POPULATION STRUCTURE IN YELLOWEYE ROCKFISH (*Sebastes ruberrimus*) DRIVEN BY LIMITED DISPERSAL AND SELECTION

by

MATTHEW RAY SIEGLE

B.A., The University of Colorado at Boulder, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2011

ABSTRACT

Natural selection and local adaptation influence fish abundances and distributions on both short and long-term time scales: first, by influencing recruitment dynamics and second, by affecting long-term population persistence in fluctuating environmental conditions. Little is known, however, about the significance of adaptive population divergence in the marine environment. In this study, I utilized an F_{ST} outlier approach to detect putatively selected loci from an Amplified Fragment Length Polymorphism (AFLP) marker dataset of yelloweye rockfish (*Sebastes ruberrimus*) collected from southeast Alaska south to Oregon. During the data analysis phase, a plate bias was detected and efforts to eliminate this effect were unsuccessful. Therefore, each PCR plate was analyzed separately, resulting in seven independent analyses. A total of 966-1580 AFLP loci were identified for the seven plates, and 0-10 loci (0-0.79%) per plate were identified as F_{ST} outliers by program, BAYESCAN. Based on results from the outlier analyses, datasets composed of neutral loci (those with no support for being under selection) and outlier loci (those exceeding a minimum posterior threshold of 0.7, corresponding to a moderate amount of support for being under selection) were established for each plate. Global F_{ST} values are approximately five times greater for the outlier datasets (mean F_{ST}=0.56) than for putatively neutral loci (mean F_{ST}=0.10). The genetic clustering program, STRUCTURE, identified similar patterns of population structure both for the neutral and outlier datasets. The outlier datasets, however, lead to overall higher assignment probabilities of individual fish to genetic groups than observed with the neutral datasets. In addition, several similar clustering patterns (e.g. southern vs. northern genetic clusters, presence of isolated clusters) were observed across several plates, providing independent evidence for regionally restricted gene flow. Overall, my results suggest that population structure driven by neutral processes may be reinforced by natural selection, which has implications for the management of yelloweye rockfish fishery stocks and the scale of connectivity within the Rockfish Conservation Area marine reserve network.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Jonathan Shurin and Dr. Eric Taylor, not only for their guidance over the last three years, but for their unconditional support while I developed my own project.

I'd like to thank my committee member, Dr. Michael Whitlock, for consistently providing great ideas that improved my project and for always being available to answer questions.

I want to thank Lynne Yamanaka for making the tissue samples available and for providing the majority of the funding for the genetic work. I'd also like to thank NSERC and Fisheries and Oceans, Canada, Inshore Rockfish program for additional funding.

There were many lab members, friends and colleagues that helped with R code, statistical analyses, field work, provided comments on the manuscript, were there for incredibly helpful discussion and provided much support throughout the trials and tribulations of AFLP trouble shooting.

Lab members, colleagues and friends:
Stefan Dick, Dr. Russel Markel, Dana Haggardy, Dr. Travis Ingram, Jon Mee, JS Moore, Julie Lee-Yaw, Rich Fitzjohn, Dr. Rebecca Martone, Dr. Alison Haupt, Dave Toews, Aleeza Gerstein, Gwylim Blackburn, Anne Dalziel, Kieran Samuk, Gina Conte, Theraesa Coyle, Andy LeBlanc, Dr. Alistair Blachford, Monica Yau, Carla Crossman, Lynne Yamanaka, Lisa Lacko, Dr. Mary O'Connor, Dr. Sean Rogers

I am greatly indebted to the Zoology Computing Unit (Dr. Alistair Blachford, Andy LeBlanc and Richard Sullivan).Without the resources and expertise provided by the ZCU I don't know how I would have finished my thesis.

Lastly, I would like to thank my fellow students that put in the time and effort to make our social events the successful events that they are. Beers and the annual Huts Party (amongst others) make the Zoology department an incredibly friendly place to work, and provide a necessary outlet for dealing with the stresses of graduate school.

INTRODUCTION

From an anthropocentric perspective the ocean lacks obvious barriers to dispersal. Moreover, approximately 70% of marine fishes and invertebrates exhibit extended pelagic larval durations (PLD) (Thorson 1950; Scheltema 1986), and have the potential for long distance larval transport via ocean currents. Together, these observations precipitated the idea that marine populations are demographically "open", exchanging larvae over vast geographic areas (Levin 2006). Following this paradigm, it was assumed that marine populations are genetically homogenous, exhibit huge effective population sizes, and that adaptation operates over large heterogeneous areas. This model was bolstered by early genetic work that found low levels of neutral genetic differentiation in many marine fishes, suggesting extensive gene flow over large geographic areas (Ward et al. 1994; Waples 1998). Gene flow is expected to dampen adaptive differentiation by breaking up favorable allelic combinations through recombination with migrant alleles (Slatkin 1973; Lenormand 2002). Thus, local adaptation is expected to be rare in the marine environment despite the diverse environmental conditions experienced over the geographic ranges of many marine species (Garant et al. 2007).

While the open population paradigm persisted for many decades, it is now well established that marine metapopulation dynamics can be exceedingly complicated, with evidence of self-recruitment (Jones et al. 1999; Swearer et al. 2002; Cowen et al. 2006), ephemeral population connectivity (Mitarai et al. 2008), and variation in post-settlement mortality (Planes & Lenfant 2002; Hamilton et al. 2008; Hauser & Carvalho 2008). Insight into this complexity requires us to change how we think about population boundaries in the ocean. Populations are often delineated on the basis of neutral genetic differentiation, most commonly with the statistic, F_{ST} (which describes the proportion of total genetic variance among a set of individuals that is attributable to differences among localities). Neutral genetic markers (those not influenced by selection) are preferentially used by researchers because the confounding effects that selection can have on allele frequencies are not present, and because neutrality is required for many population genetic models used to estimate demographic parameters, such as effective population size and

migration rate (Nielsen et al. 2009a). Whereas high neutral F_{ST} values certainly correspond to strong genetic structuring, low F_{ST} values (and an interpretation of low genetic structuring) may potentially be confounded by one of several scenarios, which might not be informative for inferring contemporary patterns of gene flow: (i) populations may be isolated, but not enough time has passed for signals of differentiation to develop; (ii) there is ongoing gene flow across sampling locations; (iii) migration amongst localities is enough to homogenize neutral allele frequencies, but not enough to alter the demographics of the receiving population (e.g., in terms of vital rates). (Conover et al. 2006). The last scenario is particularly troubling as it has been demonstrated that only a small amount of gene flow is sufficient to homogenize neutral allele frequencies (Mills & Allendorf 1996; Waples 1998; Palumbi 2003). This could arise from consistent, but ecologically insignificant dispersal over evolutionary time, or rare dispersal events that provide a large influx of recruits into an area. Additionally, low F_{ST} values commonly observed for many marine fishes may be a product of large effective population sizes that constrain the effects of drift, rather than high gene flow within a metapopulation (Allendorf & Phelps 1981). It is not surprising then that many population genetic studies of marine fishes utilizing neutral markers find limited genetic differentiation, even over broad spatial scales.

Most phenotypic traits are quantitative in nature; encoded by many genes that interact with each other and the environment to produce a phenotype (Naish & Hard 2008). Understanding the genetic variation of quantitative traits is difficult, and the link between divergence at neutral loci and loci underlying quantitative traits is tenuous at best (Merilä & Crnokrak 2001; Leinonen et al. 2008; Nielsen et al. 2009a). Furthermore, the spatial scale of neutral genetic variation is often decoupled from the scale of adaptive genetic variation, and differentiation at adaptive loci can occur on spatial scales much finer than those observed for neutral loci (Lynch 1996; McKay & Latta 2002; Latta 2003). Evolutionary processes may have variable effects on different regions of the genome, resulting in contrasting degrees of genetic differentiation. This pattern is termed heterogenous genomic divergence, and may arise through the effects of selection on genes underlying phenotypic traits (Nosil et al. 2009). Consequently, loci influenced by

selection across divergent environments (and regions genetically linked to these loci) may behave quite differently than neutral portions of the genome, and exhibit different population genetic signatures (Nosil et al. 2009). Thus, it is possible for gene flow to homogenize neutral allele frequencies at a larger geographic scale, while genetic differentiation is maintained at loci underlying phenotypic traits by selection on smaller geographic scales consistent with variation in environmental selective forces (Storz 2005).

Examples of fishery induced evolution (Marshall & Browman 2007), postsettlement selection (Marshall et al. 2010), and fisheries driven by discrete adaptations within a population complex (termed biocomplexity, Hilborn et al. 2003) have shed light on the importance of selection and adaptation in influencing fish abundances and distributions (Hauser & Carvalho 2008; Nielsen et al. 2009a). Selection may influence recruitment by functioning as a dispersal barrier and selecting against maladapted dispersers (Planes & Lenfant 2002; Vigliola et al. 2007; Hamilton et al. 2008; Allen & Marshall 2010). In this case, population genetic structure would be less related to larval dispersal patterns and more related to local environmental factors. For example, genetic differentiation of Atlantic and Icelandic cod (*Gadus morhua*) was shown to be two orders of magnitude greater at the selected *Pan I* locus (F_{ST}=0.261) than at nine microsatellites (F_{ST}=0.003), possibly due to an underlying adaptive role *Pan I* alleles have in different temperature regimes (Pampoulie et al. 2006). An order of magnitude greater differentiation was observed at the *Pan I* locus for walleye pollock (*Theragra chalcogramma*) and at the heat shock cognate 70 locus for European flounder (*Platichtys flesus*) than was observed at microsatellite loci (Canino et al. 2005, Hemmer-Hansen et al. 2007). These studies suggest that local selection pressures provide useful markers for teasing apart apparent genetic homogeneity, especially under scenarios of high gene flow.

Yelloweye rockfish, oceanographic and environmental context of the northeastern Pacific Ocean

Yelloweye rockfish (*Sebastes ruberrimus*) are widely distributed across the northeastern Pacific, inhabiting coastal rocky reefs found from the Aleutian Islands in Alaska through southern California; an area encompassing extensive oceanographic and environmental variation (Love et al. 2002). The adults are highly sedentary, and will often spend their time resting on the bottom substrate. Yelloweye rockfish are

generalist predators, preying upon both vertebrates and invertebrates, including: herring, sandlance, shrimps, crabs, and other rockfishes (Love et al. 2002). The size of an adult's territory depends upon the productivity of the environment, and generally, the more productive an area is the more sedentary the fish will be. Yelloweye rockfish are thought to be largely solitary, although aggregations of around 30 individuals have been observed on Bowie Seamount (Love et al. 2002). As with most benthic rockfishes, yelloweye rockfish are extremely slow-growing and late to mature reproductively, which can take around 20 years. Yelloweye rockfish are one of the largest of the rockfish species, and may live to over 100 years of age. The largest individuals can reach nearly a meter in length (Love et al. 2002). The long life span, and relatively large size of yelloweye rockfish make generalizations about age at length difficult (Love et al. 2002).

Across the range of yelloweye rockfish, ocean circulation in the outer coastal waters is dominated by the eastern flowing Subarctic Current, which divides along the continental shelf into the counter-clockwise flowing Alaska current and the clockwise flowing California current (Thomson 1981). The nearshore waters of the outer coast are regulated by the directional flow of the Davidson and Vancouver Island coastal currents (Figure 1) (Freeland et al. 1984; Jamieson & Phillips 1988). The continental slope and shelf off Vancouver Island, British Columbia (BC) are highly productive due to seasonal upwelling. In contrast, circulation patterns within the protected waters of the Strait of Georgia are typified by tidal currents and, in the summer, estuarine currents produced by freshwater runoff. In addition, the Victoria Sill, located at the eastern entrance to the Strait of Juan de Fuca, rises up from the bottom to -100 meters from the water surface and functions as a barrier to deep water renewal within the Georgia Basin (Thomson 1981). Moreover, there is tremendous freshwater input into the Georgia Basin during the summer months, owing mainly to the outflow of the Fraser River. Furthermore, the topography and bathymetry of British Columbia includes a large number of bays, fjords, islands and inlets, which complicate coastal circulation patterns in contrast to the relatively linear coastline of the continental United States (Hickey 1979).

