	6			ENSGMOG00000000890	>5000		*	
NS:71406_6814	6	dipeptidyl-peptidase 7 [Source:HGNC Symbol;Acc:14892]	DPP7	ENSGMOG00000014562	0		*	
Gdist:273172_69	6						*	
Gdist:459709_9284	7	B-cell translocation gene 3 [Source:ZFIN;Acc:ZDB-GENE-031113-19]	BTG3	ENSGMOG00000012227	4554		*	
NS:51949_161	7	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 1 [Source:ZFIN;Acc:ZDB-GENE-091116-69]	SLC7A1	ENSGMOG00000001821	0		*	
NS:92360_7249	7	listerin E3 ubiquitin protein ligase 1 [Source:HGNC Symbol;Acc:HGNC:13082]	LTN1	ENSGMOG00000000609	0		*	
Gdist:36011_3217	8	nucleolin-like isoform		ENSGMOG00000017831	0		*	
Gdist:579902_284	9	Ras protein-specific guanine nucleotide-releasing factor 1 [Source:HGNC Symbol;Acc:HGNC:9875]	RASGRF1	ENSGMOG00000016711	52		*	

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:306771_5207	10			ENSGMOG00000012746	427		*	
Gdist:140375_2410	11			ENSGMOG00000009859	2263		*	
Gdist:590304_342	11			ENSGMOG00000009125	1304		*	
Gdist:110193_6520	11	n-acetyltransferase 14		ENSGMOG00000004097	0		*	
NS:31932_2344	11	multicilin		ENSGMOG00000001405	0		*	
Gdist:103517_1096	12			ENSGMOG00000009604	>5000		*	
Gdist:137649_334	12	Novel gene		ENSGMOG00000015125	>5000		*	
Gdist:137622_751	12	selectin E [Source:HGNC Symbol;Acc:10718]	SELE	ENSGMOG00000015254	1578		*	
NS:444316_5262	12	nuclear receptor ror-beta-like	RORCA	ENSGMOG00000000779	0		*	
Gdist:286305_5489	13			ENSGMOG00000007891	>5000		*	
Gdist:391093_808	14	chromosome 15 open reading frame 43 [Source:HGNC Symbol;Acc:HGNC:28520]	C15orf43	ENSGMOG00000013260	>5000		*	
Gdist:59175_1492	15			ENSGMOG00000004422	1815		*	
NS:185397_2861	15	cytochrome c oxidase assembly homolog 15 (yeast) [Source:ZFIN;Acc:ZDB-GENE-040426-955]	COX15	ENSGMOG00000018132	0		*	
Gdist:364062_1292	15	None		None			*	
Gdist:91983_228	16			ENSGMOG00000008806	4553		*	
Gdist:188129_6651	16			ENSGMOG00000011710	1465		*	

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
Gdist:124586_1425	16	zinc finger fyve domain-containing protein 1-like	ZFYVE1	ENSGMOG00000014672	922		*	
Gdist:67946_1117	18	uroporphyrinogen III synthase [Source:ZFIN;Acc:ZDB-GENE-040323-2]	UROS	ENSGMOG00000002467	>5000		*	
Gdist:163151_1281	18	ankyrin repeat and SOCS box-containing 12a [Source:ZFIN;Acc:ZDB-GENE-030131-1647]	ASB12A	ENSGMOG00000016736	0		*	
Gdist:266583_1995	18	neurexin 1 [Source:HGNC Symbol;Acc:HGNC:8008]	NRXN1	ENSGMOG00000014265	0	*	*	
Gdist:561858_1727	18	RNA binding protein, fox-1 homolog 3b [Source:ZFIN;Acc:ZDB-GENE-091204-46]	RBFOX3L	ENSGMOG00000019304	0		*	
CAN:rs11905 6128	19	patatin-like phospholipase domain containing 8 [Source:ZFIN;Acc:ZDB-GENE-070705-553]	PNPLA8	ENSGMOG00000008090	0		*	
GENE:658523_3790	20	SRY (sex determining region Y)-box 1a [Source:ZFIN;Acc:ZDB-GENE-040718-186]	SOX1A	ENSGMOG00000020606	0		*	
NS:658569_622	20	inositol polyphosphate-4-phosphatase type I Aa [Source:ZFIN;Acc:ZDB-GENE-060312-5]	INPP4AA	ENSGMOG00000012587	0		*	
GENE:398989_637	21	estrogen receptor 2 (ER beta) [Source:HGNC Symbol;Acc:3468]	ESR2	ENSGMOG00000019003	958		*	
GENE:398988_3229	21	estrogen receptor 2 (ER beta) [Source:HGNC Symbol;Acc:3468]	ESR2	ENSGMOG00000019003	0		*	
Gdist:535143_690	21	none	none				*	
Gdist:126647_343	22			ENSGMOG00000002054	>5000		*	
NS:420549_8523	22	RecQ protein-like 4 [Source:HGNC Symbol;Acc:9949]	RECQL4	ENSGMOG00000001340	0		*	
Gdist:278657_	23	acyl-coenzyme a thioesterase 4-like	ACOT4	ENSGMOG00	>5000		*	

Locus	Linkage group	Gene annotation	HGNC	Gene ID	Distance to gene	Outlier	Fst > 0.3	SNP in inversion
455 Gdist:401774_764 CAN:rs119055284 MITO:872249_6730	23			000005776 ENSGMOG00000009264	>5000		*	

Table S3. Results for enrichment analysis of highly divergent SNPs. *P*-value, term ID, type (BP=biological process, MF= molecular function, CC= cellular component), and name are listed for each GO term.

P-value	Term ID	Term type	Term name
1	GO:0008150	BP	biological process
1	GO:0051179	BP	localization
1	GO:0033036	BP	macromolecule localization
1	GO:0008104	BP	protein localization
1	GO:0051674	BP	localization of cell
1	GO:0051234	BP	establishment of localization
1	GO:0006810	BP	transport
1	GO:0071705	BP	nitrogen compound transport
1	GO:0043574	BP	peroxisomal transport
1	GO:0071702	BP	organic substance transport
1	GO:0015931	BP	nucleobase-containing compound transport
1	GO:0055085	BP	transmembrane transport
1	GO:0007034	BP	vacuolar transport
1	GO:0006811	BP	ion transport
1	GO:0006812	BP	cation transport
1	GO:0072511	BP	divalent inorganic cation transport
1	GO:0030001	BP	metal ion transport
1	GO:0070838	BP	divalent metal ion transport
1	GO:0006816	BP	calcium ion transport
1	GO:0015672	BP	monovalent inorganic cation transport
1	GO:0006813	BP	potassium ion transport
1	GO:0006820	BP	anion transport
1	GO:0015711	BP	organic anion transport
1	GO:0034220	BP	ion transmembrane transport
1	GO:0098660	BP	inorganic ion transmembrane transport
1	GO:0098655	BP	cation transmembrane transport
1	GO:0098662	BP	inorganic cation transmembrane transport
1	GO:0070588	BP	calcium ion transmembrane transport
1	GO:0098656	BP	anion transmembrane transport
1	GO:0045184	BP	establishment of protein localization
1	GO:0015031	BP	protein transport
1	GO:0071806	BP	protein transmembrane transport
1	GO:0051641	BP	cellular localization
1	GO:0051668	BP	localization within membrane
1	GO:0070727	BP	cellular macromolecule localization
1	GO:0034613	BP	cellular protein localization
1	GO:0033365	BP	protein localization to organelle
1	GO:0072665	BP	protein localization to vacuole

P-value	Term ID	Term type	Term name
1	GO:0072594	BP	establishment of protein localization to organelle
1	GO:0072666	BP	establishment of protein localization to vacuole
1	GO:0034067	BP	protein localization to Golgi apparatus
1	GO:0051649	BP	establishment of localization in cell
1	GO:0046907	BP	intracellular transport
1	GO:0006886	BP	intracellular protein transport
1	GO:0065002	BP	intracellular protein transmembrane transport
1	GO:0006605	BP	protein targeting
1	GO:0006623	BP	protein targeting to vacuole
1	GO:0002376	BP	immune system process
1	GO:0022610	BP	biological adhesion
1	GO:0007155	BP	cell adhesion
1	GO:0098609	BP	cell-cell adhesion
1	GO:0098742	BP	cell-cell adhesion via plasma-membrane adhesion molecules
1	GO:0007156	BP	homophilic cell adhesion via plasma membrane adhesion molecules
1	GO:0032502	BP	developmental process
1	GO:0048856	BP	anatomical structure development
1	GO:0009888	BP	tissue development
1	GO:0061448	BP	connective tissue development
1	GO:0007498	BP	mesoderm development
1	GO:0060429	BP	epithelium development
1	GO:0009653	BP	anatomical structure morphogenesis
1	GO:0048646	BP	anatomical structure formation involved in morphogenesis
1	GO:0023052	BP	signaling
1	GO:0071840	BP	cellular component organization or biogenesis
1	GO:0044085	BP	cellular component biogenesis
1	GO:0070271	BP	protein complex biogenesis
1	GO:0050896	BP	response to stimulus
1	GO:0009605	BP	response to external stimulus
1	GO:0009719	BP	response to endogenous stimulus
1	GO:0071495	BP	cellular response to endogenous stimulus
1	GO:0042221	BP	response to chemical
1	GO:1901700	BP	response to oxygen-containing compound
1	GO:1901698	BP	response to nitrogen compound
1	GO:0060359	BP	response to ammonium ion
1	GO:0010033	BP	response to organic substance
1	GO:0033993	BP	response to lipid
1	GO:0014070	BP	response to organic cyclic compound
1	GO:1905144	BP	response to acetylcholine
1	GO:0070848	BP	response to growth factor

P-value	Term ID	Term type	Term name
1	GO:0071774	BP	response to fibroblast growth factor
1	GO:0009725	BP	response to hormone
1	GO:0048545	BP	response to steroid hormone
1	GO:0006950	BP	response to stress
1	GO:0040011	BP	locomotion
1	GO:0042330	BP	taxis
1	GO:0006935	BP	chemotaxis
1	GO:0050919	BP	negative chemotaxis
1	GO:0009987	BP	cellular process
1	GO:0051716	BP	cellular response to stimulus
1	GO:0070887	BP	cellular response to chemical stimulus
1	GO:1901701	BP	cellular response to oxygen-containing compound
1	GO:1901699	BP	cellular response to nitrogen compound
1	GO:0071242	BP	cellular response to ammonium ion
1	GO:1905145	BP	cellular response to acetylcholine
1	GO:0071310	BP	cellular response to organic substance
1	GO:0071363	BP	cellular response to growth factor stimulus
1	GO:0044344	BP	cellular response to fibroblast growth factor stimulus
1	GO:0071396	BP	cellular response to lipid
1	GO:0071407	BP	cellular response to organic cyclic compound
1	GO:0032870	BP	cellular response to hormone stimulus
1	GO:0071383	BP	cellular response to steroid hormone stimulus
1	GO:0033554	BP	cellular response to stress
1	GO:0006974	BP	cellular response to DNA damage stimulus
1	GO:0007154	BP	cell communication
1	GO:0016043	BP	cellular component organization
1	GO:0043933	BP	macromolecular complex subunit organization
1	GO:0071822	BP	protein complex subunit organization
1	GO:0006325	BP	chromatin organization
1	GO:0022607	BP	cellular component assembly
1	GO:0065003	BP	macromolecular complex assembly
1	GO:0034622	BP	cellular macromolecular complex assembly
1	GO:0006461	BP	protein complex assembly
1	GO:0043623	BP	cellular protein complex assembly
1	GO:0017004	BP	cytochrome complex assembly
1	GO:0008535	BP	respiratory chain complex IV assembly
1	GO:0051259	BP	protein oligomerization
1	GO:0051260	BP	protein homooligomerization
1	GO:0006996	BP	organelle organization
1	GO:0007031	BP	peroxisome organization
1	GO:0072662	BP	protein localization to peroxisome

P-value	Term ID	Term type	Term name
1	GO:0072663	BP	establishment of protein localization to peroxisome
1	GO:0006625	BP	protein targeting to peroxisome
1	GO:0051276	BP	chromosome organization
1	GO:0071103	BP	DNA conformation change
1	GO:0032392	BP	DNA geometric change
1	GO:0032508	BP	DNA duplex unwinding
1	GO:0061024	BP	membrane organization
1	GO:0008152	BP	metabolic process
1	GO:0044238	BP	primary metabolic process
1	GO:0009056	BP	catabolic process
1	GO:0006807	BP	nitrogen compound metabolic process
1	GO:0071704	BP	organic substance metabolic process
1	GO:1901135	BP	carbohydrate derivative metabolic process
1	GO:1901360	BP	organic cyclic compound metabolic process
1	GO:0043170	BP	macromolecule metabolic process
1	GO:0010467	BP	gene expression
1	GO:0043412	BP	macromolecule modification
1	GO:0098732	BP	macromolecule deacylation
1	GO:0016569	BP	covalent chromatin modification
1	GO:0019538	BP	protein metabolic process
1	GO:0036211	BP	protein modification process
1	GO:0006508	BP	proteolysis
1	GO:1901575	BP	organic substance catabolic process
1	GO:1901136	BP	carbohydrate derivative catabolic process
1	GO:1901564	BP	organonitrogen compound metabolic process
1	GO:0005975	BP	carbohydrate metabolic process
1	GO:0009058	BP	biosynthetic process
1	GO:1901576	BP	organic substance biosynthetic process
1	GO:1901362	BP	organic cyclic compound biosynthetic process
1	GO:1901566	BP	organonitrogen compound biosynthetic process
1	GO:0042398	BP	cellular modified amino acid biosynthetic process
1	GO:0009059	BP	macromolecule biosynthetic process
1	GO:0044237	BP	cellular metabolic process
1	GO:0006793	BP	phosphorus metabolic process
1	GO:0019637	BP	organophosphate metabolic process
1	GO:0046434	BP	organophosphate catabolic process
1	GO:0090407	BP	organophosphate biosynthetic process
1	GO:0006796	BP	phosphate-containing compound metabolic process
1	GO:0052646	BP	alditol phosphate metabolic process
1	GO:0006072	BP	glycerol-3-phosphate metabolic process