Latitude, variation in seasonal upwelling, bathymetric and topographic rugosity, freshwater and sediment input, and bottom substrate are all important physical characteristics that define environmental gradients (e.g., temperature, nutrient and salinity) and are highly variable along the coastal waters of the northeastern Pacific. Consequently, across the geographic and environmental range of yelloweye rockfish there exists considerable potential for processes to influence the role both of neutral processes and natural selection in yelloweye rockfish population structure. Little is known, however, about yelloweye rockfish dispersal and recruitment (Love et al. 2002). Although sedentary as adults, dispersal takes place during the larval phase. Pelagic larval duration is unknown for yelloweye rockfish, however, a one to three month PLD has been observed for other demersal *Sebastes* species (Gunderson & Vetter 2006), and this may approximate the PLD of yelloweye rockfish. This extended PLD suggests the potential for long distance larval transport, and high rates of gene flow. Previous population genetic work by Fisheries and Oceans Canada (DFO) supported a high gene flow scenario. Yamanaka et al. (2006) found low differentiation across nine microsatellite DNA loci and report pairwise F_{ST} values ranging from zero to 0.024. Slight population structure was detected between fish sampled in the Strait of Georgia, and fish sampled elsewhere, resulting in two groups: a (inside) Strait of Georgia group and an (outside) outer coast group (SE Alaska, Haida Gwaii, and west coast Vancouver Island down to Oregon).

A useful framework within which to consider the scale and pattern of yelloweye rockfish population structure is the British Columbia Marine Ecosystem Classification system (BCMEC), described by Wainwright et al. (1995) and Zacharias et al. (1997). The BCMEC system is hierarchical, consisting of four nested divisions based on criteria identified as significant drivers of ecological processes (Table 1, Harding & Hirvonen 1996; Harper et al. 1993; Hirvonen et al. 1995). The smallest scale division recognized in the BCMEC is the Ecosection, of which there are 13 in BC (Figure 2). Due to the complexity of the British Columbia (BC) coastline, the area encompassed by each Ecosection is highly variable, ranging from 1,500 km^2 (Strait of Juan de Fuca) to over 170,000 km^2 (Subarctic Pacific). Each Ecosection is delineated

based on physiographic (bathymetric), oceanographic (current velocity, upwelling regimes, freshwater input, water turnover, wave exposure) and biological (migratory routes, productivity, species diversity) features. In addition, a fifth element (the Ecounit) was created to quantitatively compare the Ecosections on specific physical and oceanographic factors. The five Ecounits are wave exposure, depth, relief, currents, and substrate type. Each Ecosection can be evaluated on the proportion of its area that falls into one of several classes for each Ecounit. For example, wave exposure is composed of high (fetch > 500 km, ocean swell environment), moderate (fetch between 50 and 500 km, some ocean swell), and low (fetch <50 km, protected areas with very little swell) classes (Table 2). The proportion of area belonging to the high, moderate or low class of wave exposure within each Ecosection is calculated, and can be compared between Ecosections.

The extensive geographic distribution and highly variable environmental conditions yelloweye rockfish inhabit, along with putatively low neutral genetic structuring make them an ideal species for studying the role of adaptive genetic differentiation in the presence of high neutral gene flow. Loci that exhibit greater differentiation than is expected under neutral conditions (i.e. the locus is under selection or genetically linked to a locus under selection) are termed F_{ST} outlier loci. Consequently, I hypothesized that differences in allele frequencies at outlier loci would be associated with sample locations belonging to the most disparate Ecosections. I also expected that neutral gene flow would be driven largely by ocean circulation patterns, while the movement of putatively adaptive loci would be constrained by local selection. The inside/outside grouping in yelloweye rockfish identified by Yamanaka et al. (2006) is consistent with the strong alongshore movement observed in the outer coastal waters, and isolation of the Georgia Basin due to the presence of the Victoria Sill, which may represent a major oceanographic barrier to larval dispersal. Consequently, within these two broad oceanographic regimes, I expected that finer scale ecological variation would result in further population subdivision and reflect major environmental divisions captured by the Ecosection designations.

To test these ideas, I performed a genome scan for putatively selected loci utilizing an F_{ST} outlier approach on a large number of amplified fragment length polymorphism (AFLP) genetic markers. Amplified Fragment Length Polymorphisms represent a rapid and relatively inexpensive technique for generating large numbers of markers randomly distributed throughout the genome (Meudt & Clarke 2007).

Genetic analysis (described below) using AFLPs requires no *a priori* DNA sequence information, and have been used extensively for non-model species with no available genetic resources (Meudt & Clarke 2007). Because a large number of genetic markers can be generated with the AFLP method, they have been successfully employed in so-called genome scan studies, where a large number of loci are used to detect the footprints of selection and identify potentially adaptive loci. Wilding et al. (2001) found approximately 5% of AFLP loci assayed display a higher than expected level of differentiation within two morphs of the marine snail, *Littorina*, suggesting divergent selection reduces gene flow between the morphs across the vertical gradients within the intertidal. Bonin et al. (2006) used AFLPs in a genome scan and identified eight loci as candidates for adaptation to altitude in the common frog (*Rana temporaria*). Paris et al. (2010) detected five loci positively associated with resistance to a bio-insecticide in the mosquito, *Aedes rusticus*.

The main objectives of this study are to (a) identify markers putatively under selection and (b) address if adaptive genetic differentiation differs from neutral genetic variation and how that informs our assessment of population structure of yelloweye rockfish in the northeastern Pacific Ocean. I specifically address four main questions: (i) Will F_{ST} outlier loci be detected across the sampling range as a result of selection? (ii) If F_{ST} outlier loci are detected, do they exhibit greater isolation-by-distance (genetic differentiation increases with geographic distance (Wright 1943; Rousset 1997)) than is detected with neutral loci? (iii) Will neutral loci resolve a similar inside/outside grouping that was detected by the DFO microsatellite dataset (coincident with major oceanographic regime differences, and isolation of the Georgia Basin due to the Victoria Sill)? (iv) If detected, do outlier loci resolve population subdivision in accordance with major environmental divisions within the inside and outside populations?

To my knowledge, this approach has only been applied to marine fishes in a handful of cases, and this is the first study to address adaptive genetic differentiation within *Sebastes* taxa.

MATERIALS AND METHODS

Sample collection

Yelloweye rockfish samples were collected in the coastal waters of British Columbia (including Bowie Seamount, approximately 220 km west of Haida Gwaii), southeast Alaska, Washington, and Oregon (Figure 3, Table 3). Fish were sampled during DFO research surveys and opportunistically from commercial fishery vessels from 1998-2009. Fish were sampled either onboard the fishing vessel or dockside and fin clips were stored in 95% ethanol for genetic analyses.

AFLP methods

Vos et al. (1995) describe the development of genetic markers though the selective amplification of restriction fragments. They termed this type of genetic marker, amplified fragment length polymorphisms (AFLPs), which are a relatively inexpensive and rapid way to accumulate markers randomly dispersed throughout the genome. To generate AFLPs genomic DNA is digested with restriction enzymes and ligated to short oligonucleotide adapters at the restriction sites. These adapters (whose sequences are known) are then used to generate primers for two successive rounds of polymerase chain reaction (PCR). The first round, called the preselective amplification, uses a primer with an additional base that extends past the adapter, theoretically reducing the number of fragments that will amplify by a factor of four. The final PCR is known as the selective amplification, and primers that contain an additional two selective bases are used, further reducing the number of DNA fragments that will amplify by a factor of 16. The final PCR products are separated via electrophoresis. Each fragment represents a locus, and an individual is scored at each locus for the "presence" or "absence" of that particular fragment. The AFLPs are dominant markers, as the presence of a fragment masks the potential absence of a fragment at the homologous locus.

Total genomic DNA was extracted from fin clips using a standard phenol-chloroform extraction protocol. The DNA concentrations of samples were quantified on a Nanodrop 8000 (ThermoScientific) and adjusted to a final concentration of 25-75 ng/μl.

Restriction digest and ligation

Total genomic DNA was digested with the restriction enzymes EcoR1 and Mse1 at 37°C for three hours, followed by a 70°C, 15 minute enzyme inactivation (all reagents from New England Biolabs (NEB), 20μl reaction: ~500ng genomic DNA, 10μM BSA, 1X EcoR1 buffer, 2.5 and 6.5 units of EcoR1 and Mse1, respectively). Specific oligo adapters were ligated to the ends of the restriction fragments with T4 DNA ligase at 37°C for 3 hours (5 units T4 DNA ligase, 1X T4 ligation buffer, 100 μM M-adapter, and 100 μM E-adapter).

Pre-selective amplification

The restriction/ligation product was diluted 10X for use as the template in the preselective amplification. PCR was performed in a 20 μl reaction of 0.25 mM each dATP, dCTP, dCTP, dTTP (NEB) with 1.5 mM $MgCl_2$, 1X reaction buffer, and 5 units Taq polymerase (NEB). The PCR program consisted of: 94°C-2 minute, followed by 20 cycles of 94°C-30 sec, 56°C-30 sec, 72°C-1 minute, and a final 72°C 10 minute extension.

Selective amplification

The preselective amplification product was diluted 20X for the selective amplification (10 μl reaction: touchdown PCR program). Four EcoRI and four Mse1selective primers, each with three selective nucleotide bases were used in combination for a total of 16 primer pairs (Table 4). E-selective primers were fluorescently labeled, and fragment sizing was performed on an ABI 3730xl (Applied Biosystems) by Génome Québec (Montreal, Québec, Canada).

AFLP scoring

The electropherograms of AFLP profiles were first visualized by eye in GENEMAPPER, version 3.7 (Applied Biosystems) and individuals with obvious PCR failures were removed prior to scoring. Fragment sizes and bin positions were determined by the AFLP analysis method in GENEMAPPER. Bins were set at a width of one base pair (a bin corresponds to one AFLP locus – all peaks within this size range are scored as homologous alleles) and only fragments within the size range of 50-400 base pairs having a minimum

peak height of 50 relative fluorescence units (rfus) were retained. Peaks with heights less than 50 rfus were not scored by GENEMAPPER. Peak height data for all initial bins was used in the program AFLPSCORE (Whitlock et al. 2008) to filter the initial bin set and determine phenotype scores for each individual. AFLPSCORE utilizes a peak height normalization factor to account for variation in PCR amplification intensity that may contribute to errors in AFLP scoring. For example, discerning a true AFLP fragment from background noise within an electropherogram profile containing peaks that exhibit consistently low fluorescence can be difficult. Additionally, determining which fluorescence threshold to use for assigning a fragment 'presence' or 'absence' score may introduce erroneous errors when two electropherograms exhibit differences in overall fluorescence intensities. AFLPSCORE minimizes the subjectivity created by manual scoring when differences in amplification intensity exist by automating the criteria used to assign presence/absence scores.