P-value	Term ID	Term type	Term name
1	GO:0046168	BP	glycerol-3-phosphate catabolic process
1	GO:0016310	BP	phosphorylation
1	GO:0051186	BP	cofactor metabolic process
1	GO:0006732	BP	coenzyme metabolic process
1	GO:0046483	BP	heterocycle metabolic process
1	GO:0034641	BP	cellular nitrogen compound metabolic process
1	GO:0043603	BP	cellular amide metabolic process
1	GO:0006518	BP	peptide metabolic process
1	GO:0044260	BP	cellular macromolecule metabolic process
1	GO:0044267	BP	cellular protein metabolic process
1	GO:0006464	BP	cellular protein modification process
1	GO:0018065	BP	protein-cofactor linkage
1	GO:0009249	BP	protein lipoylation
1	GO:0070647	BP	protein modification by small protein conjugation or removal
1	GO:0032446	BP	protein modification by small protein conjugation
1	GO:0016567	BP	protein ubiquitination
1	GO:0035601	BP	protein deacylation
1	GO:0006476	BP	protein deacetylation
1	GO:0006468	BP	protein phosphorylation
1	GO:0016570	BP	histone modification
1	GO:0016575	BP	histone deacetylation
1	GO:0006575	BP	cellular modified amino acid metabolic process
1	GO:0006790	BP	sulfur compound metabolic process
1	GO:0006749	BP	glutathione metabolic process
1	GO:0006725	BP	cellular aromatic compound metabolic process
1	GO:0033013	BP	tetrapyrrole metabolic process
1	GO:0006778	BP	porphyrin-containing compound metabolic process
1	GO:0006139	BP	nucleobase-containing compound metabolic process
1	GO:0090304	BP	nucleic acid metabolic process
1	GO:0016070	BP	RNA metabolic process
1	GO:0016071	BP	mRNA metabolic process
1	GO:0006396	BP	RNA processing
1	GO:0008380	BP	RNA splicing
1	GO:0000375	BP	RNA splicing, via transesterification reactions
1	GO:0000377	BP	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile
1	GO:0006397	BP	mRNA processing
1	GO:0000398	BP	mRNA splicing, via spliceosome
1	GO:0006259	BP	DNA metabolic process
1	GO:0006310	BP	DNA recombination

P-value	Term ID	Term type	Term name
1	GO:0071139	BP	resolution of recombination intermediates
1	GO:0006281	BP	DNA repair
1	GO:0006284	BP	base-excision repair
1	GO:0006289	BP	nucleotide-excision repair
1	GO:0006302	BP	double-strand break repair
1	GO:0006298	BP	mismatch repair
1	GO:0000725	BP	recombinational repair
1	GO:0000724	BP	double-strand break repair via homologous recombination
1	GO:0044249	BP	cellular biosynthetic process
1	GO:0044272	BP	sulfur compound biosynthetic process
1	GO:0044271	BP	cellular nitrogen compound biosynthetic process
1	GO:0043604	BP	amide biosynthetic process
1	GO:0043043	BP	peptide biosynthetic process
1	GO:0019184	BP	nonribosomal peptide biosynthetic process
1	GO:0006750	BP	glutathione biosynthetic process
1	GO:0051188	BP	cofactor biosynthetic process
1	GO:0009108	BP	coenzyme biosynthetic process
1	GO:0018130	BP	heterocycle biosynthetic process
1	GO:0034645	BP	cellular macromolecule biosynthetic process
1	GO:0006260	BP	DNA replication
1	GO:0019438	BP	aromatic compound biosynthetic process
1	GO:0034654	BP	nucleobase-containing compound biosynthetic process
1	GO:0032774	BP	RNA biosynthetic process
1	GO:0097659	BP	nucleic acid-templated transcription
1	GO:0006351	BP	transcription, DNA-templated
1	GO:0033014	BP	tetrapyrrole biosynthetic process
1	GO:0006779	BP	porphyrin-containing compound biosynthetic process
1	GO:0032501	BP	multicellular organismal process
1	GO:0003008	BP	system process
1	GO:0050877	BP	neurological system process
1	GO:0007600	BP	sensory perception
1	GO:0019233	BP	sensory perception of pain
1	GO:0040007	BP	growth
1	GO:0065007	BP	biological regulation
1	GO:0065009	BP	regulation of molecular function
1	GO:0044093	BP	positive regulation of molecular function
1	GO:0050790	BP	regulation of catalytic activity
1	GO:0051336	BP	regulation of hydrolase activity
1	GO:0043087	BP	regulation of GTPase activity
1	GO:0043085	BP	positive regulation of catalytic activity

P-value	Term ID	Term type	Term name
1	GO:0051345	BP	positive regulation of hydrolase activity
1	GO:0043547	BP	positive regulation of GTPase activity
1	GO:0090630	BP	activation of GTPase activity
1	GO:0065008	BP	regulation of biological quality
1	GO:0090066	BP	regulation of anatomical structure size
1	GO:0051235	BP	maintenance of location
1	GO:0045185	BP	maintenance of protein location
1	GO:0044699	BP	single-organism process
1	GO:0098602	BP	single organism cell adhesion
1	GO:0016337	BP	single organismal cell-cell adhesion
1	GO:0044763	BP	single-organism cellular process
1	GO:0006928	BP	movement of cell or subcellular component
1	GO:0048870	BP	cell motility
1	GO:0016477	BP	cell migration
1	GO:0001667	BP	ameboidal-type cell migration
1	GO:0008078	BP	mesodermal cell migration
1	GO:0050900	BP	leukocyte migration
1	GO:0097529	BP	myeloid leukocyte migration
1	GO:0097530	BP	granulocyte migration
1	GO:1990266	BP	neutrophil migration
1	GO:0060326	BP	cell chemotaxis
1	GO:0030595	BP	leukocyte chemotaxis
1	GO:0071621	BP	granulocyte chemotaxis
1	GO:0030593	BP	neutrophil chemotaxis
1	GO:0044802	BP	single-organism membrane organization
1	GO:1902589	BP	single-organism organelle organization
1	GO:0032535	BP	regulation of cellular component size
1	GO:0008361	BP	regulation of cell size
1	GO:0007049	BP	cell cycle
1	GO:0022402	BP	cell cycle process
1	GO:0007059	BP	chromosome segregation
1	GO:0098813	BP	nuclear chromosome segregation
1	GO:0000819	BP	sister chromatid segregation
1	GO:0007062	BP	sister chromatid cohesion
1	GO:0016049	BP	cell growth
1	GO:0030030	BP	cell projection organization
1	GO:1902578	BP	single-organism localization
1	GO:0044765	BP	single-organism transport
1	GO:0006818	BP	hydrogen transport
1	GO:0015992	BP	proton transport
1	GO:1902600	BP	hydrogen ion transmembrane transport
1	GO:0015849	BP	organic acid transport

P-value	Term ID	Term type	Term name
1	GO:0046942	BP	carboxylic acid transport
1	GO:0006865	BP	amino acid transport
1	GO:1903825	BP	organic acid transmembrane transport
1	GO:1905039	BP	carboxylic acid transmembrane transport
1	GO:0003333	BP	amino acid transmembrane transport
1	GO:1901264	BP	carbohydrate derivative transport
1	GO:0015780	BP	nucleotide-sugar transport
1	GO:0015781	BP	pyrimidine nucleotide-sugar transport
0.405	GO:0015786	BP	UDP-glucose transport
1	GO:0071804	BP	cellular potassium ion transport
1	GO:0071805	BP	potassium ion transmembrane transport
1	GO:0017038	BP	protein import
1	GO:1902580	BP	single-organism cellular localization
1	GO:1902582	BP	single-organism intracellular transport
1	GO:0044743	BP	intracellular protein transmembrane import
1	GO:0016558	BP	protein import into peroxisome matrix
1	GO:0072657	BP	protein localization to membrane
1	GO:0043113	BP	receptor clustering
1	GO:0051651	BP	maintenance of location in cell
1	GO:0032507	BP	maintenance of protein location in cell
1	GO:0045053	BP	protein retention in Golgi apparatus
1	GO:0044710	BP	single-organism metabolic process
1	GO:0044281	BP	small molecule metabolic process
1	GO:0006082	BP	organic acid metabolic process
1	GO:0043436	BP	oxoacid metabolic process
1	GO:0019752	BP	carboxylic acid metabolic process
1	GO:0032787	BP	monocarboxylic acid metabolic process
1	GO:0042440	BP	pigment metabolic process
1	GO:0046148	BP	pigment biosynthetic process
1	GO:0042168	BP	heme metabolic process
0.405	GO:0046160	BP	heme a metabolic process
1	GO:0044711	BP	single-organism biosynthetic process
1	GO:0044283	BP	small molecule biosynthetic process
1	GO:0016053	BP	organic acid biosynthetic process
1	GO:0046394	BP	carboxylic acid biosynthetic process
1	GO:0072330	BP	monocarboxylic acid biosynthetic process
1	GO:0006783	BP	heme biosynthetic process
0.405	GO:0006784	BP	heme a biosynthetic process
1	GO:0055114	BP	oxidation-reduction process
1	GO:0006629	BP	lipid metabolic process
1	GO:0044255	BP	cellular lipid metabolic process
1	GO:0030258	BP	lipid modification

P-value	Term ID	Term type	Term name
1	GO:0046834	BP	lipid phosphorylation
1	GO:0046486	BP	glycerolipid metabolic process
1	GO:0006644	BP	phospholipid metabolic process
1	GO:0006650	BP	glycerophospholipid metabolic process
1	GO:0046488	BP	phosphatidylinositol metabolic process
1	GO:0046854	BP	phosphatidylinositol phosphorylation
1	GO:0006631	BP	fatty acid metabolic process
0.808	GO:0009106	BP	lipoate metabolic process
1	GO:0008610	BP	lipid biosynthetic process
1	GO:0045017	BP	glycerolipid biosynthetic process
1	GO:0008654	BP	phospholipid biosynthetic process
1	GO:0046474	BP	glycerophospholipid biosynthetic process
0.552	GO:0006661	BP	phosphatidylinositol biosynthetic process
0.05	GO:0036092	BP	phosphatidylinositol-3-phosphate biosynthetic process
1	GO:0006633	BP	fatty acid biosynthetic process
0.808	GO:0009107	BP	lipoate biosynthetic process
1	GO:0044707	BP	single-multicellular organism process
1	GO:0044700	BP	single organism signaling
1	GO:0007267	BP	cell-cell signaling
1	GO:0099536	BP	synaptic signaling
1	GO:0099537	BP	trans-synaptic signaling
1	GO:0098916	BP	anterograde trans-synaptic signaling
1	GO:0007268	BP	chemical synaptic transmission
1	GO:0007270	BP	neuron-neuron synaptic transmission
1	GO:0035249	BP	synaptic transmission, glutamatergic
1	GO:0198738	BP	cell-cell signaling by wnt
1	GO:0043473	BP	pigmentation
1	GO:0048066	BP	developmental pigmentation
1	GO:0044767	BP	single-organism developmental process
1	GO:0007164	BP	establishment of tissue polarity
1	GO:0048729	BP	tissue morphogenesis
1	GO:0002009	BP	morphogenesis of an epithelium
1	GO:0060026	BP	convergent extension
1	GO:0001738	BP	morphogenesis of a polarized epithelium
1	GO:0001736	BP	establishment of planar polarity
1	GO:0060322	BP	head development
1	GO:0048589	BP	developmental growth
1	GO:0060560	BP	developmental growth involved in morphogenesis
1	GO:0003401	BP	axis elongation
1	GO:0060028	BP	convergent extension involved in axis elongation
1	GO:0070121	BP	Kupffer's vesicle development

P-value	Term ID	Term type	Term name
1	GO:0048869	BP	cellular developmental process
1	GO:0032989	BP	cellular component morphogenesis
1	GO:0032990	BP	cell part morphogenesis
1	GO:0000902	BP	cell morphogenesis
1	GO:0048858	BP	cell projection morphogenesis
1	GO:0030154	BP	cell differentiation
1	GO:0048863	BP	stem cell differentiation
1	GO:0050931	BP	pigment cell differentiation
1	GO:0030318	BP	melanocyte differentiation
1	GO:0048468	BP	cell development
1	GO:0000904	BP	cell morphogenesis involved in differentiation
1	GO:0048864	BP	stem cell development
1	GO:0048588	BP	developmental cell growth
1	GO:0007275	BP	multicellular organism development
1	GO:0035295	BP	tube development
1	GO:0035239	BP	tube morphogenesis
1	GO:0048731	BP	system development
1	GO:0048880	BP	sensory system development
1	GO:0048925	BP	lateral line system development
1	GO:0072359	BP	circulatory system development
1	GO:0001655	BP	urogenital system development
1	GO:0061008	BP	hepaticobiliary system development
1	GO:0048513	BP	animal organ development
1	GO:0048732	BP	gland development
1	GO:0001889	BP	liver development
1	GO:0060485	BP	mesenchyme development
1	GO:0048762	BP	mesenchymal cell differentiation
1	GO:0014031	BP	mesenchymal cell development
1	GO:0014033	BP	neural crest cell differentiation
1	GO:0014032	BP	neural crest cell development
1	GO:0001755	BP	neural crest cell migration
1	GO:0048368	BP	lateral mesoderm development
1	GO:0060029	BP	convergent extension involved in organogenesis
1	GO:0007399	BP	nervous system development
1	GO:0048483	BP	autonomic nervous system development
1	GO:0022008	BP	neurogenesis
1	GO:0042063	BP	gliogenesis
1	GO:0010001	BP	glial cell differentiation
1	GO:0021782	BP	glial cell development
1	GO:0048699	BP	generation of neurons
1	GO:0030182	BP	neuron differentiation
1	GO:0046530	BP	photoreceptor cell differentiation

P-value	Term ID	Term type	Term name
1	GO:0048666	BP	neuron development
1	GO:0048667	BP	cell morphogenesis involved in neuron differentiation
1	GO:0042461	BP	photoreceptor cell development
1	GO:0031175	BP	neuron projection development
1	GO:0061564	BP	axon development
1	GO:0048812	BP	neuron projection morphogenesis
1	GO:0007409	BP	axonogenesis
1	GO:1990138	BP	neuron projection extension
1	GO:0048675	BP	axon extension
1	GO:0097485	BP	neuron projection guidance
1	GO:1902284	BP	neuron projection extension involved in neuron projection guidance
1	GO:0007411	BP	axon guidance
1	GO:0048846	BP	axon extension involved in axon guidance
1	GO:0001501	BP	skeletal system development
1	GO:0051216	BP	cartilage development
1	GO:0055123	BP	digestive system development
1	GO:0072001	BP	renal system development
1	GO:0007417	BP	central nervous system development
1	GO:0048709	BP	oligodendrocyte differentiation
1	GO:0007422	BP	peripheral nervous system development
1	GO:0014037	BP	Schwann cell differentiation
1	GO:0014044	BP	Schwann cell development
1	GO:0048484	BP	enteric nervous system development
1	GO:0009790	BP	embryo development
1	GO:0009792	BP	embryo development ending in birth or egg hatching
1	GO:0043009	BP	chordate embryonic development
1	GO:0060037	BP	pharyngeal system development
1	GO:0007389	BP	pattern specification process
1	GO:0003002	BP	regionalization
1	GO:0009952	BP	anterior/posterior pattern specification
1	GO:0009799	BP	specification of symmetry
1	GO:0009855	BP	determination of bilateral symmetry
1	GO:0007368	BP	determination of left/right symmetry
1	GO:0071910	BP	determination of liver left/right asymmetry
1	GO:0003140	BP	determination of left/right asymmetry in lateral mesoderm
1	GO:0031016	BP	pancreas development
1	GO:0035469	BP	determination of pancreatic left/right asymmetry
1	GO:0007423	BP	sensory organ development
1	GO:0001654	BP	eye development