Within AFLPSCORE two thresholds are defined by the user: the locus selection threshold and the phenotype-calling threshold. The locus selection threshold is used to remove loci that exhibit low mean peak heights (which can be hard to distinguish from background noise) and are likely to contribute to unrepeatable genotypes. For each bin, the mean peak height from all individuals is calculated from the normalized data and only those bins greater than the locus selection threshold are retained. The phenotype-calling threshold is used as a cutoff when assigning presence/absence scores for the loci retained after the locus selection filter. An absolute peak height threshold (e.g. 100 normalized rfus) or a relative peak height threshold (e.g. 10% of the bin's mean peak height) may be used. Replicated samples are used to assess scoring error rates under a suite of different thresholds, allowing the user to optimize scoring thresholds and minimize genotyping errors. For my study, all primer pairs were analyzed independently and a range of 91-181 replicated individuals (14.5%-28.9% of total sample) were used to assess error rates (please see Appendix A for a more thorough description of AFLPs and the scoring procedure used in AFLPSCORE).

Plate bias in AFLP genotyping

After an initial dataset was created from AFLPSCORE I proceeded with downstream analyses, including multi-dimensional scaling analyses and assignment tests. Upon analysis of this initial dataset, several ambiguities were identified due to plate-to-plate variation in amplification intensity (i.e. plate effect bias): genetic similarity was greatest amongst individuals assayed on the same PCR plate regardless of sampling location or other geographic or biological differences between sample locations. There are a number of steps in the AFLP procedure where laboratory variation or error may introduce these plate effects: differences in DNA quality, incomplete restriction digestion, variation in the ligation of adapters to the restriction fragments, differential amplification of the preselective and selective PCRs, variation in the migration of fragments in the capillary of the genotyping machine between runs, or even by different conditions the PCR plates were exposed to while in transit to the sequencing facility, Génome Québec.

DNA quality was checked on a Nanodrop 8000, and all samples were diluted to similar concentrations. All restriction digests, ligations and preselective amplifications were performed with the same stock of reagents, and done within a three week time period. Proper care was taken to ensure the reagents were in good condition. Products from representative samples for the digestion and PCR amplifications were run on agarose gels to check for digestion and amplification. All selective amplifications were shipped overnight to Génome Québec frozen, on ice packs, and wrapped in foil to prevent UV degradation of the fluorescent label. There were, however, several problems with the shipping company, including one lost package and two accidentally sent ground instead of priority overnight, and it is possible that the different shipments were exposed to varying conditions en route. Additionally, it is possible that there were inconsistencies in how the PCR products were injected into the capillary sequencer and the electrophoretic conditions under which the samples were run. Despite these efforts to ensure consistency, variation at one of the laboratory steps, or a culmination of slight variations during many of the steps facilitated enough variation amongst my PCR plates as to precipitate this plate bias.

Despite the potential for plate effects to exist, there is surprisingly little written about dealing with these biases in the literature. A few AFLP scoring programs indirectly deal with this effect by reducing variation in amplification intensities between samples (GENOGRAPHER, Behham et al. 2009; AFLPSCORE; and RAWGENO, Arrigo et al. 2009) by normalizing peak heights. While I did employ AFLPSCORE, accounting for variation in amplification intensity does not remedy other potential sources of plate effects, including variation in background noise or the preferential amplification of fragments between plates.

To resolve this problem, I used a generalized linear model to assess the predictive power of "plate" on the genotypes for all initial AFLP loci. An alpha value of 0.05 (i.e., plates for which there was no significant effect of plate identity) were retained for the final dataset. This step reduced the number of loci in the final dataset by 75%, but, unfortunately, failed to completely eliminate the plate bias.

To address the remaining plate bias, I decided to treat each plate as an independent test (a "project") of my hypotheses, and eliminate between plate comparisons altogether. With only one plate per project, I effectively removed the two major sources potentially contributing to a plate effect: variable PCR background noise and differential selective amplification between plates. The major limitations of this approach are the reduction in sample size for each sample location and the inability to analyze and compare all sample locations together. The strength of the plate effect was so great, however, that it would have been impossible to differentiate any biological signal from that of the plate effect.

My analyses consisted of a total of seven plates, six of which contain four sample locations with 10-24 individuals (mean of 23) per sample location. The remaining plate contained two sample locations with 37 and 46 individuals per location (Table 5). All my analyses were performed for each of the seven plates independently, and while this approach negates the ability to analyze and compare all sample locations together, there is the benefit of having several of the major oceanographic and environmental contrasts replicated in independent analyses (for a more detailed discussion regarding plate effects and my scoring procedure please see Appendix B).

Summary of final AFLP scoring method

All electropherograms were first screened by eye to remove obvious PCR failures and individuals with unusual profiles (e.g., peaks absent from portions of the size range that contain many peaks in the majority of individuals, individual profiles contain high variation in peak height). The first preliminary bin set was generated in GENEMAPPER and corresponding peak height data for all individuals were used in AFLPSCORE to eliminate error prone loci and determine final AFLP phenotypes. Individuals were then filtered by the amount of missing data and only individuals with less than 15% missing data were retained. The bin set for each plate was filtered one last time, and, as recommended by Bonin et al. (2004), only bins with a fragment presence frequency greater than 5% were retained. Thus, a binary table was generated for each plate containing only individuals with less than 15% missing data and only bins that had a minimum fragment presence frequency greater than 5%.

Data Analysis

AFLP genome scan for outlier loci

The goal of this study was to identify putatively adaptive loci (F_{ST} outliers) and compare the scale of adaptive genetic differentiation to that derived from neutral loci. The program BAYESCAN (Foll & Gaggiotti 2008, http://cmpg.unibe.ch/software/bayescan) was used to detect F_{ST} outliers. BAYESCAN is an extension of the hierarchical Bayesian method described by Beaumont & Balding (2004). The program estimates population-specific F_{ST} coefficients, allowing for different demographic scenarios and levels of genetic drift in each population. Each value of genetic differentiation, F_{ST} (*i, j*) for locus *i* in population *j* is decomposed as a linear combination of a locus effect (α_i) and a population effect (β_j). Outlier loci are identified by focusing on the posterior distribution of α_i. If a locus effect (where $\alpha_I \neq 0$) is necessary to explain the overall pattern of genetic diversity, then the locus is interpreted to be under selection and identified as a putatively adaptive marker. A positive value of α_i suggests that locus *i* is subject to positive selection and a negative value suggests the locus is subject to stabilizing selection. To test for the significance of α_i for each locus *i*, the posterior probability that α_i is different from zero is calculated using

a Reversible Jump Markov Chain Monte Carlo (RJ-MCMC) algorithm. The RJ-MCMC algorithm compares the posterior probabilities of two models for each locus: a neutral model (alpha = 0, M1), and a selection model (alpha ≠ 0, M2). Model choice is performed by using "Bayes factors" (BF), where the BF provides a scale of evidence for one model over another. The BF for model M2, for a given data set *N*, is given by the equation:

(1) BF=P(N|M2)/P(N|M1)

If BF=2, then the data favors M2 over M1 at odds of two to one. The idea of spurious results resulting from multiple simultaneous hypothesis tests, however, needs to be incorporated into posterior probability distribution. This is achieved in BAYESCAN by using the posterior odds (PO) instead of BFs to make decisions:

(2) PO=P(M2|N)/P(M1|N)

Posterior odds are the ratio of posterior probabilities and indicate how more likely the model with selection (M2) is compared to the neutral model (M1).

The estimation of model parameters was automated on the basis of short pilot runs (20 pilot runs, length 5,000). I used a sample size of 5,000 and a thinning interval of 50, resulting in a total chain length of 300,000 iterations. When utilizing BAYESCAN, Foll & Gaggiotti (2008) recommend defining F_{ST} outlier thresholds by setting a target False Discovery Rate (FDR), and looking for the highest posterior odds threshold achieving this threshold. For this study, a FDR of 0.05 was used.

In order to get an overall average effect of non-neutral loci on population structure, I created outlier and neutral datasets for each plate based on the results of the BAYESCAN analyses. I included loci in the outlier dataset that did not exceed the minimum posterior probability for outlier status in the BAYESCAN analysis, yet still exhibited moderate support for non-neutrality (posterior probability >0.70). This threshold was chosen because I wanted to include as many loci as possible in the outlier datasets that still exhibited some support for non-neutrality, and not have to rely on analyses based on one or two loci due to the presence of the plate bias. A higher minimum threshold would have forced me to drop many

putative outliers.

The neutral dataset corresponds to loci found by BAYESCAN to have virtually no support for being under selection (posterior probability <0.2). Analyses were run for both the outlier and neutral datasets to evaluate the influence of putatively adaptive loci on population structure, and contrast that with inferences based on putatively neutral AFLP markers.

Population genetic data analysis

Mean genetic diversity was assessed with AFLP-SURV (Vekemans 2002) based on (i) the proportion of polymorphic loci (PLP) and (ii) the expected heterozygosity averaged across loci (H_J). The AFLPs are dominant markers, and the occurrence of heterozygotes is, therefore, not detectable. Consequently, the Bayesian method of Zhivotovsky (1999) was implemented in AFLP-SURV to estimate allele frequencies. Genetic differentiation between sample locations was estimated as the unbiased F_{ST} estimator of Weir & Cockerham (1984), and 95% confidence intervals were obtained by bootstrapping 1000 replicates over loci. Isolation-by-distance (IBD) refers to the relationship between genetic differentiation and geographic distance between localities. This pattern is often found in wild populations (Kinlan & Gaines, 2003) and has been described in several marine fishes (Pogson et al. 2001; Buonacorrsi et al. 2002; Bradubury & Bentzen, 2007; Johansson et al. 2008). To test for IBD in yelloweye rockfish, I performed Mantel (1967) tests between the pairwise comparisons of linearized F_{ST} (F_{ST} / (1-F_{ST})) and geographic distance (in kilometers) in the R environment (R Development Core Team 2011). I ran 10,000 permutations to assess significance.

The program STRUCTURE version 2.3 (Pritchard et al. 2000; Falush et al. 2007; admixture model, no sample location priors) was used to determine the number of genetic clusters (*K*) among my sample locations and to estimate individual assignment probabilities for each fish to each resolved cluster. The number of putative clusters assessed ranged from 1 to 5 for the outlier and neutral datasets. Each run consisted of a 150,000 step burn-in with an additional 150,000 steps, and 10 iterations were run for each *K*-value. I also used the online resource STRUCTURE HARVESTER (Earl 2011,

www.taylor0.biology.ucla.edu/struct_harvest/index.php) to calculate the ΔK statistic (the rate of change in the log probability of data between consecutive *K*-values) according to Evanno et al. (2005), that addresses a potential statistical artifact produced by STRUCTURE that results in higher likelihoods and variance with larger *K* values, and can complicate inferring the true number of clusters in a data set.

RESULTS

AFLP delineation and scoring

A total of 672 samples were collected from 15 locations across the northeastern Pacific. The most northern sampling location was near Sitka, Alaska (57° 10.82' N, 136° 4.46' W), and the most southern location was off the coast of central Oregon. Most of the sampling effort was focused around Vancouver Island. Forty-six individuals (6.8%) contained at least 15% missing data (i.e. poor quality PCR profiles that corresponded to 15% of the total number of AFLP loci), and were dropped from the study, leaving a total of 626 fish that were genotyped.

A total of 6281 bins were identified in AFLPSCORE from the original 16 primer pairs. Six of these primer pairs, however, had mismatch error rates greater than 10% and were dropped from the final dataset (Table 6). The remaining 10 primer pairs yielded a total of 4,209 bins, which were filtered within each plate's binary (presence/absence) table according to the fragment presence frequency for each bin. Bins that contained a fragment presence frequency of less than 5% were discarded. The final number of bins for each plate ranged from 966-1,580 (mean of 1286). The average fragment presence frequency for each plate ranged from 0.31-0.45 (mean of 0.39) (Table 5).