P-value	Term ID	Term type	Term name
1	GO:0043010	BP	camera-type eye development
1	GO:0060041	BP	retina development in camera-type eye
1	GO:0043583	BP	ear development
1	GO:0048598	BP	embryonic morphogenesis
1	GO:0007369	BP	gastrulation
1	GO:0042074	BP	cell migration involved in gastrulation
1	GO:0021675	BP	nerve development
1	GO:0021545	BP	cranial nerve development
1	GO:0009887	BP	animal organ morphogenesis
1	GO:0090596	BP	sensory organ morphogenesis
1	GO:0048592	BP	eye morphogenesis
1	GO:0048593	BP	camera-type eye morphogenesis
1	GO:0001754	BP	eye photoreceptor cell differentiation
1	GO:0042462	BP	eye photoreceptor cell development
1	GO:0048645	BP	animal organ formation
1	GO:0060536	BP	cartilage morphogenesis
1	GO:0048565	BP	digestive tract development
1	GO:0048546	BP	digestive tract morphogenesis
1	GO:0071907	BP	determination of digestive tract left/right asymmetry
1	GO:0007507	BP	heart development
1	GO:0003007	BP	heart morphogenesis
1	GO:0060914	BP	heart formation
1	GO:0061371	BP	determination of heart left/right asymmetry
1	GO:0060973	BP	cell migration involved in heart development
1	GO:0003318	BP	cell migration to the midline involved in heart development
1	GO:0060974	BP	cell migration involved in heart formation
1	GO:0003260	BP	cardioblast migration
1	GO:0060975	BP	cardioblast migration to the midline involved in heart field formation
0.405	GO:0003261	BP	cardiac muscle progenitor cell migration to the midline involved in heart field formation
1	GO:0048568	BP	embryonic organ development
1	GO:0048562	BP	embryonic organ morphogenesis
1	GO:0042471	BP	ear morphogenesis
1	GO:0001822	BP	kidney development
1	GO:0060993	BP	kidney morphogenesis
1	GO:0048793	BP	pronephros development
1	GO:0072114	BP	pronephros morphogenesis
1	GO:0048839	BP	inner ear development
1	GO:0060872	BP	semicircular canal development
1	GO:0071599	BP	otic vesicle development

P-value	Term ID	Term type	Term name
1	GO:0042472	BP	inner ear morphogenesis
1	GO:0048752	BP	semicircular canal morphogenesis
1	GO:0071600	BP	otic vesicle morphogenesis
1	GO:0048840	BP	otolith development
1	GO:0007420	BP	brain development
1	GO:0030900	BP	forebrain development
1	GO:0021537	BP	telencephalon development
1	GO:0021988	BP	olfactory lobe development
1	GO:0035844	BP	cloaca development
1	GO:0021772	BP	olfactory bulb development
1	GO:0060042	BP	retina morphogenesis in camera-type eye
1	GO:0050789	BP	regulation of biological process
1	GO:0032879	BP	regulation of localization
1	GO:0051049	BP	regulation of transport
1	GO:0034762	BP	regulation of transmembrane transport
1	GO:0043269	BP	regulation of ion transport
1	GO:0034765	BP	regulation of ion transmembrane transport
1	GO:0051239	BP	regulation of multicellular organismal process
1	GO:0050794	BP	regulation of cellular process
1	GO:0010646	BP	regulation of cell communication
1	GO:0051128	BP	regulation of cellular component organization
1	GO:0031344	BP	regulation of cell projection organization
1	GO:1902275	BP	regulation of chromatin organization
1	GO:0051270	BP	regulation of cellular component movement
1	GO:0007165	BP	signal transduction
1	GO:1903831	BP	signal transduction involved in cellular response to ammonium ion
1	GO:0095500	BP	acetylcholine receptor signaling pathway
1	GO:0035556	BP	intracellular signal transduction
1	GO:0019932	BP	second-messenger-mediated signaling
1	GO:0048016	BP	inositol phosphate-mediated signaling
1	GO:0048017	BP	inositol lipid-mediated signaling
1	GO:0048015	BP	phosphatidylinositol-mediated signaling
1	GO:0007264	BP	small GTPase mediated signal transduction
1	GO:0007265	BP	Ras protein signal transduction
1	GO:0007266	BP	Rho protein signal transduction
1	GO:0007186	BP	G-protein coupled receptor signaling pathway
1	GO:0007218	BP	neuropeptide signaling pathway
1	GO:0007187	BP	G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger
1	GO:0007188	BP	adenylate cyclase-modulating G-protein coupled receptor signaling pathway

P-value	Term ID	Term type	Term name
1	GO:0007193	BP	adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway
1	GO:0038003	BP	opioid receptor signaling pathway
1	GO:0007213	BP	G-protein coupled acetylcholine receptor signaling pathway
1	GO:0030522	BP	intracellular receptor signaling pathway
1	GO:0007166	BP	cell surface receptor signaling pathway
1	GO:0007215	BP	glutamate receptor signaling pathway
1	GO:0007216	BP	G-protein coupled glutamate receptor signaling pathway
1	GO:0007196	BP	adenylate cyclase-inhibiting G-protein coupled glutamate receptor signaling pathway
1	GO:1905114	BP	cell surface receptor signaling pathway involved in cell-cell signaling
1	GO:0016055	BP	Wnt signaling pathway
1	GO:0035567	BP	non-canonical Wnt signaling pathway
1	GO:0060070	BP	canonical Wnt signaling pathway
1	GO:0007167	BP	enzyme linked receptor protein signaling pathway
1	GO:0007169	BP	transmembrane receptor protein tyrosine kinase signaling pathway
1	GO:0008543	BP	fibroblast growth factor receptor signaling pathway
1	GO:0007224	BP	smoothened signaling pathway
0.748	GO:0071526	BP	semaphorin-plexin signaling pathway
1	GO:0009755	BP	hormone-mediated signaling pathway
1	GO:0043401	BP	steroid hormone mediated signaling pathway
1	GO:0048583	BP	regulation of response to stimulus
1	GO:0032101	BP	regulation of response to external stimulus
1	GO:0090287	BP	regulation of cellular response to growth factor stimulus
1	GO:0050793	BP	regulation of developmental process
1	GO:0022603	BP	regulation of anatomical structure morphogenesis
1	GO:1905330	BP	regulation of morphogenesis of an epithelium
1	GO:0022604	BP	regulation of cell morphogenesis
1	GO:0045595	BP	regulation of cell differentiation
1	GO:0060284	BP	regulation of cell development
1	GO:0010769	BP	regulation of cell morphogenesis involved in differentiation
1	GO:2000026	BP	regulation of multicellular organismal development
0.679	GO:2000027	BP	regulation of organ morphogenesis
1	GO:0090175	BP	regulation of establishment of planar polarity
1	GO:0060071	BP	Wnt signaling pathway, planar cell polarity pathway
1	GO:0051960	BP	regulation of nervous system development

P-value	Term ID	Term type	Term name
1	GO:0050767	BP	regulation of neurogenesis
1	GO:0045664	BP	regulation of neuron differentiation
0.405	GO:0046532	BP	regulation of photoreceptor cell differentiation
0.405	GO:0042478	BP	regulation of eye photoreceptor cell development
1	GO:0010975	BP	regulation of neuron projection development
1	GO:0050770	BP	regulation of axonogenesis
1	GO:0023051	BP	regulation of signaling
1	GO:0009966	BP	regulation of signal transduction
1	GO:1902531	BP	regulation of intracellular signal transduction
1	GO:0051056	BP	regulation of small GTPase mediated signal transduction
1	GO:0046578	BP	regulation of Ras protein signal transduction
1	GO:0035023	BP	regulation of Rho protein signal transduction
1	GO:0008589	BP	regulation of smoothened signaling pathway
1	GO:0030111	BP	regulation of Wnt signaling pathway
1	GO:2000050	BP	regulation of non-canonical Wnt signaling pathway
0.808	GO:2000095	BP	regulation of Wnt signaling pathway, planar cell polarity pathway
1	GO:0060828	BP	regulation of canonical Wnt signaling pathway
1	GO:0040036	BP	regulation of fibroblast growth factor receptor signaling pathway
1	GO:0050804	BP	modulation of synaptic transmission
1	GO:0051966	BP	regulation of synaptic transmission, glutamatergic
1	GO:0040012	BP	regulation of locomotion
1	GO:0050920	BP	regulation of chemotaxis
1	GO:1902667	BP	regulation of axon guidance
1	GO:2000145	BP	regulation of cell motility
1	GO:0030334	BP	regulation of cell migration
1	GO:0019222	BP	regulation of metabolic process
1	GO:0080090	BP	regulation of primary metabolic process
1	GO:0060255	BP	regulation of macromolecule metabolic process
1	GO:0051246	BP	regulation of protein metabolic process
1	GO:0010468	BP	regulation of gene expression
1	GO:0051171	BP	regulation of nitrogen compound metabolic process
1	GO:0031323	BP	regulation of cellular metabolic process
1	GO:0032268	BP	regulation of cellular protein metabolic process
1	GO:0031399	BP	regulation of protein modification process
1	GO:0090311	BP	regulation of protein deacetylation
1	GO:0031056	BP	regulation of histone modification
1	GO:0031063	BP	regulation of histone deacetylation

P-value	Term ID	Term type	Term name
1	GO:0019219	BP	regulation of nucleobase-containing compound metabolic process
1	GO:0051252	BP	regulation of RNA metabolic process
1	GO:0009889	BP	regulation of biosynthetic process
1	GO:0031326	BP	regulation of cellular biosynthetic process
1	GO:0010556	BP	regulation of macromolecule biosynthetic process
1	GO:2000112	BP	regulation of cellular macromolecule biosynthetic process
1	GO:2001141	BP	regulation of RNA biosynthetic process
1	GO:1903506	BP	regulation of nucleic acid-templated transcription
1	GO:0006355	BP	regulation of transcription, DNA-templated
1	GO:0051090	BP	regulation of sequence-specific DNA binding transcription factor activity
1	GO:0040008	BP	regulation of growth
1	GO:0048638	BP	regulation of developmental growth
1	GO:0001558	BP	regulation of cell growth
1	GO:0061387	BP	regulation of extent of cell growth
1	GO:0030516	BP	regulation of axon extension
1	GO:0048841	BP	regulation of axon extension involved in axon guidance
1	GO:0030155	BP	regulation of cell adhesion
1	GO:0022407	BP	regulation of cell-cell adhesion
1	GO:0048519	BP	negative regulation of biological process
1	GO:0040013	BP	negative regulation of locomotion
1	GO:0048523	BP	negative regulation of cellular process
1	GO:0051271	BP	negative regulation of cellular component movement
1	GO:0010648	BP	negative regulation of cell communication
1	GO:0051129	BP	negative regulation of cellular component organization
1	GO:0031345	BP	negative regulation of cell projection organization
1	GO:0007162	BP	negative regulation of cell adhesion
1	GO:0022408	BP	negative regulation of cell-cell adhesion
1	GO:0051093	BP	negative regulation of developmental process
1	GO:0045596	BP	negative regulation of cell differentiation
1	GO:0010721	BP	negative regulation of cell development
1	GO:0010771	BP	negative regulation of cell morphogenesis involved in differentiation
1	GO:0048585	BP	negative regulation of response to stimulus
1	GO:0032102	BP	negative regulation of response to external stimulus
1	GO:0050922	BP	negative regulation of chemotaxis
1	GO:0051241	BP	negative regulation of multicellular organismal process

P-value	Term ID	Term type	Term name
1	GO:0051961	BP	negative regulation of nervous system development
1	GO:0050768	BP	negative regulation of neurogenesis
1	GO:0045665	BP	negative regulation of neuron differentiation
1	GO:0010977	BP	negative regulation of neuron projection development
1	GO:0050771	BP	negative regulation of axonogenesis
1	GO:1902668	BP	negative regulation of axon guidance
1	GO:0023057	BP	negative regulation of signaling
1	GO:0009968	BP	negative regulation of signal transduction
1	GO:0045879	BP	negative regulation of smoothened signaling pathway
1	GO:0030178	BP	negative regulation of Wnt signaling pathway
1	GO:0090090	BP	negative regulation of canonical Wnt signaling pathway
1	GO:0045926	BP	negative regulation of growth
1	GO:0048640	BP	negative regulation of developmental growth
1	GO:0030308	BP	negative regulation of cell growth
1	GO:0030517	BP	negative regulation of axon extension
1	GO:0048843	BP	negative regulation of axon extension involved in axon guidance
1	GO:0048518	BP	positive regulation of biological process
1	GO:0048522	BP	positive regulation of cellular process
1	GO:0010647	BP	positive regulation of cell communication
1	GO:0051272	BP	positive regulation of cellular component movement
1	GO:0009893	BP	positive regulation of metabolic process
1	GO:0010604	BP	positive regulation of macromolecule metabolic process
1	GO:0010628	BP	positive regulation of gene expression
1	GO:0048584	BP	positive regulation of response to stimulus
1	GO:0023056	BP	positive regulation of signaling
1	GO:0009967	BP	positive regulation of signal transduction
1	GO:0030177	BP	positive regulation of Wnt signaling pathway
1	GO:2000052	BP	positive regulation of non-canonical Wnt signaling pathway
0.405	GO:2000096	BP	positive regulation of Wnt signaling pathway, planar cell polarity pathway
1	GO:0040017	BP	positive regulation of locomotion
1	GO:2000147	BP	positive regulation of cell motility
1	GO:0030335	BP	positive regulation of cell migration
1	GO:0051091	BP	positive regulation of sequence-specific DNA binding transcription factor activity
1	GO:0005575	CC	cellular_component
1	GO:0005576	CC	extracellular region
1	GO:0043226	CC	organelle

P-value	Term ID	Term type	Term name
1	GO:0043228	CC	non-membrane-bounded organelle
1	GO:0043227	CC	membrane-bounded organelle
1	GO:0031982	CC	vesicle
1	GO:0045202	CC	synapse
1	GO:0005856	CC	cytoskeleton
1	GO:0045111	CC	intermediate filament cytoskeleton
1	GO:0099080	CC	supramolecular complex
1	GO:0099081	CC	supramolecular polymer
1	GO:0099512	CC	supramolecular fiber
1	GO:0005623	CC	cell
1	GO:0016020	CC	membrane
1	GO:0098805	CC	whole membrane
1	GO:0019867	CC	outer membrane
1	GO:0032991	CC	macromolecular complex
1	GO:1902494	CC	catalytic complex
1	GO:1990234	CC	transferase complex
1	GO:0061695	CC	transferase complex, transferring phosphorus-containing groups
1	GO:1990391	CC	DNA repair complex
1	GO:1990351	CC	transporter complex
1	GO:0043234	CC	protein complex
1	GO:0044464	CC	cell part
1	GO:0097458	CC	neuron part
1	GO:0031975	CC	envelope
0.405	GO:0030313	CC	cell envelope
1	GO:0005622	CC	intracellular
1	GO:0044424	CC	intracellular part
1	GO:0005737	CC	cytoplasm
1	GO:0044444	CC	cytoplasmic part
1	GO:0005667	CC	transcription factor complex
1	GO:0043229	CC	intracellular organelle
1	GO:0043232	CC	intracellular non-membrane-bounded organelle
1	GO:0043231	CC	intracellular membrane-bounded organelle
1	GO:0042579	CC	microbody
1	GO:0005777	CC	peroxisome
1	GO:0005634	CC	nucleus
1	GO:0005739	CC	mitochondrion
1	GO:0005793	CC	endoplasmic reticulum-Golgi intermediate compartment
1	GO:0005694	CC	chromosome
1	GO:0000793	CC	condensed chromosome
1	GO:0012505	CC	endomembrane system

P-value	Term ID	Term type	Term name
1	GO:0005783	CC	endoplasmic reticulum
1	GO:0042995	CC	cell projection
1	GO:0043005	CC	neuron projection
1	GO:1990204	CC	oxidoreductase complex
1	GO:0009331	CC	glycerol-3-phosphate dehydrogenase complex
1	GO:0071944	CC	cell periphery
0.405	GO:0030312	CC	external encapsulating structure
0.405	GO:0044462	CC	external encapsulating structure part
0.405	GO:0009279	CC	cell outer membrane
1	GO:0005886	CC	plasma membrane
1	GO:0044422	CC	organelle part
1	GO:0044446	CC	intracellular organelle part
1	GO:0044427	CC	chromosomal part
1	GO:0030915	CC	Smc5-Smc6 complex
1	GO:0044428	CC	nuclear part
1	GO:0000109	CC	nucleotide-excision repair complex
0.405	GO:0071942	CC	XPC complex
0.405	GO:0000111	CC	nucleotide-excision repair factor 2 complex
1	GO:0044430	CC	cytoskeletal part
1	GO:0099513	CC	polymeric cytoskeletal fiber
1	GO:0005882	CC	intermediate filament
1	GO:0044438	CC	microbody part
1	GO:0044439	CC	peroxisomal part
1	GO:0044429	CC	mitochondrial part
1	GO:0031967	CC	organelle envelope
1	GO:0005635	CC	nuclear envelope
1	GO:0005643	CC	nuclear pore
1	GO:0005740	CC	mitochondrial envelope
1	GO:0031090	CC	organelle membrane
1	GO:0098588	CC	bounding membrane of organelle
1	GO:0031903	CC	microbody membrane
1	GO:0005778	CC	peroxisomal membrane
1	GO:0019866	CC	organelle inner membrane
1	GO:0031966	CC	mitochondrial membrane
1	GO:0005743	CC	mitochondrial inner membrane
1	GO:0044456	CC	synapse part
1	GO:0098793	CC	presynapse
1	GO:0044421	CC	extracellular region part
1	GO:0005615	CC	extracellular space
1	GO:0044425	CC	membrane part
1	GO:0044459	CC	plasma membrane part
1	GO:0098590	CC	plasma membrane region

P-value	Term ID	Term type	Term name
1	GO:0097060	CC	synaptic membrane
1	GO:0042734	CC	presynaptic membrane
1	GO:0031224	CC	intrinsic component of membrane
1	GO:0031300	CC	intrinsic component of organelle membrane
1	GO:0031231	CC	intrinsic component of peroxisomal membrane
0.405	GO:0031230	CC	intrinsic component of cell outer membrane
1	GO:0016021	CC	integral component of membrane
1	GO:0031301	CC	integral component of organelle membrane
1	GO:0005779	CC	integral component of peroxisomal membrane
0.405	GO:0045203	CC	integral component of cell outer membrane
1	GO:0031226	CC	intrinsic component of plasma membrane
1	GO:0005887	CC	integral component of plasma membrane
1	GO:0019898	CC	extrinsic component of membrane
1	GO:1902495	CC	transmembrane transporter complex
1	GO:0034702	CC	ion channel complex
1	GO:0034703	CC	cation channel complex
1	GO:0034705	CC	potassium channel complex
1	GO:0008076	CC	voltage-gated potassium channel complex
1	GO:0005942	CC	phosphatidylinositol 3-kinase complex
1	GO:0003674	MF	molecular_function
1	GO:0004871	MF	signal transducer activity
1	GO:0005215	MF	transporter activity
1	GO:0022892	MF	substrate-specific transporter activity
1	GO:0022857	MF	transmembrane transporter activity
1	GO:0022803	MF	passive transmembrane transporter activity
1	GO:0015267	MF	channel activity
1	GO:0022836	MF	gated channel activity
1	GO:0022832	MF	voltage-gated channel activity
1	GO:0022891	MF	substrate-specific transmembrane transporter activity
1	GO:0005342	MF	organic acid transmembrane transporter activity
1	GO:0022838	MF	substrate-specific channel activity
1	GO:0015075	MF	ion transmembrane transporter activity
1	GO:0008509	MF	anion transmembrane transporter activity
1	GO:0008514	MF	organic anion transmembrane transporter activity
1	GO:0046943	MF	carboxylic acid transmembrane transporter activity
1	GO:0015171	MF	amino acid transmembrane transporter activity
1	GO:0008324	MF	cation transmembrane transporter activity
1	GO:0022890	MF	inorganic cation transmembrane transporter activity
1	GO:0015077	MF	monovalent inorganic cation transmembrane transporter activity

P-value	Term ID	Term type	Term name
1	GO:0015078	MF	hydrogen ion transmembrane transporter activity
1	GO:0072509	MF	divalent inorganic cation transmembrane transporter activity
1	GO:0046873	MF	metal ion transmembrane transporter activity
1	GO:0015085	MF	calcium ion transmembrane transporter activity
1	GO:0015079	MF	potassium ion transmembrane transporter activity
1	GO:0005216	MF	ion channel activity
1	GO:0005261	MF	cation channel activity
1	GO:0005262	MF	calcium channel activity
1	GO:0005267	MF	potassium channel activity
1	GO:0005244	MF	voltage-gated ion channel activity
1	GO:0022843	MF	voltage-gated cation channel activity
1	GO:0005249	MF	voltage-gated potassium channel activity
0.41	GO:0005251	MF	delayed rectifier potassium channel activity
1	GO:0045499	MF	chemorepellent activity
1	GO:0098772	MF	molecular function regulator
1	GO:0005085	MF	guanyl-nucleotide exchange factor activity
1	GO:0005088	MF	Ras guanyl-nucleotide exchange factor activity
1	GO:0005089	MF	Rho guanyl-nucleotide exchange factor activity
1	GO:0030234	MF	enzyme regulator activity
1	GO:0060589	MF	nucleoside-triphosphatase regulator activity
1	GO:0030695	MF	GTPase regulator activity
1	GO:0008047	MF	enzyme activator activity
1	GO:0005096	MF	GTPase activator activity
1	GO:0000988	MF	transcription factor activity, protein binding
1	GO:0000989	MF	transcription factor activity, transcription factor binding
1	GO:0003712	MF	transcription cofactor activity
1	GO:0009055	MF	electron carrier activity
1	GO:0060089	MF	molecular transducer activity
1	GO:0004872	MF	receptor activity
1	GO:0099600	MF	transmembrane receptor activity
1	GO:0022834	MF	ligand-gated channel activity
1	GO:0015276	MF	ligand-gated ion channel activity
1	GO:0005217	MF	intracellular ligand-gated ion channel activity
1	GO:0099094	MF	ligand-gated cation channel activity
1	GO:0038023	MF	signaling receptor activity
1	GO:0003707	MF	steroid hormone receptor activity
1	GO:0001653	MF	peptide receptor activity
1	GO:0004888	MF	transmembrane signaling receptor activity
1	GO:0099604	MF	ligand-gated calcium channel activity

P-value	Term ID	Term type	Term name
1	GO:0015278	MF	calcium-release channel activity
1	GO:0005220	MF	inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity
1	GO:0004930	MF	G-protein coupled receptor activity
1	GO:0008528	MF	G-protein coupled peptide receptor activity
1	GO:0004985	MF	opioid receptor activity
0.405	GO:0001626	MF	nociceptin receptor activity
1	GO:0008066	MF	glutamate receptor activity
1	GO:0098988	MF	G-protein coupled glutamate receptor activity
1	GO:0001640	MF	adenylate cyclase inhibiting G-protein coupled glutamate receptor activity
1	GO:0001642	MF	group III metabotropic glutamate receptor activity
1	GO:0017154	MF	semaphorin receptor activity
1	GO:0003824	MF	catalytic activity
0.405	GO:0070283	MF	radical SAM enzyme activity
1	GO:0016491	MF	oxidoreductase activity
1	GO:0016675	MF	oxidoreductase activity, acting on a heme group of donors
1	GO:0016676	MF	oxidoreductase activity, acting on a heme group of donors, oxygen as acceptor
1	GO:0016651	MF	oxidoreductase activity, acting on NAD(P)H
1	GO:0016653	MF	oxidoreductase activity, acting on NAD(P)H, heme protein as acceptor
1	GO:0016627	MF	oxidoreductase activity, acting on the CH-CH group of donors
1	GO:0015002	MF	heme-copper terminal oxidase activity
1	GO:0004129	MF	cytochrome-c oxidase activity
1	GO:0016614	MF	oxidoreductase activity, acting on CH-OH group of donors
1	GO:0016616	MF	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor
1	GO:0004367	MF	glycerol-3-phosphate dehydrogenase [NAD+] activity
1	GO:0016705	MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen
1	GO:0004497	MF	monooxygenase activity
1	GO:0016787	MF	hydrolase activity
1	GO:0017171	MF	serine hydrolase activity
1	GO:0008233	MF	peptidase activity
1	GO:0070011	MF	peptidase activity, acting on L-amino acid peptides
1	GO:0004175	MF	endopeptidase activity
1	GO:0008238	MF	exopeptidase activity
1	GO:0004177	MF	aminopeptidase activity
1	GO:0008239	MF	dipeptidyl-peptidase activity

P-value	Term ID	Term type	Term name
1	GO:0004180	MF	carboxypeptidase activity
1	GO:0008236	MF	serine-type peptidase activity
1	GO:0070008	MF	serine-type exopeptidase activity
1	GO:0004185	MF	serine-type carboxypeptidase activity
1	GO:0004252	MF	serine-type endopeptidase activity
1	GO:0016817	MF	hydrolase activity, acting on acid anhydrides
1	GO:0016818	MF	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides
1	GO:0016462	MF	pyrophosphatase activity
1	GO:0017111	MF	nucleoside-triphosphatase activity
1	GO:0004386	MF	helicase activity
1	GO:0070035	MF	purine NTP-dependent helicase activity
1	GO:0003678	MF	DNA helicase activity
1	GO:0043138	MF	3'-5' DNA helicase activity
1	GO:0009378	MF	four-way junction helicase activity
1	GO:0016887	MF	ATPase activity
1	GO:0042623	MF	ATPase activity, coupled
1	GO:0008094	MF	DNA-dependent ATPase activity
1	GO:0008026	MF	ATP-dependent helicase activity
1	GO:0004003	MF	ATP-dependent DNA helicase activity
1	GO:0043140	MF	ATP-dependent 3'-5' DNA helicase activity
1	GO:0016788	MF	hydrolase activity, acting on ester bonds
1	GO:0042578	MF	phosphoric ester hydrolase activity
1	GO:0016791	MF	phosphatase activity
1	GO:0052866	MF	phosphatidylinositol phosphate phosphatase activity
1	GO:0034593	MF	phosphatidylinositol bisphosphate phosphatase activity
1	GO:0034596	MF	phosphatidylinositol phosphate 4-phosphatase activity
1	GO:0016316	MF	phosphatidylinositol-3,4-bisphosphate 4-phosphatase activity
1	GO:0016874	MF	ligase activity
0.932	GO:0016879	MF	ligase activity, forming carbon-nitrogen bonds
1	GO:0016881	MF	acid-amino acid ligase activity
0.405	GO:0004363	MF	glutathione synthase activity
0.808	GO:0016979	MF	lipoate-protein ligase activity
1	GO:0016740	MF	transferase activity
1	GO:0016782	MF	transferase activity, transferring sulfur-containing groups
1	GO:0016783	MF	sulfurtransferase activity
0.405	GO:0016992	MF	lipoate synthase activity
1	GO:0016772	MF	transferase activity, transferring phosphorus-containing groups