Loci putatively under selection

With a FDR of 0.05, the identification of outlier loci corresponded to a minimum posterior probability of 0.89-0.96 of a locus being putatively subject to selection across the seven plates. At least one outlier was identified for six of the seven plates (Table 7), and the number of outliers detected ranged from 0-10 (0-0.79% of total number of loci). Five outlier loci were identified on more than one plate: one common to Plates 01 and 05, and four common to Plates 06 and 08. The number of loci with a posterior probability of being under selection greater than 0.7 (minimum for inclusion in the outlier datasets) ranged from 3-18 across the six plates where outliers were detected. In contrast, most loci had posterior probability values considerably lower than 0.70 and the number of neutral loci comprising each plate's neutral dataset ranged from 928-1,576.

Population genetic diversity and differentiation

AFLP-SURV was used to calculate total heterozygosity (H_t) which was approximately twice as high for the outlier datasets as the neutral datasets (mean for outlier datasets: 0.54; mean for neutral datasets: 0.27) (Table 8). The proportion of polymorphic loci (PLP) and Nei's gene diversity (*Hj*) were positively correlated for sample locations within each plate's neutral dataset and each plate's outlier dataset. (Spearman's rank correlation coefficient: neutral dataset, r_s=0.89, p-value<0.001; outlier dataset, r_s=0.85, p-value<0.001).

Global F_{ST} values for the outlier datasets ranged from 5-10x greater than those of the neutral datasets (mean for outlier datasets: 0.56; mean for neutral datasets: 0.10) (Table 8). Mantel tests for IBD were non-significant both for the neutral and outlier datasets for all plates (mean p-value for neutral: 0.67; outlier: 0.39). There was a trend, however, towards greater significance in the outlier datasets in two out of six plates (Plates 02 and 06; r = 0.86 and 0.54, respectively) (Table 9).

Population structure

Regional differentiation

The STRUCTURE analysis of population structure provided the greatest amount of support for two or three genetic clusters across all seven plates, both for the outlier and neutral datasets. Log probabilities steadily increased with increasing *K*-values, however, Δ*K* values were consistently highest between a *K* of 2 and a *K* of 3 for all neutral and outlier datasets. Geographic clustering patterns were observed on all seven plates, and were stronger with the outlier dataset, as expected because the outlier loci were identified on the basis of being the most differentiated between sample locations (Table 10).

The predominant oceanographic division was between the outside waters west of Vancouver Island and the inside waters east of Vancouver Island. Across the seven plates, this "outside-inside" comparison was tested four times where individual plates contained samples from outside and inside waters around Vancouver Island (Plate 01: BS/ES/TA vs. TI; Plate 02: CI/GC/WA vs. GA; Plate 03: OR/TK vs. LB/MI; and Plate 06: CI/OR vs. LB/TI). Three of these comparisons exhibited the outside vs. inside split.

By contrast, no "inside-outside" split was observed on one plate (Plate 02, where GA and WA clustered together) (Table 11). Interestingly, on both plates that contained two inside sampling locations (Plate 03: LB and MI; Plate 06: LB and TI), the sample locations did not cluster together (Appendix C, pages 74 and 84). In addition, all seven plates contained at least two outside sampling locations, yet none of these plates resulted in all individuals from outside sites being assigned to the same genetic cluster (Table 11).

Ecosection differentiation

For 24 different Ecosection comparisons across the seven plates, only 12 discernible clusters corresponded exclusively with one Ecosection (Table 11). For example, Plate 01 contained four sample locations from four different Ecosections and was most likely represented by three different genetic clusters. Of these four locations, individuals from Bowie Seamount (BS) and Texada Island (TI) both formed a unique cluster. Fish from Esperanza (ES) and Tasu (TA), however, both strongly assigned to the third cluster. So, while BS and TI belong to different Ecosections and formed clusters distinct from one another, fish from ES and TA formed a single genetic cluster yet belong to distinct Ecosections (Appendix C, page 66). In addition, on Plate 03, fish from Oregon strongly clustered based on Ecosection, yet fish from Mittlenatch (which exhibited strong assignment to a single cluster) did not cluster with individuals from the Lower Bute Inlet location, despite belonging to the same Ecosection (Appendix C, page 74).

Several clustering patterns were observed on more than one plate, providing independent verification of some of the clustering patterns. For instance, Bowie Seamount was identified as a unique cluster on Plate 01 and Plate 05 (Appendix C, pages 66 and 80). Clusters discerned by latitude (relative southern and northern clusters) were identified on three plates. Southern clusters include GA/WA (Plate 02), CJ (Plate 04), and OR (Plate 06) vs. northern clusters CI/GC (Plate 02), AK (Plate 04), and LB/CI/TI (Plate 06) (Appendix C, pages 70, 78 and 84).

DISCUSSION

What is the effect of the plate bias on AFLP scoring and data interpretation?

The plate bias detected in my dataset was of such severity that efforts to mitigate its effects were unsuccessful. My solution to account for plate bias was to eliminate between plate comparisons altogether. Given the amount of AFLP literature dedicated to minimizing scoring error resulting from variation in amplification intensity (Bonin et al. 2004, 2005, 2007; Ehrich et al. 2006; Meudt & Clarke 2007; Holland et al. 2008; Whitlock et al. 2008; Arrigo et al. 2009; Foll et al. 2010; Gaggiotti 2010; Herrmann et al. 2010), it is surprising that little is written about a plate bias, which is indicative of plate specific fragment amplification in addition to potential overall amplification intensity differences. The only source I found that even loosely addresses this type of bias is the RawGeno Manual (see Arrigo et al. 2009). Arrigo et al. discuss how the presence of individuals with a low or high number of detected fragments may represent failed or unusual PCR amplifications. Within RawGeno, it is recommended that individuals below and above certain quantiles should be removed prior to further scoring and analysis. The criteria for what constitutes a 'low' or 'high' number of detected fragments, however, remains unclear, and basing these thresholds on quantiles entails that it is always possible to trim away the ends of the distribution without knowing what constitutes an 'acceptable' number of fragments. Moreover, it is not necessarily only differences in the number of amplified fragments between plates that may introduce a plate bias, but a difference in specific fragment amplification is also an important factor.

My solution to the plate bias issue was to remove between plate comparisons altogether, a solution that introduced two major complications: interpreting the mismatch error rates and characterizing general results across all the individual plate results. Estimating scoring accuracy is an important part of any genotyping study, as even relatively minor scoring errors can have a significant effect on the results (Bonin et al. 2004; Pompanon et al. 2005). Unfortunately, the ambiguous nature of the plate bias and subsequent independent plate analyses makes interpretation of the scoring accuracy difficult to assess. For my study, mismatch error rates (mean 8.3%) are higher than has been found with other AFLP studies (Hansen et al.

1999; Bonin et al. 2004, 2007), which put error rates in the 2-5% range. As my replicate individuals were on separate plates (two pairs of replicate plates with fundamentally different amplification profiles), it is most likely the pervasive plate bias that is responsible for the inflated error rates. In this vein, my error rates probably represent upper bounds, and within-plate scoring accuracy is higher, but there is no way to quantify the exact error rates.

Further, the bin set for each plate was filtered for low frequency bins, and bins containing a low number of presence scores were discarded. As the rarity of presence scores within a bin increases, there is a corresponding decrease in the likelihood that that bin represents a true AFLP locus, and more likely corresponds to stochastic amplification (resulting from high intensity stutter bands, background noise, or a fragment resulting from non-specific restriction site cutting). Thus, in addition to a repeatability filter, bins were filtered for presence rarity, and the final bin sets met a recommended minimum presence frequency of 5% (Bonin et al. 2004). In summary, while the plate bias complicates interpretation of the scoring error rates, it is probable that they represent an upper bound of error rates, and within-plate scoring accuracy is actually higher. In addition, removing low frequency bins increases the chances that the remaining bins represent "true" AFLP loci, and are not simply the product of stochastic amplification.

The greatest effect of scoring errors on my study is the potential incorrect identification of outlier loci. Due to the ambiguity surrounding my error rates I focused more on broad patterns based on many loci (the outlier and neutral datasets) than on specifics with regards to individual loci. It is hoped that even with the incorrect identification of some outlier loci, the average effect of non-neutrality on population structure was accurately assessed.

Under ideal conditions, there would be no plate bias and all PCR plates would have been analyzed as one project. Including all of the sample locations in a single analysis would, theoretically, lead to a greater number of outlier loci detected (as sample locations occupying the total amount of geographic and environmental variation would be included), and I could have performed more straight-forward analyses of population structure of yelloweye rockfish in the northeastern Pacific Ocean. Unfortunately, my only

solution to the plate bias was to analyze each plate independently, resulting in smaller sample sizes and restricting analyses to locations that do not encompass the entire sample range. The number of outlier loci I detected for each plate (0-0.79% of the total number of loci) was lower than what has been reported in other genome scan studies (average of 3-10%, Nosil et al. 2009). This is consistent with expectations given the reduction in sample size for 13 out of the 15 sample locations.

The greatest challenge in dealing with this plate bias resides in trying to generalize and summarize patterns from seven plates' worth of independent analyses. This is partly why I used the oceanographic framework (supported by DFO's microsatellite DNA data) and the environmental divisions of the BCMEC to frame my hypotheses, as both the oceanographic and environmental boundaries are clearly defined, and allowed me to directly test spatial patterns of genetic clustering by counting the number of coincidences where the genetic cluster boundaries (identified from the STRUCTURE analyses) fall within the inside/outside or Ecosection divisions.

Expected versus observed patterns of population structure and the feasibility of the BCMEC framework

Neutral and outlier loci were expected to resolve different patterns of population structure within the oceanographic and environmental framework of the BCMEC. Neutral loci were expected to confirm an "inside-outside" division corresponding to high alongshore flow in the outer coastal waters and respective isolation of the Georgia Basin, due to the Victoria Sill in the south and the shallow bathymetry of the Broughton Strait in the north (Yamanaka et al. 2006). Outlier loci were hypothesized to resolve population structure on a finer scale within the inside-outside groups in accordance with the major environmental divisions embodied in the Ecosection designations. The inside-outside division was detected in three out of four plates that contained both inside and outside sampling locations and, therefore, my study suggests that it likely represents a significant barrier to gene flow. As my data are based on a different genetic system than Yamanaka et al. (2006), the consistent inside-outside division shown in both marker sets adds strength to this inference. Neutral population structure within the outside oceanographic division,

however, did not meet expectations. I hypothesized that for neutral loci, fish sampled from outside locations would assign to the same genetic cluster. This hypothesis reflects high alongshore water flow in the outer coastal waters, the small amount of neutral gene flow sufficient to homogenize neutral allele frequencies, and reflects the previous findings of Yamanaka et al (2006). Moreover, I hypothesized that outlier loci that are putatively under selection would be necessary to resolve population subdivision against the background of high neutral gene flow within the outside division. Furthermore, I expected these outlier loci to exhibit an isolation-by-distance pattern. While I did not detect any IBD signals for any of the neutral or outlier datasets, there was a slight trend towards greater significance with a couple of the outlier datasets. The IBD comparisons, however, were restricted to localities between plates, resulting in only six pairwise comparisons within six of the seven plates. Detecting IBD with so few comparisons would be difficult.

Out of 24 total Ecosection comparisons across the seven plates, only 12 comparisons yielded an association between Ecosections and unique genetic clusters. The BCMEC framework was used as a starting point for partitioning the environmental variation observed in my study region and assigning break points between the major divisions. There are several reasons, however, to think that the Ecosection divisions might not capture the relevant environmental variation potentially related to adaptive differentiation in yelloweye rockfish. Namely, there is tremendous variation in area between the different Ecosections, which can lead to problems with spatial scaling. For example, the Bowie Seamount sample location is located within the Subarctic Pacific Ecosection (encompassing over 170,000 km^2), which is predominantly characterized by features of the abyssal plain and the subarctic current. The Bowie Seamount, however, is an isolated seamount that rises over 3,000 meters from the bottom to less than 50 meters from the surface. Fine-scale oceanography associated with the Bowie Seamount (e.g. the Haida eddies, Dower et al. 2001) most likely has greater influence over local environmental conditions that regulate rockfish recruitment and settlement. By contrast, the Lower Bute Inlet and Mittlenatch sampling locations are positioned in the Johnstone Strait Ecosection, an area encompassing approximately 2,500 km^2. The oceanographic and environmental features used to define this Ecosection are much more likely

to encompass the entire Ecosection compared with the characterization of areas that are two to three orders of magnitude larger.