P-value	Term ID	Term type	Term name
1	GO:0016301	MF	kinase activity
1	GO:0035004	MF	phosphatidylinositol 3-kinase activity
1	GO:0016773	MF	phosphotransferase activity, alcohol group as acceptor
1	GO:0052813	MF	phosphatidylinositol bisphosphate kinase activity
1	GO:0046934	MF	phosphatidylinositol-4,5-bisphosphate 3-kinase activity
1	GO:0052742	MF	phosphatidylinositol kinase activity
1	GO:0016303	MF	1-phosphatidylinositol-3-kinase activity
1	GO:0016307	MF	phosphatidylinositol phosphate kinase activity
1	GO:0035005	MF	1-phosphatidylinositol-4-phosphate 3-kinase activity
1	GO:0004672	MF	protein kinase activity
1	GO:0004674	MF	protein serine/threonine kinase activity
1	GO:0005198	MF	structural molecule activity
1	GO:0005488	MF	binding
1	GO:1901681	MF	sulfur compound binding
1	GO:1901363	MF	heterocyclic compound binding
1	GO:0097367	MF	carbohydrate derivative binding
1	GO:0097159	MF	organic cyclic compound binding
1	GO:0046906	MF	tetrapyrrole binding
1	GO:0020037	MF	heme binding
1	GO:1901265	MF	nucleoside phosphate binding
1	GO:0003676	MF	nucleic acid binding
1	GO:0001067	MF	regulatory region nucleic acid binding
1	GO:0003677	MF	DNA binding
1	GO:0003684	MF	damaged DNA binding
1	GO:0003697	MF	single-stranded DNA binding
1	GO:0043565	MF	sequence-specific DNA binding
1	GO:0003690	MF	double-stranded DNA binding
1	GO:1990837	MF	sequence-specific double-stranded DNA binding
1	GO:0000975	MF	regulatory region DNA binding
1	GO:0044212	MF	transcription regulatory region DNA binding
1	GO:0001012	MF	RNA polymerase II regulatory region DNA binding
1	GO:0000976	MF	transcription regulatory region sequence-specific DNA binding
1	GO:0000977	MF	RNA polymerase II regulatory region sequence-specific DNA binding
1	GO:0043167	MF	ion binding
1	GO:0043168	MF	anion binding
1	GO:0035639	MF	purine ribonucleoside triphosphate binding
1	GO:0043169	MF	cation binding
1	GO:0046872	MF	metal ion binding

P-value	Term ID	Term type	Term name
1	GO:0005509	MF	calcium ion binding
1	GO:0000287	MF	magnesium ion binding
1	GO:0046914	MF	transition metal ion binding
1	GO:0005506	MF	iron ion binding
1	GO:0008270	MF	zinc ion binding
1	GO:0033218	MF	amide binding
1	GO:0042277	MF	peptide binding
0.808	GO:1900750	MF	oligopeptide binding
1	GO:0042923	MF	neuropeptide binding
1	GO:0072341	MF	modified amino acid binding
0.808	GO:0043295	MF	glutathione binding
1	GO:0051540	MF	metal cluster binding
1	GO:0051536	MF	iron-sulfur cluster binding
1	GO:0051539	MF	4 iron, 4 sulfur cluster binding
1	GO:0050840	MF	extracellular matrix binding
1	GO:0005515	MF	protein binding
1	GO:0046983	MF	protein dimerization activity
1	GO:0042393	MF	histone binding
1	GO:0005102	MF	receptor binding
1	GO:0038191	MF	neuropilin binding
1	GO:0030215	MF	semaphorin receptor binding
1	GO:0042802	MF	identical protein binding
1	GO:0042803	MF	protein homodimerization activity
1	GO:0043236	MF	laminin binding
1	GO:0048037	MF	cofactor binding
1	GO:0050662	MF	coenzyme binding
1	GO:0044877	MF	macromolecular complex binding
1	GO:0003682	MF	chromatin binding
1	GO:0036094	MF	small molecule binding
1	GO:0001882	MF	nucleoside binding
1	GO:0001883	MF	purine nucleoside binding
1	GO:0032549	MF	ribonucleoside binding
1	GO:0032550	MF	purine ribonucleoside binding
1	GO:0000166	MF	nucleotide binding
1	GO:0032553	MF	ribonucleotide binding
1	GO:0017076	MF	purine nucleotide binding
1	GO:0030554	MF	adenyl nucleotide binding
1	GO:0032555	MF	purine ribonucleotide binding
1	GO:0032559	MF	adenyl ribonucleotide binding
1	GO:0005524	MF	ATP binding
1	GO:0051287	MF	NAD binding
1	GO:0001071	MF	nucleic acid binding transcription factor activity

P-value	Term ID	Term type	Term name
1	GO:0003700	MF	transcription factor activity, sequence-specific DNA binding
1	GO:0098531	MF	transcription factor activity, direct ligand regulated sequence-specific DNA binding
1	GO:0000981	MF	RNA polymerase II transcription factor activity, sequence-specific DNA binding
1	GO:0004879	MF	RNA polymerase II transcription factor activity, ligand-activated sequence-specific DNA binding

Table S4. Mean results of three replicate STRUCTURE analyses for K=2.

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
N009964	0.9880	0.0120
N009976	0.8982	0.1018
N009965a	0.9961	0.0039
N009977	0.9988	0.0012
N009966	0.9082	0.0918
N009978	0.8913	0.1087
N009965b	0.8489	0.1511
N009967	0.9013	0.0987
N009979	0.9954	0.0046
N009968	0.9991	0.0009
N009969	0.8986	0.1014
N009970	0.9973	0.0027
N009971	0.9987	0.0013
N009960	0.9970	0.0030
N009972	0.8816	0.1184
N009961	0.9243	0.0757
N009973	0.9987	0.0013
N009962	0.9976	0.0024
N009974	0.9989	0.0011
N009963	0.9977	0.0023
N009975	0.9990	0.0010
T009876	0.9972	0.0028
T009877	0.9985	0.0015
T009878	0.9980	0.0020
T009879	0.9071	0.0929
T009868	0.9971	0.0029
T009880	0.9168	0.0832
T009869	0.9834	0.0166
T009881	0.9738	0.0262
T009870	0.9912	0.0088
T009882	0.9992	0.0008
T009871	0.9969	0.0031
T009883	0.9818	0.0182
T009872	0.9956	0.0044
T009884	0.9772	0.0228
T009873	0.9959	0.0041
T009885	0.9975	0.0025
T009874	0.9967	0.0033
T009886	0.9864	0.0136
T009875	0.9985	0.0015

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
T009887	0.9994	0.0006
2JA001	0.9963	0.0037
2JA002	0.9979	0.0021
2JA003	0.9757	0.0243
2JA004	0.9989	0.0011
2JA005	0.9739	0.0261
2JA006	0.9830	0.0170
2JA007	0.9988	0.0012
2JA009	0.8921	0.1079
2JA010	0.9575	0.0425
2JA011	0.9726	0.0274
2JA013	0.8892	0.1108
2JA014	0.9978	0.0022
2JA015	0.9981	0.0019
2JA016	0.9081	0.0919
2JA017	0.9960	0.0040
2JA018	0.9166	0.0834
2JA020	0.9074	0.0926
2JB001	0.9990	0.0010
2JB002	0.9696	0.0304
2JB003	0.9953	0.0047
2JB004	0.9988	0.0012
2JB005	0.9132	0.0868
2JB006	0.9948	0.0052
2JB007	0.9992	0.0008
2JB008	0.9978	0.0022
2JB009	0.9985	0.0015
2JB010	0.9372	0.0628
2JB011	0.9292	0.0708
2JB012	0.9988	0.0012
2JB013	0.9957	0.0043
2JB014	0.9985	0.0015
2JB015	0.9986	0.0014
2JB016	0.9690	0.0310
2JB017	0.9881	0.0119
2JB018	0.9852	0.0148
2JB019	0.9511	0.0489
2JB020	0.8787	0.1213
SMP11200	0.0023	0.9977
SMP11212	0.0005	0.9995
SMP11201	0.0005	0.9995

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
SMP11213	0.0005	0.9995
SMP11202	0.0012	0.9988
SMP11214	0.0004	0.9996
SMP11203	0.0548	0.9452
SMP11216	0.0022	0.9978
SMP11204	0.0010	0.9990
SMP11205	0.0011	0.9989
SMP11206	0.0468	0.9532
SMP11207	0.0005	0.9995
SMP11196	0.0020	0.9980
SMP11215	0.0028	0.9972
SMP11197	0.0006	0.9994
SMP11208	0.0154	0.9846
SMP11198	0.0432	0.9568
SMP11210	0.0005	0.9995
SMP11209	0.0036	0.9964
SMP11211	0.0020	0.9980
GBJ11419	0.0506	0.9494
GBJ11420	0.0006	0.9994
GBJ11421	0.0035	0.9965
GBJ11422	0.0007	0.9993
GBJ11411	0.0005	0.9995
GBJ11423	0.0005	0.9995
GBJ11412	0.0037	0.9963
GBJ11424	0.1082	0.8918
GBJ11413	0.0059	0.9941
GBJ11426	0.0006	0.9994
GBJ11414	0.0011	0.9989
GBJ11427	0.0512	0.9488
GBJ11428	0.0437	0.9563
GBJ11416	0.0005	0.9995
GBJ11429	0.0011	0.9989
GBJ11417	0.0004	0.9996
GBJ11430	0.0012	0.9988
GBJ11418	0.0004	0.9996
GBJ11431	0.0006	0.9994
GBM001	0.0009	0.9991
GBM003	0.0168	0.9832
GBM004	0.0010	0.9990
GBM005	0.0008	0.9992
GBM006	0.0013	0.9987

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
GBM007	0.0005	0.9995
GBM008	0.0006	0.9994
GBM009	0.0227	0.9773
GBM010	0.0005	0.9995
GBM012	0.0526	0.9474
GBM013	0.0009	0.9991
GBM014	0.0006	0.9994
GBM015	0.0441	0.9559
GBM016	0.0007	0.9993
GBM017	0.0007	0.9993
GBM018	0.0004	0.9996
GBM019	0.0019	0.9981
GBM020	0.0096	0.9904
GBM021	0.0018	0.9982
GBM022	0.0006	0.9994
GBM023	0.0024	0.9976
GBM024	0.0029	0.9971
GBM025	0.0007	0.9993
GBM026	0.0014	0.9986
CCF0910119	0.9913	0.0087
CCF0910120	0.9988	0.0012
CCF0910121	0.9976	0.0024
CCF0910122	0.9770	0.0230
CCF0910123	0.9992	0.0008
CCF0910127	0.9944	0.0056
CCF0910124	0.9184	0.0816
CCF0910128	0.9988	0.0012
CCF0910125	0.9980	0.0020
CCF0910129	0.9951	0.0049
CCF0910126	0.9994	0.0006
CCF0910130	0.9155	0.0845
CCF0911	0.0009	0.9991
CCF0912	0.9900	0.0100
CCF0913	0.9809	0.0191
CCF0903	0.9763	0.0237
CCF0915	0.0013	0.9987
CCF0904	0.9865	0.0135
CCF0916	0.9945	0.0055
CCF0905	0.9996	0.0004
CCF0917	0.9363	0.0637
CCF0906	0.9984	0.0016

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
CCF0918	0.9889	0.0111
CCF0907	0.9986	0.0014
CCF0919	0.9974	0.0026
CCF0908	0.8937	0.1063
CCF0920	0.9994	0.0006
CCF0909	0.9782	0.0218
CCF0921	0.9122	0.0878
CCF0910	0.9812	0.0188
CCF0922	0.9888	0.0112
CCF1110160	0.8819	0.1181
CCF1110161	0.0006	0.9994
CCF1110162	0.9985	0.0015
CCF1110163	0.9251	0.0749
CCF1110156	0.8840	0.1160
CCF1110157	0.9940	0.0060
CCF1110158	0.8497	0.1503
CCF1110159	0.9901	0.0099
SR009252	0.0067	0.9933
SR009253	0.9983	0.0017
SR009254	0.0028	0.9972
SR009255	0.9331	0.0669
SR009256	0.0100	0.9900
SR009257	0.0013	0.9987
SR009258	0.0019	0.9981
SR009259	0.0010	0.9990
SR009248	0.9955	0.0045
SR009260	0.9661	0.0339
SR009249	0.0009	0.9991
SR009261	0.0015	0.9985
SR009250	0.0012	0.9988
SR009262	0.0075	0.9925
SR009251	0.0515	0.9485
SR009263	0.9974	0.0026
SR009264	0.0115	0.9885
SR009265	0.9986	0.0014
SR009266	0.9983	0.0017
SR009267	0.0007	0.9993
GBK11173	0.9614	0.0386
GBK11185	0.9095	0.0905
GBK11174_a	0.9968	0.0032
GBK11186	0.9923	0.0077

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
GBK11175	0.8873	0.1127
GBK11187	0.9283	0.0717
GBK11174_b	0.9735	0.0265
GBK11176	0.9993	0.0007
GBK11188	0.9925	0.0075
GBK11177	0.9921	0.0079
GBK11189	0.9909	0.0091
GBK11178	0.9931	0.0069
GBK11190	0.9727	0.0273
GBK11179	0.9990	0.0010
GBK11191	0.9970	0.0030
GBK11180	0.9443	0.0557
GBK11192	0.9622	0.0378
GBK11181	0.9987	0.0013
GBK11182	0.9985	0.0015
GBK11183	0.9890	0.0110
GBK11184	0.9547	0.0453
GBK11338	0.9250	0.0750
GBK11350	0.8870	0.1130
GBK11339	0.0005	0.9995
GBK11351_a	0.9988	0.0012
GBK11340	0.8636	0.1364
GBK11352	0.8975	0.1025
GBK11351_b	0.0005	0.9995
GBK11353	0.9940	0.0060
GBK11342	0.9665	0.0335
GBK11343	0.9878	0.0122
GBK11344	0.9377	0.0623
GBK11345	0.0003	0.9997
GBK11334	0.9397	0.0603
GBK11346	0.9893	0.0107
GBK11335	0.9238	0.0762
GBK11347	0.9922	0.0078
GBK11336	0.8962	0.1038
GBK11348	0.0124	0.9876
GBK11337	0.9732	0.0268
GBK11349	0.9460	0.0540
SP011256	0.8900	0.1100
SP011257	0.9002	0.0998
SP011258	0.9979	0.0021
SP011259	0.9984	0.0016

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
SP011248	0.9288	0.0712
SP011260	0.9992	0.0008
SP011249	0.9989	0.0011
SP011261	0.9739	0.0261
SP011250	0.9984	0.0016
SP011262	0.9631	0.0369
SP011251	0.9988	0.0012
SP011263	0.9962	0.0038
SP011252	0.9192	0.0808
SP011264	0.9888	0.0112
SP011253	0.9992	0.0008
SP011265	0.9986	0.0014
SP011254	0.9865	0.0135
SP011266	0.9985	0.0015
SP011255	0.9988	0.0012
SP011267	0.8867	0.1133
MFE001	0.8843	0.1157
MFE002	0.9981	0.0019
MFE005	0.9708	0.0292
MFE006	0.9187	0.0813
MFE007	0.9908	0.0092
MFE008	0.9191	0.0809
MFE009	0.9795	0.0205
MFE010	0.8959	0.1041
MFE011	0.8849	0.1151
MFE012	0.9969	0.0031
MFE013	0.0126	0.9874
MFE014	0.9692	0.0308
MFE015	0.0005	0.9995
MFE016	0.9544	0.0456
MFE017	0.9952	0.0048
MFE018	0.9989	0.0011
MFE003	0.9991	0.0009
MFE004	0.0051	0.9949
MFE019	0.9501	0.0499
MFE020	0.9567	0.0433
MFE021	0.9969	0.0031
MFE022	0.9338	0.0662
MFE023	0.9739	0.0261
MFE024	0.9854	0.0146
MFF004	0.8757	0.1243

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
MFF005	0.9679	0.0321
MFF006	0.9923	0.0077
MFF007	0.9982	0.0018
MFF008	0.9971	0.0029
MFF009	0.9846	0.0154
MFF010	0.9844	0.0156
MFF011	0.9661	0.0339
MFF012	0.9398	0.0602
MFF013	0.9271	0.0729
MFF014	0.9990	0.0010
MFF015	0.9855	0.0145
MFF016	0.8869	0.1131
MFF017	0.8456	0.1544
MFF018	0.9008	0.0992
MFF019	0.9969	0.0031
MFF020	0.9822	0.0178
MFF021	0.8954	0.1046
MFF022	0.9987	0.0013
MFF024	0.9589	0.0411
MFF025	0.8719	0.1281
MFF026	0.9912	0.0088
MFF027	0.9283	0.0717
MFF029	0.9362	0.0638
MFG005	0.9538	0.0462
MFG006	0.9991	0.0009
MFG007	0.9978	0.0022
MFG008	0.9937	0.0063
MFG009	0.9286	0.0714
MFG010	0.9806	0.0194
MFG011	0.9748	0.0252
MFG012	0.8815	0.1185
MFG014	0.9992	0.0008
MFG015	0.9992	0.0008
MFG016	0.9460	0.0540
MFG017	0.9966	0.0034
MFG018	0.9937	0.0063
MFG019	0.9979	0.0021
MFG020	0.9970	0.0030
MFG021	0.9122	0.0878
MFG003	0.9831	0.0169
MFG004	0.9437	0.0563

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
MFG022	0.9987	0.0013
MFG023	0.9000	0.1000
MFG025	0.9373	0.0627
MFG026	0.9318	0.0682
MFG027	0.9393	0.0607
MFG028	0.9991	0.0009
MFH002	0.9627	0.0373
MFH003	0.8410	0.1590
MFH004	0.8631	0.1369
MFH005	0.9245	0.0755
MFH006	0.0011	0.9989
MFH007	0.0004	0.9996
MFH008	0.9640	0.0360
MFH009	0.9989	0.0011
MFH010	0.0350	0.9650
MFH011	0.0009	0.9991
MFH013	0.8905	0.1095
MFH014	0.9916	0.0084
MFH016	0.9569	0.0431
MFH017	0.0013	0.9987
MFH018	0.9965	0.0035
MFH019	0.9800	0.0200
MFH020	0.9900	0.0100
MFH021	0.9560	0.0440
MFH022	0.8327	0.1673
MFH023	0.8680	0.1320
MFH024	0.8563	0.1437
MFH025	0.0007	0.9993
MFH026	0.9354	0.0646
MFI001	0.8893	0.1107
MFI002	0.9871	0.0129
MFI003	0.0304	0.9696
MFI004	0.0325	0.9675
MFI006	0.9955	0.0045
MFI007	0.9985	0.0015
MFI009	0.0005	0.9995
MFI010	0.9988	0.0012
MFI011	0.0526	0.9474
MFI012	0.9940	0.0060
MFI013	0.8169	0.1831
MFI014	0.8454	0.1546

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
MFI015	0.9990	0.0010
MFI016	0.9175	0.0825
MFI017	0.9884	0.0116
MFI018	0.9987	0.0013
MFI019	0.8825	0.1175
MFI020	0.9980	0.0020
MFI021	0.0019	0.9981
MFI022	0.9985	0.0015
MFI023	0.9113	0.0887
MFI024	0.0029	0.9971
MFI025	0.8565	0.1435
MFI026	0.9995	0.0005
GBE001	0.0010	0.9990
GBE003	0.0090	0.9910
GBE012	0.0027	0.9973
GBE014	0.0020	0.9980
GBE016	0.0010	0.9990
GBE017	0.0181	0.9819
GBE020	0.0015	0.9985
GBE021	0.0012	0.9988
GBE022	0.0228	0.9772
GBE023	0.0009	0.9991
GBE024	0.0348	0.9652
GBE025	0.0004	0.9996
GBE026	0.0009	0.9991
GBE027	0.0015	0.9985
GBE028	0.0065	0.9935
GBE030	0.0011	0.9989
GBE031	0.0007	0.9993
GBE032	0.0046	0.9954
GBE036	0.0011	0.9989
GBA001	0.0261	0.9739
GBA003	0.0013	0.9987
GBA004	0.0027	0.9973
GBA005	0.0005	0.9995
GBA008	0.0138	0.9862
GBA009	0.0007	0.9993
GBA010	0.0687	0.9313
GBA011	0.0008	0.9992
GBA012	0.0008	0.9992
GBA013	0.0016	0.9984

Individual	Proportion Assigned to Offshore Population	Proportion Assigned to Gilbert Bay Population
GBA014	0.0011	0.9989
GBA015	0.0022	0.9978
GBA016	0.0006	0.9994
GBA017	0.0806	0.9194
GBA018	0.0229	0.9771
GBA019	0.0007	0.9993
GBA020	0.0032	0.9968
GBA021	0.0024	0.9976
GBA022	0.0013	0.9987
GBA023	0.0070	0.9930
GBA024	0.0010	0.9990
GBA025	0.1116	0.8884
GBA026	0.0018	0.9982
GBA027	0.0006	0.9994
GBA028	0.0008	0.9992
GBA029	0.0010	0.9990
GBA030	0.0003	0.9997
GBA031	0.0004	0.9996
GBA032	0.0007	0.9993
GBA033	0.0006	0.9994
GBA034	0.0007	0.9993
GBA035	0.0009	0.9991
GBA036	0.0006	0.9994
GBA037	0.0006	0.9994
GBA038	0.0005	0.9995
GBA039	0.0016	0.9984
GBA040	0.0014	0.9986
GBA041	0.0003	0.9997
GBA042	0.0006	0.9994

Table S5. Baseline assignment accuracy determined by assigner for 25 top-ranked SNPs chosen by each of the three selection methods: (A) FST ranking of all loci (B) FST ranking of SNPs showing no evidence of LD and (C) GRRF ranking.

Number of SNPs	F_{ST} (Filtered for LD)	F_{ST}	GRRF
1	82.5	82.5	76.5
2	81	80.5	82.5
3	93	86.5	87
4	97.5	89.5	89
5	97.5	94.5	92
6	97.5	97	93.5
7	98	97.5	96
8	98.5	98.5	98
9	99	98.5	98
10	98	98.5	99
11	99	96	98
12	99	95	98
13	99	95	98
14	99	94.5	98
15	99	93.5	98
16	99	93.5	98
17	99	91.5	98
18	99	89.5	98
19	99	89.5	98
20	99	89.5	98
21	99	89	99
22	99	87	99
23	100	88.5	99
24	100	87	99
25	100	88	98

Table S6. Posterior mean and standard deviation (S.D.) of estimated mixture proportion of Gilbert Bay cod in each fishery sample calculated using 23 SNP subsets and 7,568 SNPs in gsi_sim. True proportion of Gilbert Bay individuals was determined using individual assignment in STRUCTURE is listed.

Site	True proportion of Gilbert Bay individuals	Mixture proportions							
		7,568 SNPs		23 SNPs F_{ST} (filtered for LD)		23 SNPs F_{ST}		23 SNPs GRRF	
		Posterior mean	S.D.	Posterior mean	S.D.	Posterior mean	S.D.	Posterior mean	S.D.
CCF09	0.065	0.078	0.047	0.078	0.046	0.016	0.021	0.078	0.047
CCF11	0.125	0.166	0.117	0.168	0.119	0.055	0.072	0.166	0.117
GBK	0.098	0.107	0.047	0.106	0.047	0.012	0.017	0.106	0.047
MFE	0.125	0.141	0.068	0.139	0.068	0.020	0.027	0.140	0.068
MFF	0.000	0.020	0.027	0.020	0.027	0.020	0.027	0.020	0.028
MFG	0.000	0.020	0.027	0.020	0.028	0.020	0.027	0.020	0.028
MFH	0.261	0.271	0.089	0.271	0.088	0.021	0.029	0.311	0.095
MFI	0.250	0.261	0.086	0.261	0.086	0.020	0.028	0.305	0.091
SR	0.650	0.644	0.032	0.643	0.033	0.024	0.033	0.643	0.046
SP	0.000	0.024	0.103	0.024	0.101	0.024	0.033	0.038	0.102

Table S7. N_e estimates and corresponding 95% C.I.'s calculated using the LD method, LD method with a bias correcting equation and Jorde and Ryman's temporal method.

Linkage Disequilibrium Method							Temporal method			
Year	Naïve N_e	95% C.I. min.	95% C.I. max.	Bias-corrected N_e	95% C.I. min.	95% C.I. max.	Time period	N_e	95% C.I. min.	95% C.I. max.
1998	698	670.10	728.30	974.00	920.85	1033.57	1998-2004	-75.80	-73.00	∞
2004	476.4	445.00	512.50	664.77	609.56	730.86	1998-2011	357.20	344.20	370.40
2011	699.5	671.20	730.20	976.09	922.21	1036.57	1998-2015	1256.30	1210.50	1303.00
2015	655.2	610.60	706.80	914.27	835.72	1008.97	2004-2011	226.80	218.50	235.20
							2004-2015	-130.60	-125.80	∞
							2011-2015	139.00	133.90	144.10

Appendix B – Supplementary Figures

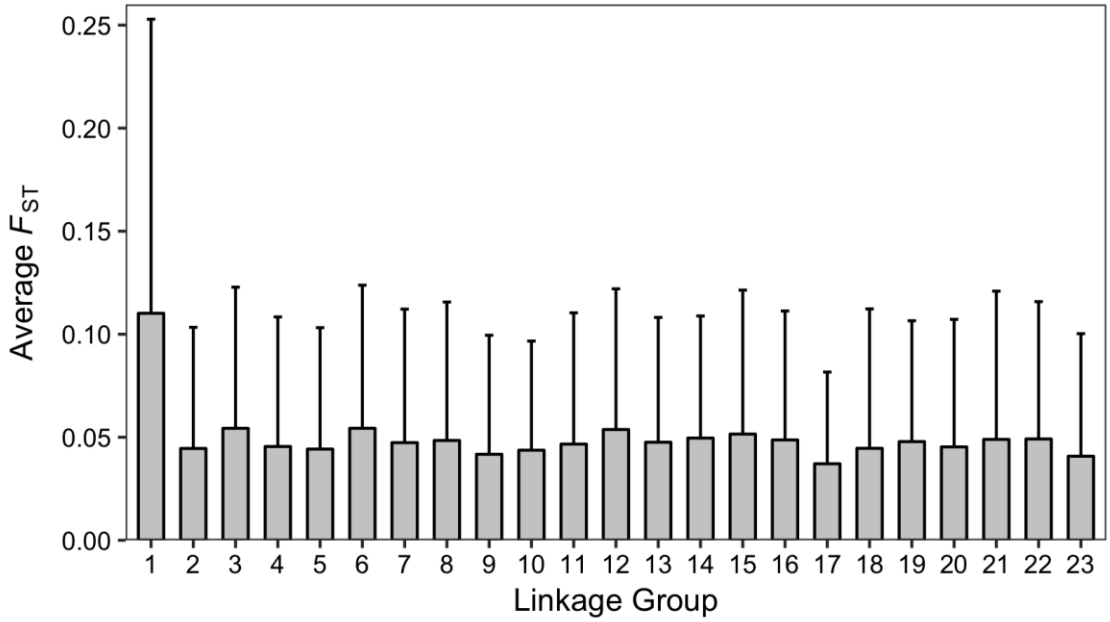


Figure S1**.** Average F_{ST} between putative populations, offshore and Gilbert Bay, for each LG. Error bars represent standard deviation of the mean.

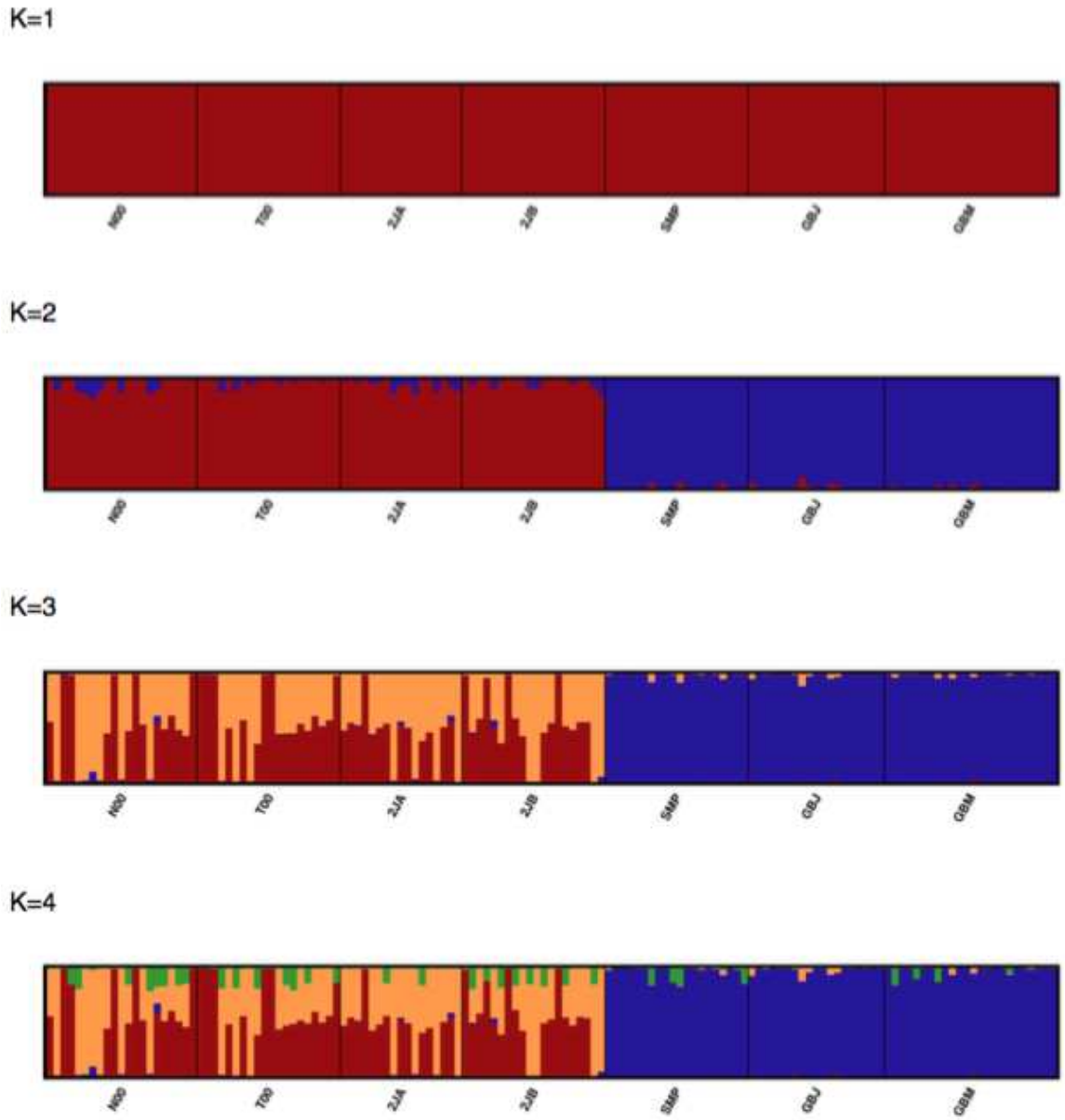


Figure S2. Plots of individual admixture determined by STRUCTURE analysis (K=1-4) for all sites using filtered data from Chapter 2.