Ideally, genome scan studies investigating adaptive divergence along environmental gradients will couple extensive and finer-scale environmental and life-history data along with the genetic sampling, and can then employ more rigorous methods for identifying putatively adaptive loci coincident with specific environmental features or life-history characteristics (see Foll & Gaggiotti 2006; Joost et al. 2007). Gaggiotti et al. (2009) sampled 19 spawning locations of Atlantic herring (*Clupea harengus*) and collected data for ten environmental and life-history variables at each of the spawning sites. They used the approach of Foll & Gaggiotti (2006) to identify environmental and/or life-history traits that were significantly correlated with outlier loci and likely major drivers of population structure. The strength of this approach lies in the ability to collect data for a suite of environmental and life-history characteristics. For many marine fishes and research programs, however, collecting this degree of environmental data is logistically and technically not feasible, as sample locations are often extremely remote, and may be hundreds of meters beneath the water surface. For species where collecting data to employ more rigorous methods for detecting selection is not feasible, gross environmental distinctions, such as that made with the BCMEC system, provide a starting point for evaluating genetic differences and may identify target areas for future, more detailed studies. My results, therefore, suggest that finer scale environmental characterization is required for more sensitive investigations of potentially adaptive loci in yelloweye rockfish.

How common is regional population structure within the rockfishes?

Despite the importance of many rockfishes both to commercial and recreational fisheries (Parker et al. 2000), there are a limited number of studies that investigate regional population genetic structure within the genus. Sivasundar & Palumbi (2010) report a wide variety of genetic patterns among 15 species of rockfishes investigated with a coarse (three site) sampling scheme along the west coast of the United States. Strong genetic breaks between southern California and Oregon were observed for the blue (*S. mystinus*) and yellowtail (*S. flavidus*) rockfishes. The rest of the 15 species exhibited moderate to weak genetic

differentiation across the sampling range. Interestingly, species within the subgenus *Sebastosomus* exhibited stronger signals of differentiation than species within the subgenus *Pteropodus*, despite having a reported longer PLD (Carr & Syms 2006). Significant mitochondrial DNA differentiation was observed in rosethorn rockfish (*S. helvomaculatus*) sampled from five locations, from the Gulf of Alaska down to San Francisco, California (Rocha-Olivares & Vetter 1999). Gilbert-Horvath et al. (2006) investigated population structure within populations of kelp rockfish (*S. atrovirens*) and did not find any substantial genetic differentiation over 800 km of California coastline.

Buonaccorsi et al. (2002) investigated population structure in copper rockfish (*S. caurinus*) sampled from Haida Gwaii to southern California, including two sites within the Georgia Basin: Puget Sound and the Canadian Gulf Islands. Outer coast sites exhibited weak yet significant differentiation from each other at six microsatellite loci (F_{ST}=0.007), while the Puget Sound and coastal samples exhibited much greater differentiation (F_{ST}=0.087). The Gulf Island site was genetically intermediate to, and distinct from, the outer coast sites and Puget Sound. A significant isolation-by-distance (IBD) pattern was observed for the four outer coast sites. Johansson et al. (2008) followed up on the results of Buonaccorsi et al. (2002) to see whether or not the coast-wide IBD pattern would hold if the number of microsatellite markers, the sample size, and the number of sample locations were doubled. Significant population subdivision (F_{ST}=0.004) was found at a coast-wide scale (northern Washington to southern California; ~2,200 km), and a significant coast-wide IBD pattern was observed. By contrast, within the finer scale of the Oregon coast (~450 km), a single panmictic population was detected with individual based clustering analyses, and no IBD pattern was observed. At a similar spatial scale (~500 km), similar findings were observed for copper rockfish sampled along the west coast of Vancouver Island, although populations located deeper within fjords showed marked differentiation from outer coast samples (S. Dick, University of BC, unpublished data).

The general consensus is that weak yet statistically significant regional genetic structure is common within the rockfishes, and results from Sivasundar & Palumbi (2010) suggest that genetic structure corresponds with major oceanographic breaks. In the temperate northeastern Pacific, oceanographic

divisions driven by gyres of the Alaska and California currents and major topographic features, such as Point Conception in southern California, greatly influence the species composition of rocky reef fishes, which loosely defines four biogeographic provinces: the warmer Cortez and San Diego provinces, and the colder Oregonian and Aleutian provinces (Gunderson & Vetter 2006). Within these provinces, mesoscale processes, such as fronts associated with upwelling or the intersection of different current systems (Shanks et al. 2000), gyres above banks and seamounts, or upwelling shadows on the lee-side of headlands may act to retain eggs and larvae (Gunderson & Vetter 2006). Interestingly, many rockfishes have distributions that span all of these provinces, while other groups of fishes have much more restricted distributions (Gunderson & Vetter 2006).

Yelloweye rockfish occur throughout the Aleutian, Oregonian and San Diego provinces and appear to have life history patterns adapted to the different temperature regimes, oceanographic processes and biotic interactions within each of these provinces accordingly (e.g., timing of parturition to coincide with variation in local retention processes; Love et al. 2002) (Gunderson & Vetter 2006). While speculative, it is possible that as the yelloweye rockfish expanded its geographic range throughout the northeastern Pacific, the advantages afforded by local retention, such as a greater likelihood of encountering familiar habitat, prey base and environmental conditions, and exposure to similar pathogen communities, selected for behavioral and life history attributes that limited dispersal away from natal habitats.

How does limited dispersal and adaptive divergence interact to structure marine populations?

Under conditions of restricted dispersal and gene flow, allele frequencies will drift apart and genetic differentiation may accumulate between more discrete locales (Wright 1931). Because gene flow is expected to reduce the establishment of favorable allelic combinations due to recombination with migrant genes, a reduction in gene flow may increase the potential for adaptive divergence to establish (Slatkin 1973, 1987; Lenormand 2002). Consequently, gene flow will be further restricted if local selection is strong enough to select against maladapted immigrants. Thus, a reduction in dispersal and gene flow may precipitate the establishment of locally adaptive allelic combinations, which further restrict gene flow by decreasing post-

settlement survival of immigrant individuals. Selection against migrant individuals has been observed between isolated populations of the same species (Nielsen et al. 2003).

The regional structure I observed with the neutral loci suggest a mesoscale model of dispersal (Gunderson & Vetter 2006), in which populations at a coast-wide scale are regionally self-recruiting, lack an IBD signature, and are expected to exhibit genetic breaks consistent with habitat or oceanographic barriers. I tested for genetic breaks between the outer coastal waters and the inside waters of the Georgia Basin and the Broughton Archipelago. Genetic clusters consistent with this division were identified in three out of four comparisons, that I was able to make, suggesting that dispersal between the outside and inside waters is restricted as first suggested from microsatellite DNA data by Yamanaka et al. (2006). Interestingly, neutral AFLP loci resolved finer scale population structure within the outside and inside divisions that were not consistent with my expectations. Neutral AFLP loci were expected to resolve the inside-outside division, but not further population subdivision within the two oceanographic divisions. It was hypothesized that outlier loci, those exhibiting substantial differentiation, would be necessary to resolve population subdivision on a finer scale within the inside and outside divisions.

The results from the outlier analyses mimic the population structure observed with the neutral loci, and suggest that the patterns of limited dispersal that allow neutral divergence also allow selective divergence, which in turn, may act to reinforce population structure primarily driven by neutral processes. The outlier datasets were comprised of 3-18 loci exhibiting a moderate amount of support for being under selection by BAYESCAN (i.e., posterior probability > 0.7). The purpose of including more outlier loci with a lower support for being influenced by selection was to evaluate the general effect of non-neutrality on population structure, and rely less on inferences regarding individual loci.

Interest in adaptive evolution within wild populations has resulted in researchers utilizing molecular methods to detect 'footprints of selection' (i.e. genetic polymorphism maintained by heterogenous environments), as demonstrating natural selection in the wild remains quite challenging (Nielsen et al. 2009a). It is not surprising, however, that there remain relatively few good examples of this in the literature

(Nielsen et al. 2009a). Adaptive traits are often quantitative and the effects of single genes can be quite small and difficult to detect (Naish & Hard 2008). Moreover, the prevalence of neutral markers, which are often used in limited numbers in population genetic studies, makes detecting the footprints of selection unlikely (Nielsen et al. 2009a).

Despite these limitations, however, studies detecting the footprints of selection do exist. Classic examples of candidate adaptive genes in marine fishes include lactate dehydrogenase B in killifish (*Fundulus heteroclitus*, Powers & Schulte 1998), the membrane protein coding *pantophysin* (Pan I) gene, which is well studied in Atlantic cod (*Gadus morhua*, Nielsen et al. 2009b), and, most recently, ectodysplasin in threespine stickleback (*Gasterosteus aculeatus*, Barrett et al. 2008). A number of studies have demonstrated that different Pan I alleles are more highly differentiated than neutral markers, and may be subject to diversifying selection (Pogson et al. 1995; Fevolden & Pogson 1997; Pogson 2001; Canino et al. 2005). Temperature and salinity gradients have been suggested to be responsible for the elevated levels of differentiation (Case et al. 2005), and different Pan I genotypes have been correlated with different fitness correlates, such as: growth, body condition and behavior (Case et al. 2006; Jónsdóttir et al. 2008; Pampoulie et al. 2008).

Genome scan approaches, however, have only been used to identify signatures of selection in a few marine fishes. A Web of Science search for 'genome scan' and 'marine fish' found only 13 records. Moen et al. (2008) performed a genome scan in Atlantic cod and found 28 out of 318 single nucleotide polymorphisms (SNPs) had elevated levels of differentiation consistent with directional selection. Also in Atlantic cod, Nielsen et al. (2009b) identified 8 out of 98 gene-associated SNPs in a genome scan with high support for being under directional selection. A genome scan with AFLP loci was conducted by McCusker & Bentzen (2010) in the Atlantic wolfish (*Anarhichas lupus*), who detected two putatively selected loci. Colbeck et al. (2011) used AFLPs for a genome scan in the capelin (*Mallotus villosus*) and found four outlier loci that exhibited divergence at adjacent sampling sites, and at scales within the range of adult movement and/or larval dispersal, putatively due to the effects of divergent selection.

Evidence of self-recruitment in marine fishes (Jones et al. 1999; Cowen 2006) necessitated the reexamination of adaptation in the marine environment and as such, examples of divergent selection and adaptive genetic differentiation are accumulating (Hauser & Carvalho 2008). Genome scan studies represent a starting place for identifying putatively adaptive loci in marine species. I identified a lower but relatively comparable proportion of outlier loci in yelloweye rockfish compared to the two previously described AFLP genome scan studies. The next step is to isolate and sequence these outlier AFLP loci and see if they match any known functional gene sequences housed in accessible databases (e.g. National Center for Biotechnology Information, www.ncbi.nlh.nih.gov). Determining if an AFLP locus represents a gene with a known function may help guide future studies by prioritizing relevant environmental gradients over which there is expected genetic variation at the candidate gene.

AFLPs versus microsatellites: why more structure with AFLPs?