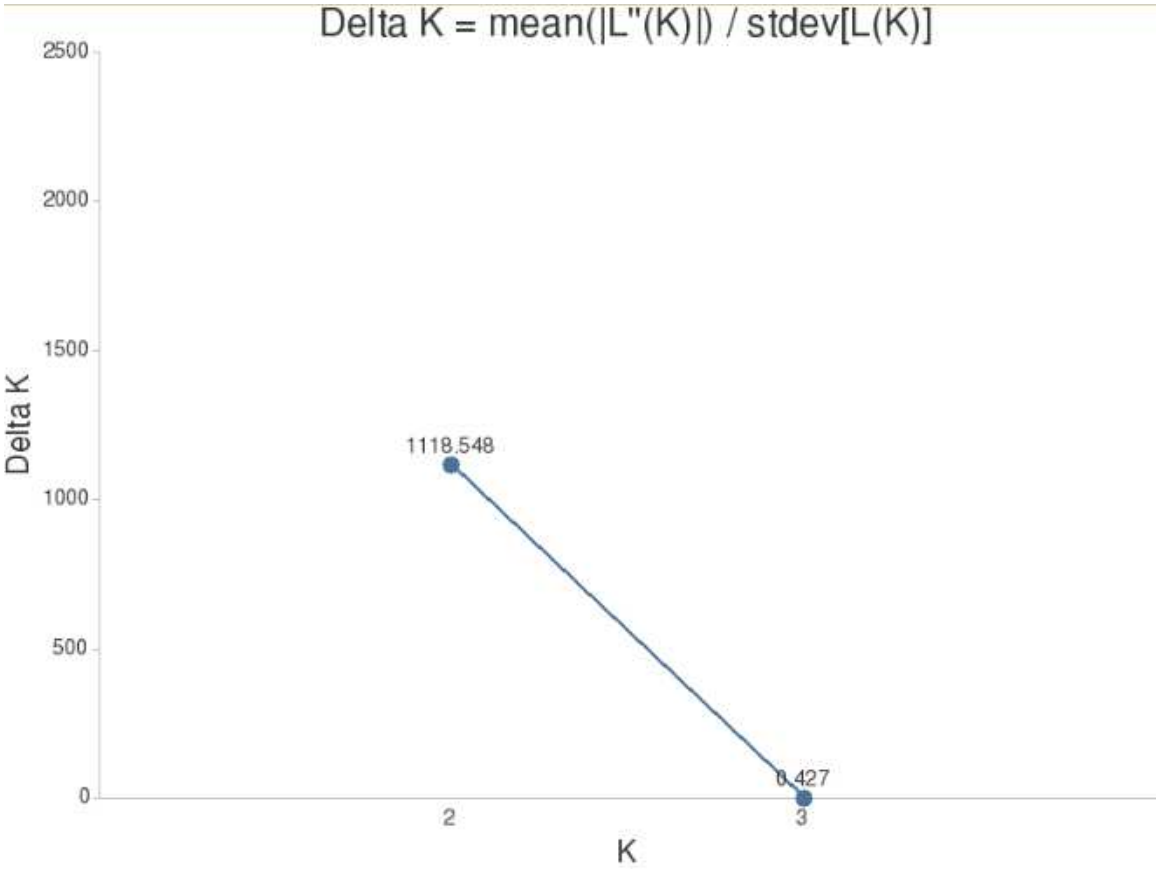
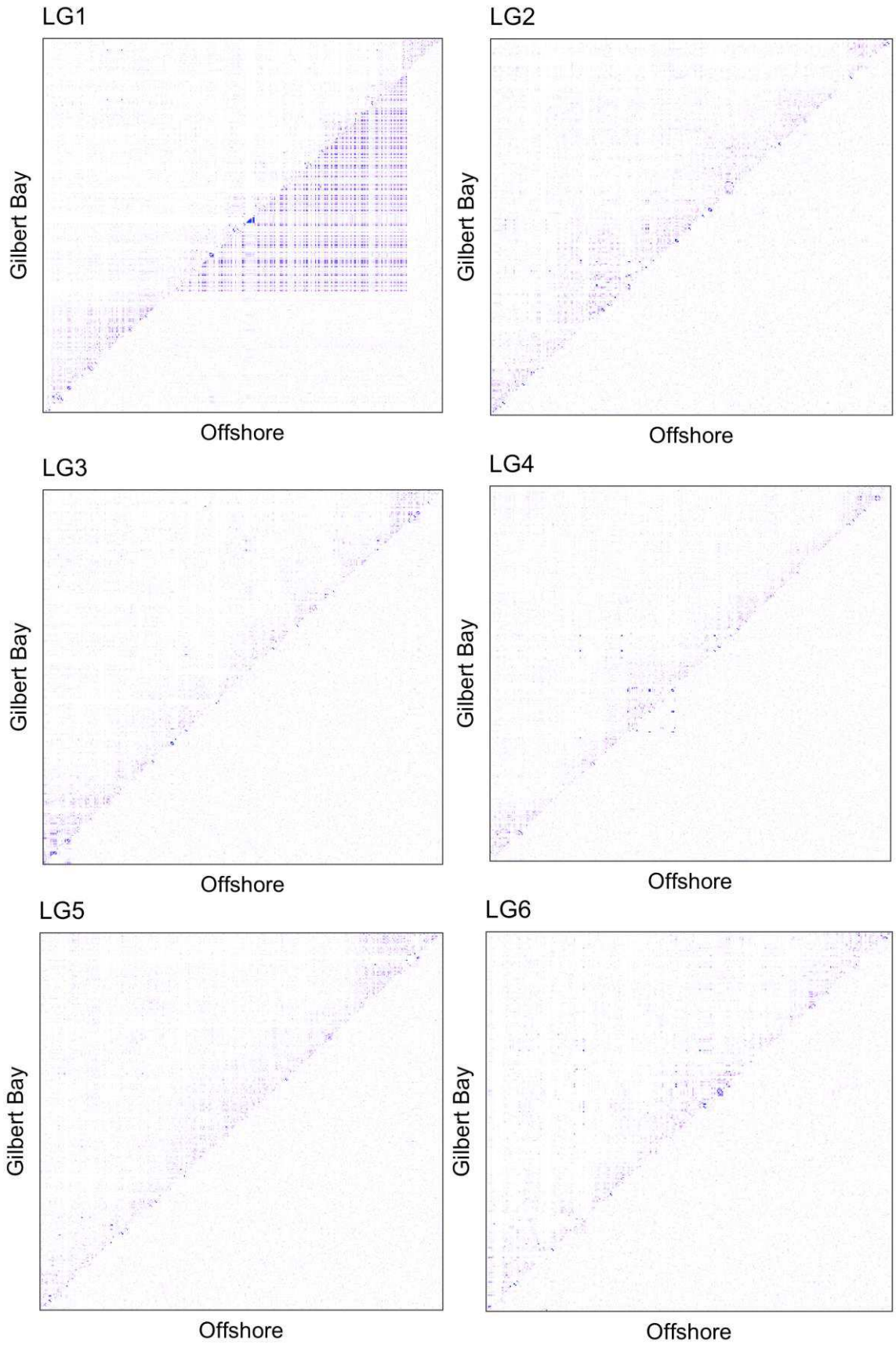
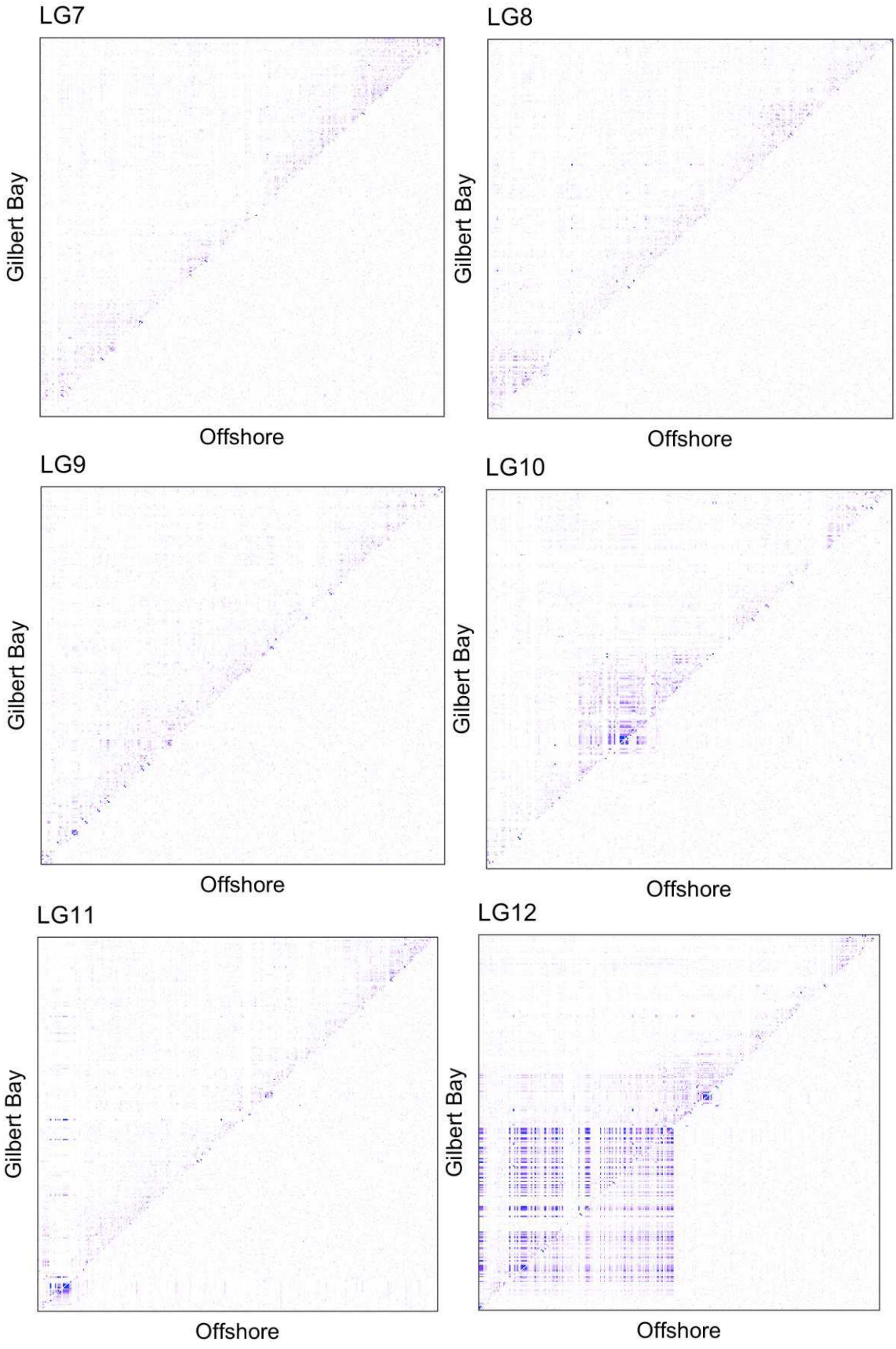
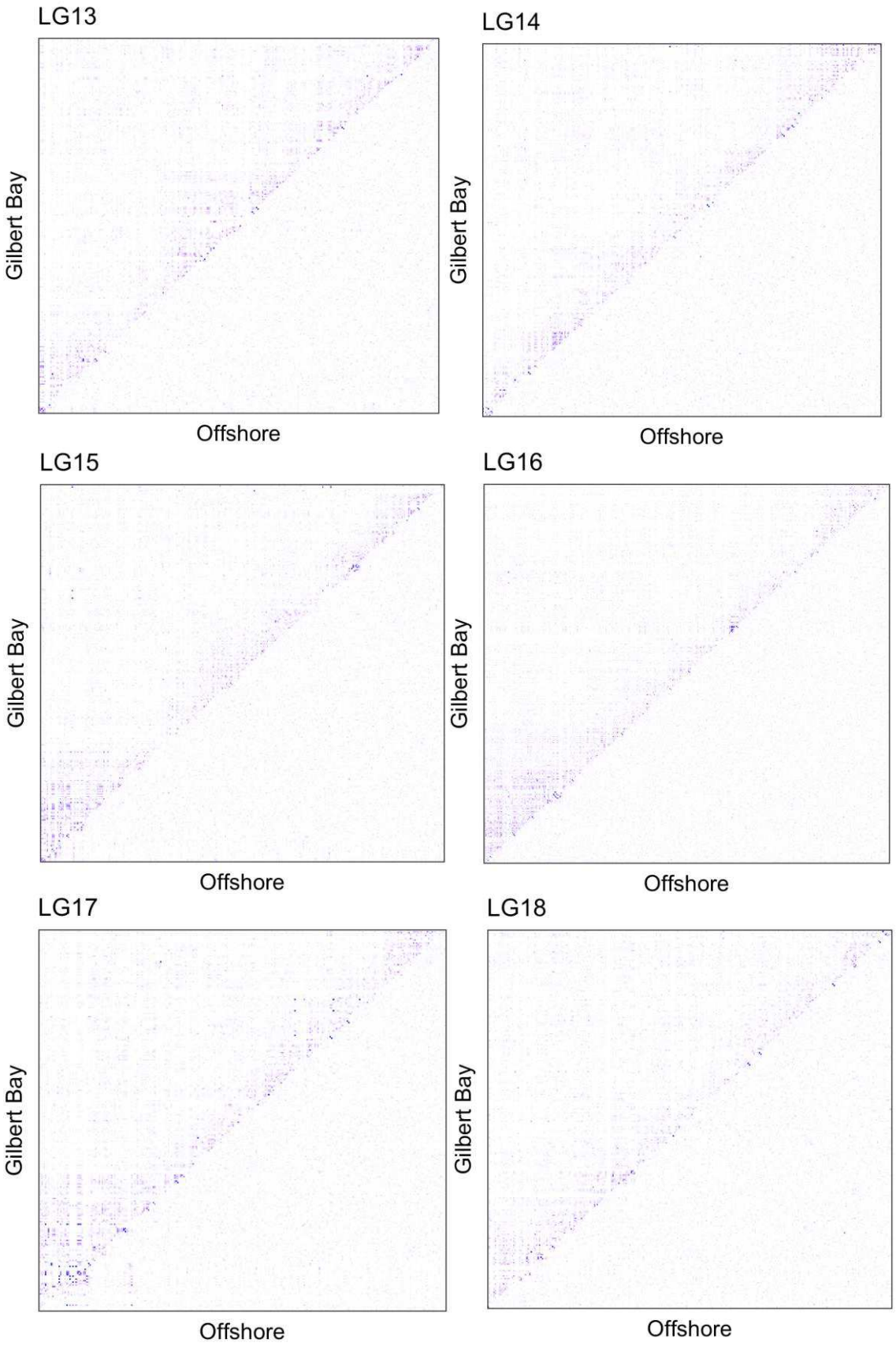
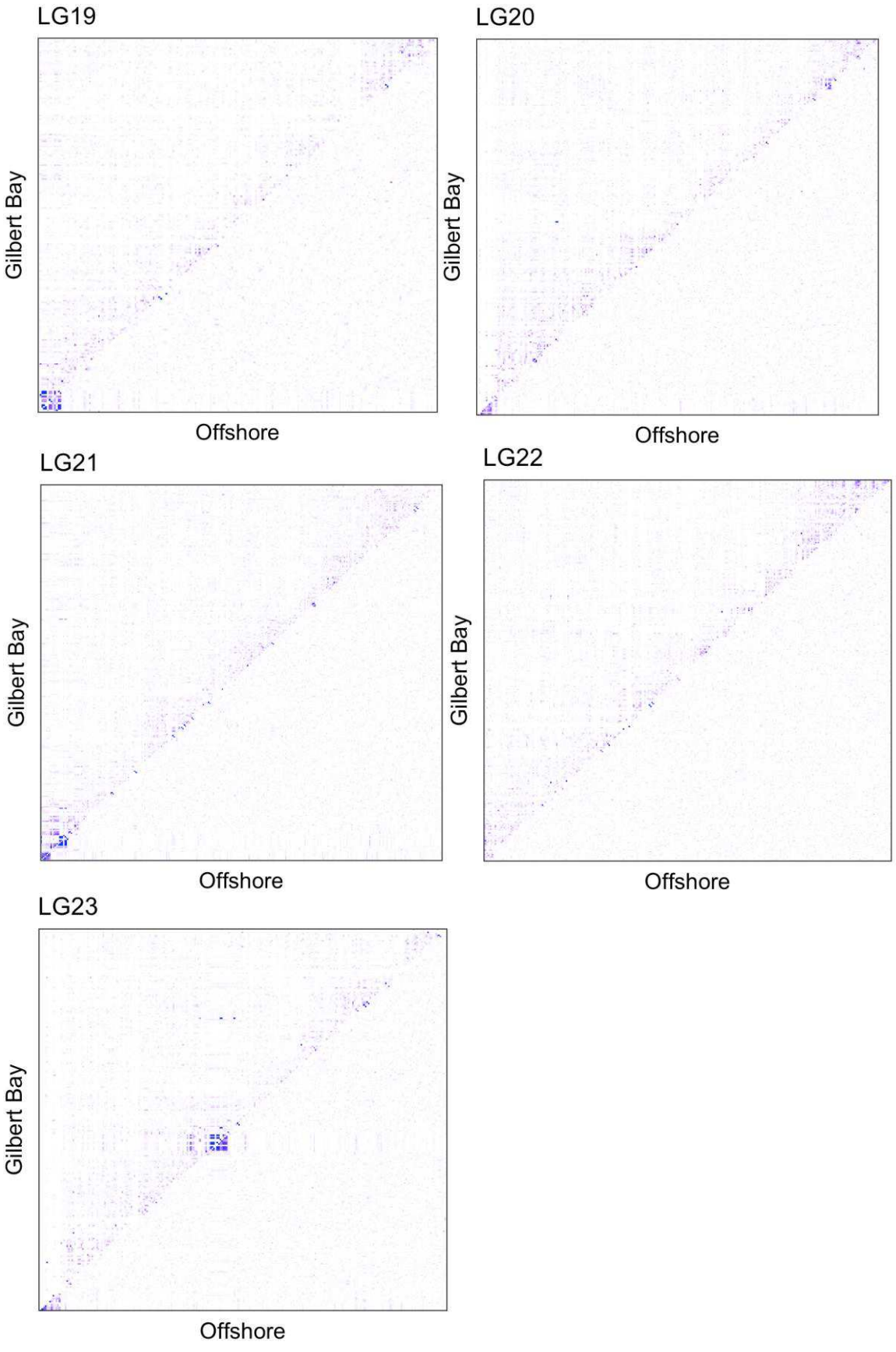


Figure S3. Plot showing the DeltaK statistic for each value of K as determined by the method of Evanno et al. (2005) implemented in CLUMPAK. Filtered data from Chapter 2 was used here.


LG1
Gilbert Bay
Offshore
LG2
Gilbert Bay
Offshore
LG3
Gilbert Bay
Offshore
LG4
Gilbert Bay
Offshore
LG5
Gilbert Bay
Offshore
LG6
Gilbert Bay
Offshore


LG7
Gilbert Bay
Offshore
LG8
Gilbert Bay
Offshore
LG9
Gilbert Bay
Offshore
LG10
Gilbert Bay
Offshore
LG11
Gilbert Bay
Offshore
LG12
Gilbert Bay
Offshore


LG13
Gilbert Bay
Offshore
LG14
Gilbert Bay
Offshore
LG15
Gilbert Bay
Offshore
LG16
Gilbert Bay
Offshore
LG17
Gilbert Bay
Offshore
LG18
Gilbert Bay
Offshore


LG19
Gilbert Bay
Offshore
LG20
Gilbert Bay
Offshore
LG21
Gilbert Bay
Offshore
LG22
Gilbert Bay
Offshore
LG23
Gilbert Bay
Offshore

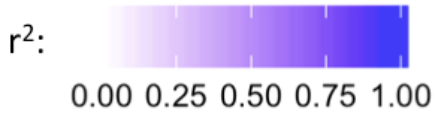


Figure S4. Pattern of pairwise LD, measured as r^2, within each linkage groups (LG1 to LG23) for each population: Gilbert Bay (above diagonal) and offshore (below diagonal).

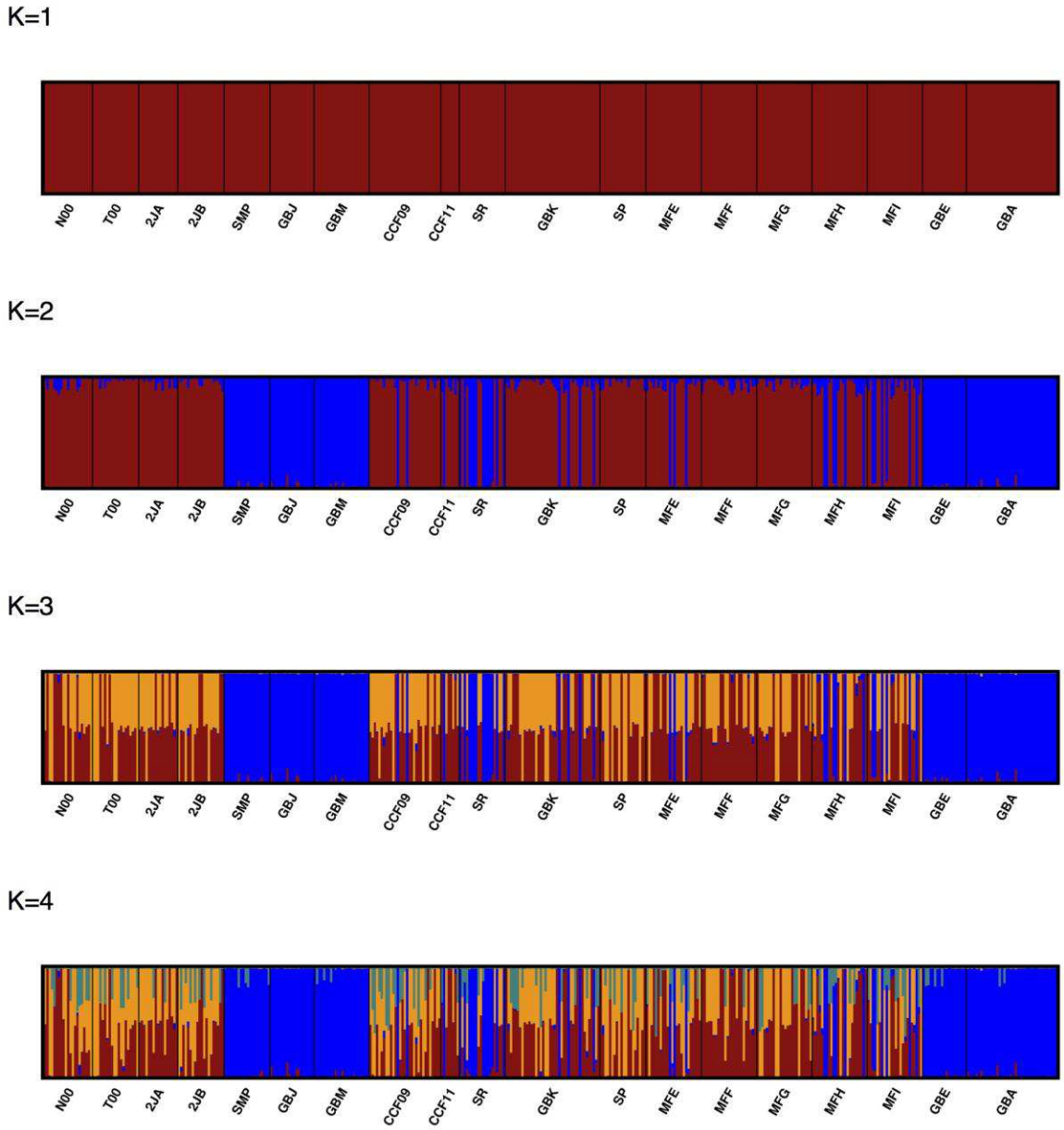


Figure S5. Plots of individual admixture determined by STRUCTURE analysis (K=1-4) for all sites using filtered data from Chapter 3.

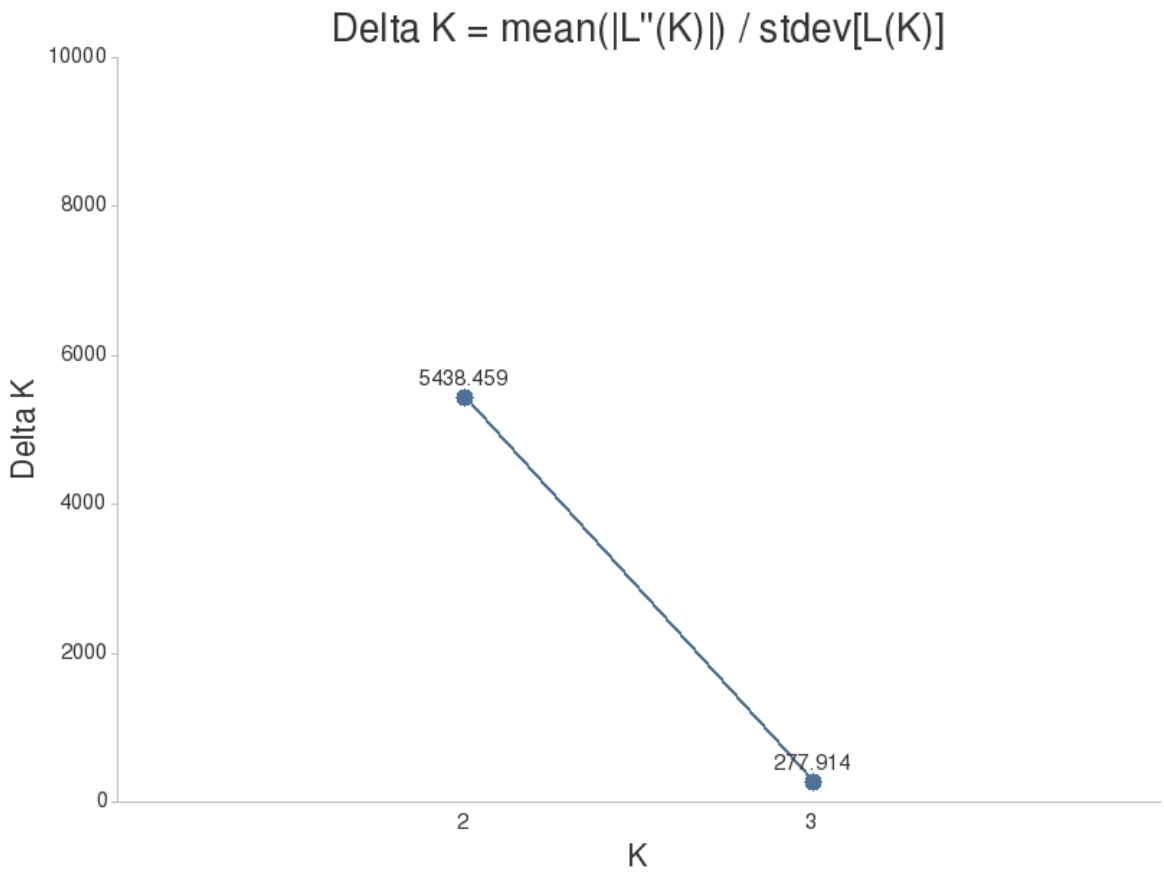


Figure S6. Plot showing the DeltaK statistic for each value of K as determined by the method of Evanno et al. (2005) implemented in CLUMPAK. Filtered data from Chapter 3 was used here.