The results of this study expand on those of Yamanaka et al. (2006). My data represent an independent resolution of the inside-outside division originally reported by Yamanaka and colleagues, but also found greater substructure within that major division. The varying degree of population structure identified in the two studies demonstrates that marker selection can have an important influence on the pattern of genetic differentiation and subsequent implications. Genetic differentiation results from the cumulative effect of different evolutionary processes. Some of these will have variable effects on different molecular markers (e.g., mutation, selection), while others (e.g., migration) are expected to affect the whole genome, and thus affect different marker types equally (Mariette 2001). Microsatellites are often highly polymorphic, exhibiting high within-population heterozygosity, which can precipitate underestimates of genetic differentiation (Hedrick 1999; Balloux et al. 2000). In addition, high mutation rates, estimated to be 10^{-4}-10^{-2} (Weber & Wong 1993) and resulting incidences of homoplasy within microsatellites (Estoup et al. 1995, 2002) may further lead to underestimation of differentiation (Mariette 2001). The use of dominant markers, however, may also bias estimates of genetic differentiation. Jenczewski et al. (1999) showed that allele frequencies based on random amplified polymorphic DNA (RAPD) phenotypes produced upwardly

biased F_{ST} estimates. Krutovskii (1999) demonstrated that RAPDs correctly estimate population differentiation only when sample sizes are greater than 30 individuals. It is clear that different marker types are not directly comparable, as they exhibit variable sensitivities to different evolutionary processes. Thus, when two or more genetic marker datasets are available, it is similarity in the direction of qualitative patterns, not absolute values that should be evaluated for congruence (Gaudeul et al. 2004).

The AFLP method has largely been embraced by the plant, fungi and bacteria communities, with the majority of studies focusing on crop and other economically important species and their pathogens (Bensch & Åkesson 2005). In contrast, fewer population genetic studies in animals have utilized AFLPs, again mostly limited to domesticated and model species. As a consequence, there are a limited number of studies directly comparing measures of genetic diversity and differentiation derived from AFLPs and microsatellites in wild populations (Gaudeul et al. 2004). Paupy et al. (2004) investigated population structure of the yellow fever mosquito (*Aedes aegypt*) with three marker types: AFLPs, microsatellites and isoenzyme markers. All three marker types reveal the same population structure. F_{ST} values derived from AFLPs, however, were 3-5x higher than that from microsatellites and isoenzyme markers. In the Atlantic wolfish, McCusker & Bentzen (2010) found weak but significant population structure with both microsatellites and AFLPs, although F_{ST} estimates were slightly higher with AFLPs. In studies of genetic diversity and differentiation of oaks, *Quercus* spp. (Mariette et al. 2002) and *Athyrium distentifolium* (Woodhead et al. 2005), G_{ST} and pairwise F_{ST} values were higher for AFLPs than microsatellites. Gaudeul et al. (2004) found higher genetic differentiation with AFLPs (θ=0.40) than with microsatellites (θ=0.23). Mariette et al. (2001) used AFLPs and microsatellites to investigate genetic diversity within and between populations of Maritime pine (*Pinus pinaster).* Absolute values of differentiation obtained from both marker types were congruent (microsatellites: G_{ST}=0.111; AFLPs: G_{ST}=0.102), as were estimates of total genetic diversity.

It is apparent from the limited number of studies available that microsatellites and AFLPs are generally in congruence and reveal similar qualitative patterns of genetic differentiation and population

structure, despite a discrepancy between absolute values. These studies, however, are biased towards plants, and do not necessarily represent marker behavior and population structure patterns in marine fishes, whose traditionally high effective population sizes and potential for long distance dispersal sets them apart from most other systems (Nielsen et al. 2009a). In fact, a Web of Science search with the keywords: "aflp", "microsatellite", and "marine fish", only produced ten studies, three of which are genetic linkage maps, and only two of which discuss a comparison of population structure derived from both AFLPs and microsatellites (and even both of those are about the same study organism by the same authors: McCusker & Bentzen 2010, 2011). The extremely limited number of studies investigating the effect of marker type on population genetic diversity and differentiation in marine systems is striking. Waples (1998) pointed out that the low F_{ST} values in marine systems reduces the signal to noise ratio, and error associated with that estimation takes on greater significance. Moreover, the use of additional genetic markers and higher sample sizes will increase statistical power, and may lead to statistically significant results, but distilling biological significance from statistical significance is another matter. It is not surprising then that population genetic studies of marine organisms often have a difficult time interpreting weak yet statistically significant differentiation (Hauser & Carvalho 2008). Thus, it appears that species predisposed to exhibiting weak population structure (e.g., marine fishes and invertebrates that have an extended PLD), are more susceptible to variable interpretations of population structure based on marker type than species distributed within more highly structured populations.

Implications for management and conservation

In BC, yelloweye rockfish are managed as two distinct stocks that align with the inside/outside genetic break identified by Yamanaka et al. (2006). This genetic break also provided the impetus for the Committee on the Status of Wildlife in Canada (COSEWIC) to recognize two Designatable Units (DUs) of yelloweye rockfish, both of which are currently listed as 'Special Concern' (COSEWIC 2008). Yelloweye rockfish, along with many other rockfish species are in a state of decline across the northeastern Pacific Ocean (Parker et al. 2000). To address conservation concerns, DFO established an inshore rockfish

conservation strategy composed of four measures: (i) improve catch monitoring programs; (ii) reduce harvest levels to below the estimates for natural mortality; (iii) improve population assessment and monitoring programs; (iv) create a network of marine protected areas to increase larval production (Groundfish Integrated Fisheries Management Plan for 2011-2013). This strategy was implemented in 2002, and there has been a reduction in total allowable catch (TAC) by 50% and 75%, for the outside and inside stocks, respectively, and there are now 164 marine protected areas (termed Rockfish Conservation Areas, RCAs) encompassing approximately 30% and 20% of inside and outside suitable rockfish habitat (Yamanaka & Logan 2010).

The results of this study have implications for both the stock delineation of yelloweye rockfish and the spatial considerations of the RCAs. I largely confirmed the inside-outside genetic break first elucidated by Yamanaka and colleagues, further supporting the isolation of the Georgia Basin from the outer coastal waters. In addition, my results suggest that effective larval dispersal distances are smaller than previously thought, and population substructure exists on a finer spatial scale, which is most relevant for the delineation of outside stocks. Further work is needed, however, to clearly resolve regional population boundaries that likely experience a higher degree of self-recruitment, the implications of which also bear on the layout of the RCA system. For example, when designing and siting marine protected areas, the area size is often constrained, either for economic, political or logistical reasons (Roberts et al. 2003). It then becomes necessary to implement a network of protected areas, which function in concert to help maintain recruitment within the protected areas or even to promote population connectivity within the greater region. The RCAs were not implemented as a functioning network, and the degree to which one RCA helps promote larval replenishment in another RCA is not well understood. If dispersal is regionally limited, then RCAs in one region may not be a significant source of larval recruits for RCAs located in different regions, across potential dispersal barriers. Furthermore, the population structure I observed with the outlier analyses mimics the patterns detected with the neutral loci, and suggest that selection may be operating within the same regional scale as neutral processes, and may act to reinforce population structure

that is primarily driven by limited dispersal. If that is the case, then postsettlement survival is expected to be lowest amongst individuals recruiting from more environmentally disparate locales. The placement of the RCAs, then, should not only reflect dispersal distances, but encompass environmental heterogeneity, as adaptive genetic variation and phenotype matching might be important components in postsettlement survivorship (Marshall et al. 2010).

Conclusion

My results suggest that high alongshore water flow does not necessarily translate into high alongshore larval transport, and larvae are probably able to take advantage of oceanographic retention mechanisms, restricting gene flow along the coast. Most likely, limited dispersal is primarily driving population structure in yelloweye rockfish, which may be reinforced by selection acting within this regional scale. The detection of the plate bias during the initial data analysis phase of my study was a significant obstacle, the effects of which had a role in driving differences among localities in allele frequencies. Despite the potential pervasiveness of this bias, it has received almost no attention in the literature, and studies have yet to report on testing for this bias or steps used to minimize its effects. By treating my PCR plates as independent "projects", I was able to independently replicate several spatial comparisons, bolstering support for the patterns of population structure that were detected (e.g. inside vs. outside; isolation of Bowie Seamount). The BCMEC framework was used as a starting point for partitioning environmental variation and looking for coincident genetic clusters. Twelve out of 24 comparisons yielded an association between Ecosections and unique genetic clusters, suggesting that putatively adaptive genetic structure may reflect regional environmental differences. In systems where environmental may be logistically difficult to obtain, using a broad classification system, like the BCMEC employed in my study is a useful starting point for studies of adaptive genetic variation, which can help future studies target genetic variation across finer spatial scales to better resolve the mechanisms promoting adaptive genetic divergence.

TABLES

Table 1: The hierarchical British Columbia Marine Ecosystem Classification (table modified from Zacharias et al. 1997).

Ecozone	Ecoprovinces	Ecoregions	Ecosections
Pacific	North Pacific	Subarctic Pacific	Subarctic Pacific
		Transitional pacific	Transitional pacific
	Pacific Shelf & mountains	Outer Pacific Marine Shelf	Continental Slope
		Inner Pacific Marine Shelf	Vancouver Island Shelf
			Queen Charlotte Sound
			Dixon Entrance
			Hecate Strait
			North Coast Fjords
			Queen Charlotte Strait
			Johnstone Strait
	Georgia-Puget Sound Basin	Georgia Basin	Juan de Fuca Strait
			Strait of Georgia

Table 2: Ecounit descriptions. Ecounits represent major oceanographic and bathymetric variables that have significant influence on biotic communities. These units were created to evaluate large scale oceanographic and environmental differences that define each Ecosection. (Table modified from Zacharias et al. 1997.)

Ecounit	Class	Description
Wave Exposure	High	Fetch >500 km. Ocean swell environment.
	Moderate	Fetch 50-500 km. Some swell areas, open sounds and straits
	Low	Fetch <50 km. Protected areas; some small sounds and straits
Depth	Photic	0-20 m
	Shallow	20-200 m
	Moderate	200-1,000 m
	Abyssal	>1,000 m
Relief	High	Abundant cover and diversity of habitats
	Low	Smooth or gently undulating bottom
Currents	High	Maximum currents >3 knots (1.54 m/s)
	Low	Maximum currents <3 knots (1.54 m/s)
Substrate	Hard	Bedrock, boulders, cobble, and some sand/gravel areas
	Sand	Sand, gravel/sand, and some muddy areas
	Mud	Mud and sandy mud
	Unknown	Not sampled

Table 3: Sample locations and respective sample sizes.

Sample Location (code)	N	Sample Location (code)	N
SE Alaska (AK)	37	Esperanza (ES)	46
Bowie Seamount (BS)	48	lower Bute Inlet (LB)	48
Tasu (TA)	48	Texada Island (TI)	48
Cape St. James (CJ)	46	Middlenatch Island (MI)	24
Calvert Island (CI)	46	Gabriola Island (GA)	17
Gordon Channel (GC)	34	Oregon (OR)	48
Top Knot (TK)	48	Washington (WA)	41
Triangle (TR)	47		

Table 4: The primer sequences of all primer combinations are shown below. The three bases at the end of the EcoR1 and Mse 1 primers (E- and M-) correspond to the three selective nucleotides used.

Primer Pair	E-**NNN** primer	M-**NNN** primer
1	GACTGCGTACCAATTC**ACC**	GATGAGTCCTGAGTAA**CAA**
2	GACTGCGTACCAATTC**ACG**	
3	GACTGCGTACCAATTC**AGC**	
4	GACTGCGTACCAATTC**ATA**	
5	GACTGCGTACCAATTC**ACC**	GATGAGTCCTGAGTAA**CAC**
6	GACTGCGTACCAATTC**ACG**	
7	GACTGCGTACCAATTC**AGC**	
8	GACTGCGTACCAATTC**ATA**	
9	GACTGCGTACCAATTC**ACC**	GATGAGTCCTGAGTAA**CAT**
10	GACTGCGTACCAATTC**ACG**	
11	GACTGCGTACCAATTC**AGC**	
12	GACTGCGTACCAATTC**ATA**	
13	GACTGCGTACCAATTC**ACC**	GATGAGTCCTGAGTAA**CAG**
14	GACTGCGTACCAATTC**ACG**	
15	GACTGCGTACCAATTC**AGC**	
16	GACTGCGTACCAATTC**ATA**	

Table 5: Sample locations and respective sample sizes ordered by plate. These sample sizes do not include individuals that contained more than 15% missing data, and were dropped from analyses. The number of loci and average fragment presence frequency for each plate are also reported.

Plate	Sample Locations	N	No. of loci	Avg. fragment presence freq.
Plate 01	BS	24	1093	0.45
	ES	24		
	TA	24		
	TI	24		
Plate 02	CI	22	1274	0.36
	GA	17		
	GC	10		
	WA	17		
Plate 03	LB	24	1510	0.31
	MI	24		
	OR	24		
	TK	24		
Plate 04	AK	37	1580	0.41
	CJ	46		
Plate 05	BS	24	1313	0.44
	TA	24		
	TK	24		
	TR	24		
Plate 06	CI	24	966	0.40
	LB	24		
	OR	24		
	TI	24		
Plate 07	ES	22	1268	0.37
	GC	24		
	TR	23		
	WA	24		
			Mean: 1286	Mean: 0.39

Table 6: Primer pairs and corresponding scoring criteria used in AFLPSCORE. The asterisks and bolded rows mark the primer pairs retained for downstream analyses.

Primer Pair	Threshold class	Locus selection threshold	Phenotype-calling threshold	No. of replicates	Error Rate	No. of initial bins
1	absolute	100	100	124	14	479
2	absolute	100	300	161	11.8	287
3	absolute	100	300	171	15.8	318
4*	**absolute**	**100**	**100**	**91**	**6.4**	**486**
5*	**absolute**	**150**	**150**	**159**	**9.7**	**492**
6	absolute	150	250	165	13.1	319
7	absolute	150	250	173	13.4	380
8*	**absolute**	**150**	**150**	**179**	**8.3**	**451**
9*	**absolute**	**100**	**100**	**181**	**10**	**485**
10*	**absolute**	**100**	**250**	**171**	**8.4**	**273**
11	absolute	150	300	171	11.4	289
12*	**absolute**	**100**	**100**	**173**	**8.5**	**432**
13*	**absolute**	**150**	**150**	**176**	**7.4**	**431**
14*	**absolute**	**150**	**150**	**167**	**7.6**	**367**
15*	**absolute**	**150**	**300**	**171**	**9.1**	**320**
16*	**absolute**	**100**	**100**	**174**	**7.7**	**472**
					*Mean: 8.3	*Total: 4209

Table 7: BAYESCAN results and corresponding dataset information. Asterisks indicate which plates share common outlier loci (one outlier was identified on both Plate 01 and Plate 05, and four outliers were identified on Plate 6 and Plate 07).

Plate	No. of outlier loci (% of total no. of loci)	Minimum posterior probability	No. of loci common between plates	No. of loci in outlier dataset	No. of loci in neutral dataset
Plate 01	6 (0.55%)	0.92	1*	18	1025
Plate 02	2 (0.16%)	0.96	0	16	1197
Plate 03	1 (0.0006%)	0.94	0	5	1455
Plate 04	0	---	---	---	1576
Plate 05	1 (0.0008%)	0.93	1*	3	1261
Plate 06	7 (0.62%)	0.89	4**	12	928
Plate 07	10 (0.79%)	0.94	4**	10	1252

Table 8: Global F_{ST} and heterozygosity values determined in AFLP-SURV for each dataset within the seven plates. Ht: total gene diversity; Hw: mean gene diversity within populations (Nei's Hs); Hb: average gene diversity among populations in excess of that observed within populations; F_{ST}: the proportion of total gene diversity that occurs among as opposed to within populations.

Plate	Dataset	Ht	Hw	Hb	F_{ST}
Plate 01	neutral	0.2605	0.2274	0.0331	0.1274
	outlier	0.5562	0.2204	0.3357	0.6007
Plate 02	neutral	0.2605	0.2274	0.0331	0.1274
	outlier	0.5655	0.1861	0.3794	0.6634
Plate 03	neutral	0.2354	0.2248	0.0106	0.0451
	outlier	0.5019	0.2783	0.2236	0.4555
Plate 04	neutral	0.2942	0.2569	0.0373	0.1254
	---	---	---	---	---
Plate 05	neutral	0.2827	0.2493	0.0334	0.1190
	outlier	0.5609	0.2219	0.3389	0.5946
Plate 06	neutral	0.2670	0.2478	0.0192	0.0718
	outlier	0.5472	0.2659	0.2813	0.5235
Plate 07	neutral	0.2692	0.2464	0.0227	0.0846
	outlier	0.5099	0.2436	0.2663	0.5079
Mean	neutral	0.2671	0.2400	0.0271	0.1001
	outlier	0.5403	0.2360	0.3042	0.5576

Table 9: Results of the Mantel tests for isolation-by-distance.

Plate	P-value	
	Neutral dataset	Outlier dataset
Plate 01	0.834	0.660
Plate 02	0.828	0.179
Plate 03	0.830	0.494
Plate 04	---	---
Plate 05	0.498	0.505
Plate 06	0.836	0.338
Plate 07	0.165	0.163
Mean	0.666	0.390

Table 10: STRUCTURE assignment scores. For each *K*-value, the proportion of individuals whose genotype assigns to a single cluster at the minimum assignment threshold is reported. Bolded values highlight the overall better assignment observed in the outlier datasets compared to the neutral datasets.

Plate	Dataset	Proportion of individuals assigning to a single cluster (for K=2)		Proportion of individuals assigning to a single cluster (for K=3)	
		Assignment threshold 0.9	Assignment threshold 0.95	Assignment threshold 0.9	Assignment threshold 0.95
Plate 01	neutral	0.58	0.51	0.49	0.43
	outlier	**0.65**	**0.57**	**0.69**	**0.64**
Plate 02	neutral	0.71	0.67	0.55	0.50
	outlier	0.71	**0.68**	**0.82**	**0.76**
Plate 03	neutral	0.69	0.60	0.39	0.28
	outlier	**0.72**	**0.65**	**0.63**	**0.32**
Plate 04	neutral	0.62	0.59	0.43	0.36
	---	---	---	---	---
Plate 05	neutral	0.64	0.0	**0.46**	**0.39**
	outlier	**0.68**	**0.03**	0.0	0.0
Plate 06	neutral	**0.67**	**0.59**	0.39	0.31
	outlier	0.60	0.58	**0.71**	**0.58**
Plate 07	neutral	0.70	0.68	0.49	0.42
	outlier	**0.91**	**0.87**	**0.80**	**0.63**

Table 11: Expectations and results for the genetic clustering patterns with the neutral and outlier loci. The neutral loci were expected to resolve the inside/outside division identified by Yamanaka et al. (2006). The outlier loci were expected to discern further population subdivision within the inside and outside divisions, as well as conform to the major environmental divisions encompassed within the BCMEC Ecosection divisions.

	Neutral loci		**Outlier loci**
Plate (no. of sample locations)	Inside vs. outside split tested? / Observed?	No. of outside sites / No. outside sites forming distinct outside cluster / No. of outside locations containing individuals which largely assign to a single cluster	No. of Ecosections / No. of clusters / No. of clusters corresponding to one Ecosection
Plate 01 (4)	Y / Y	3 / 0 / 1	4 / 3 / 2
Plate 02 (4)	Y / N	3 / 0 / 2	4 / 3 / 2
Plate 03 (4)	Y / Y	2 / 0 / 0	3 / 3 / 1
Plate 04 (2)	N / ---	2 / 0 / 2	2 / 2 / 2
Plate 05 (4)	N / ---	4 / 0 / 1	3 / 3 / 1
Plate 06 (4)	Y / Y	2 / 0 / 1	4 / 3 / 3
Plate 07 (4)	N / ---	4 / 0 / 2	3 / 3 / 1

FIGURES

Figure 1: Map of prevailing ocean currents off of British Columbia coast (Figure reproduced from Jamieson & Phillips 1988).

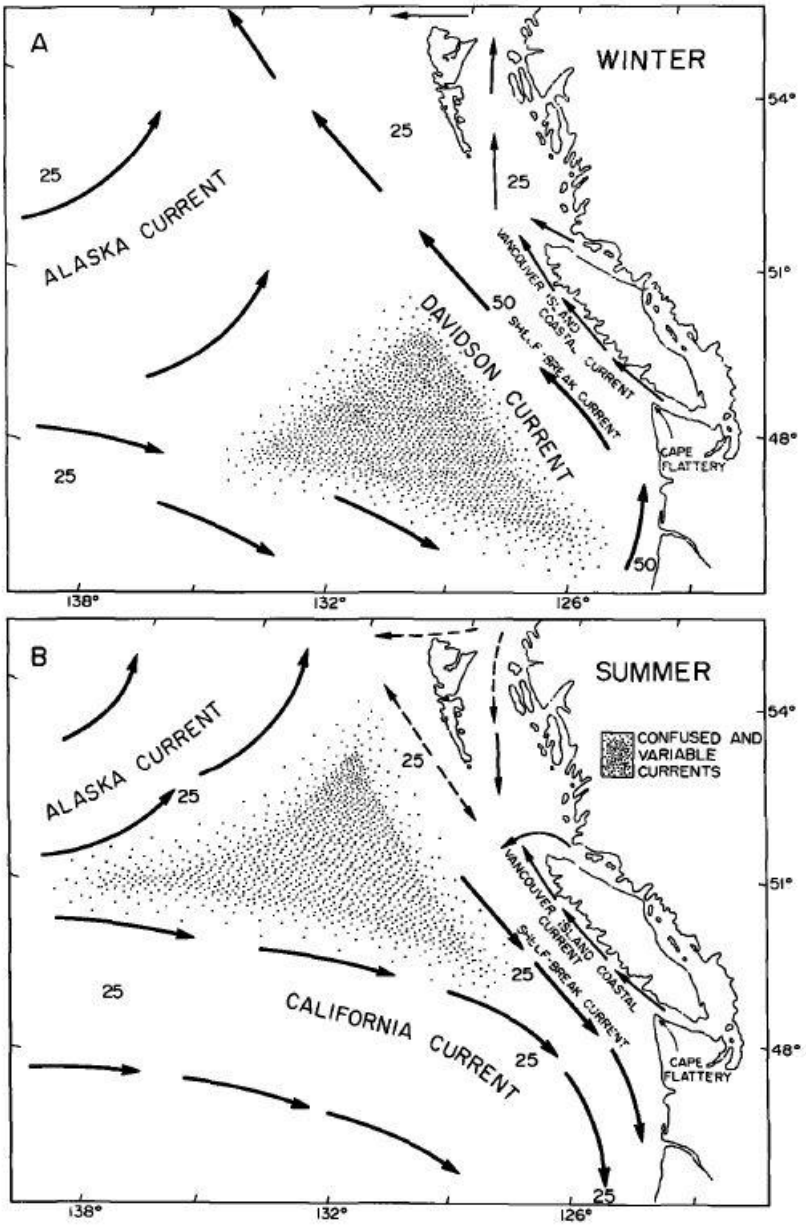


Figure 2: Marine Ecosections delineated in the British Columbia Marine Ecosystem Classification (BCMEC). (Figure modified from Zacharias et al. 1997) The Alaska (AK) and Oregon (OR) sampling locations were designated as their own "Ecosection", as these locations are not part of the BCMEC system.

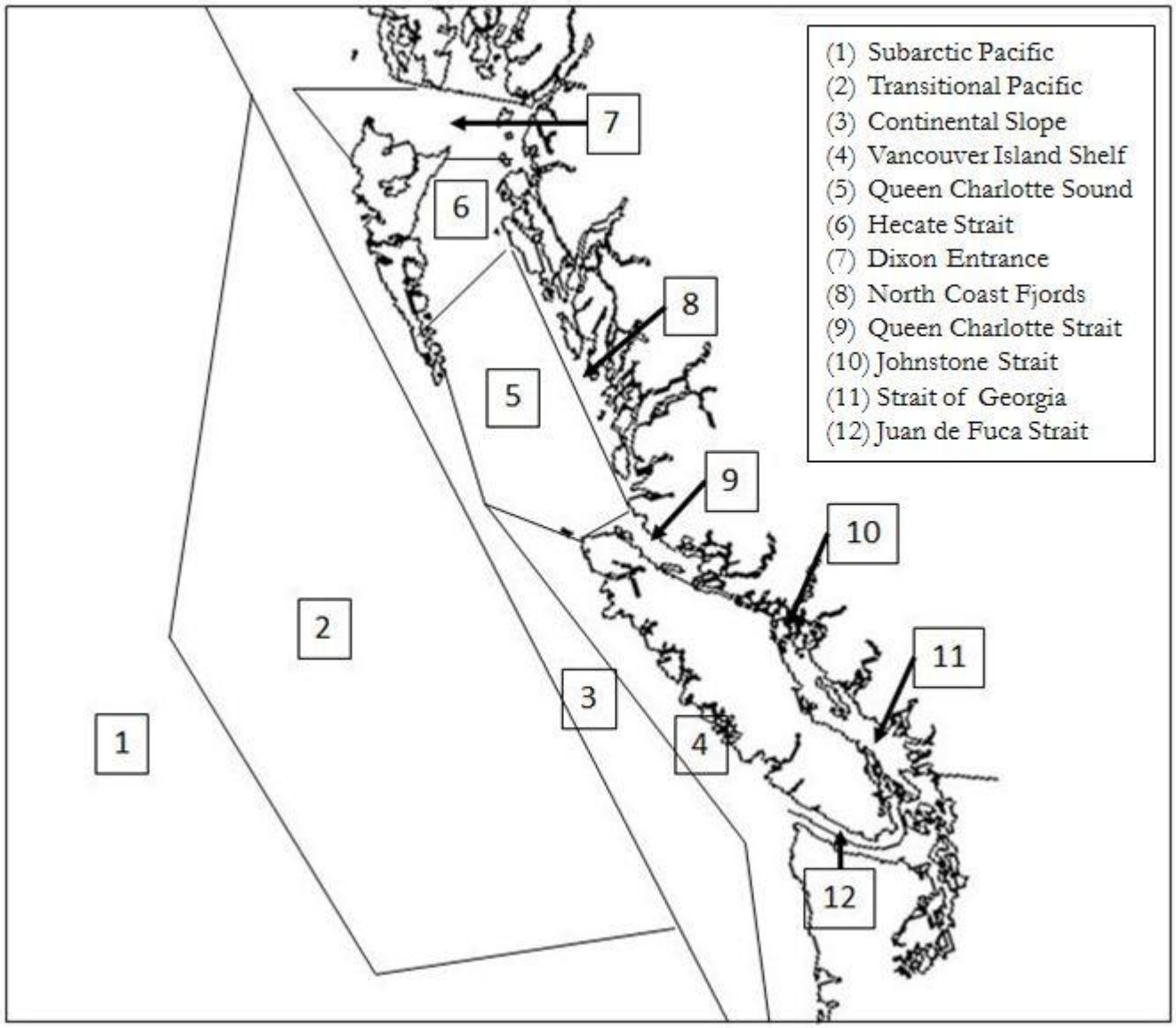
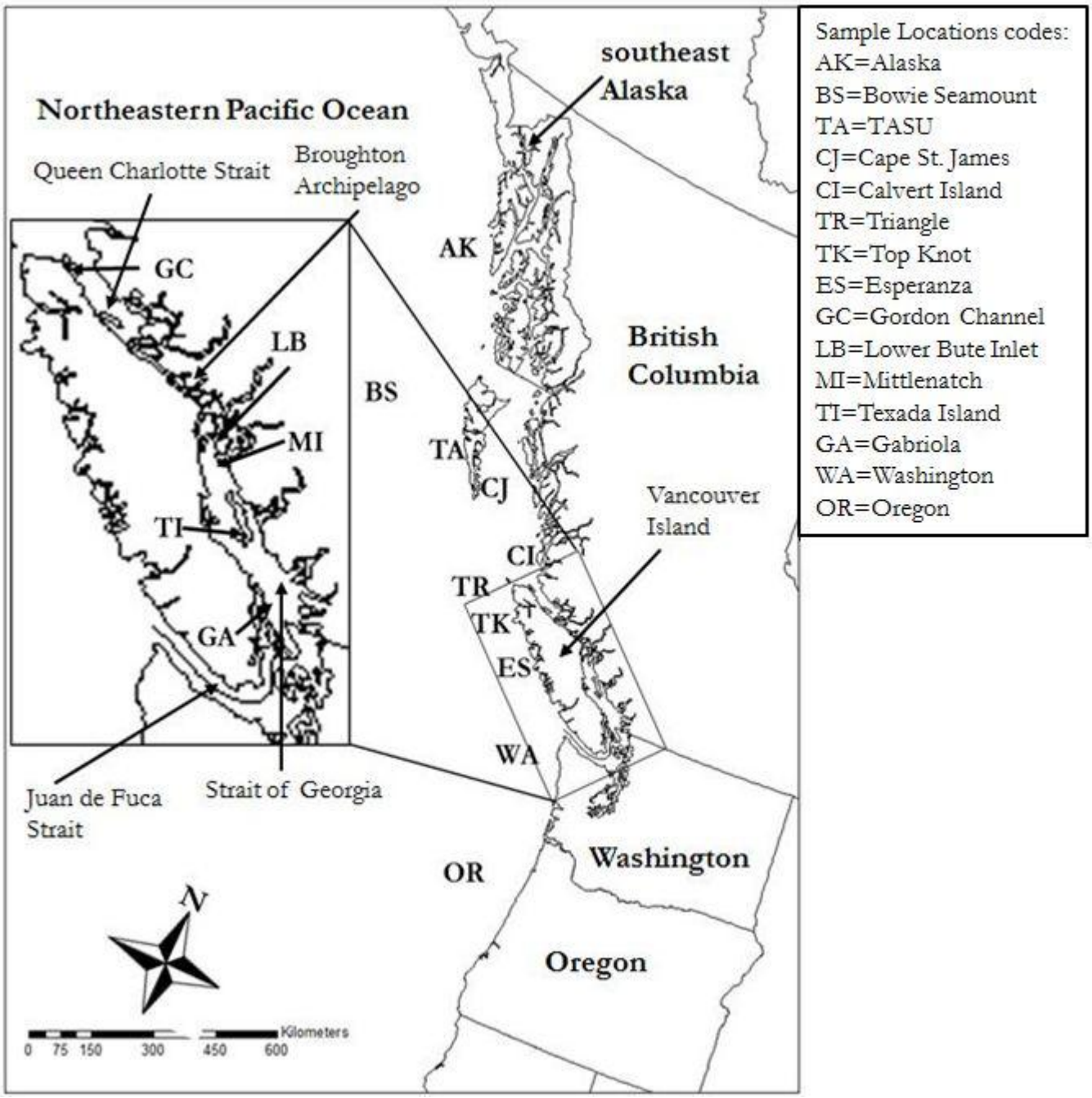
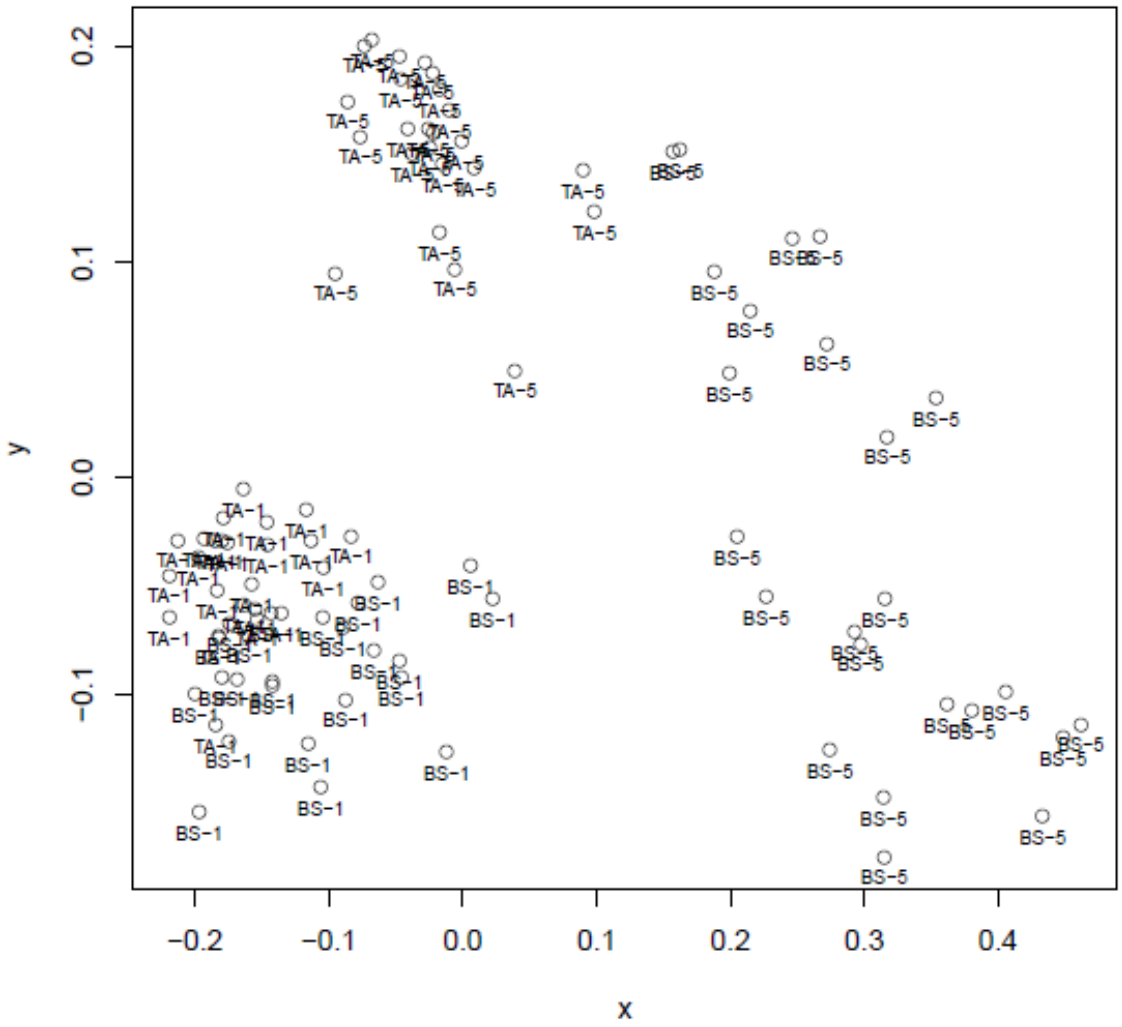


Figure 3: Map of yelloweye rockfish sampling locations.

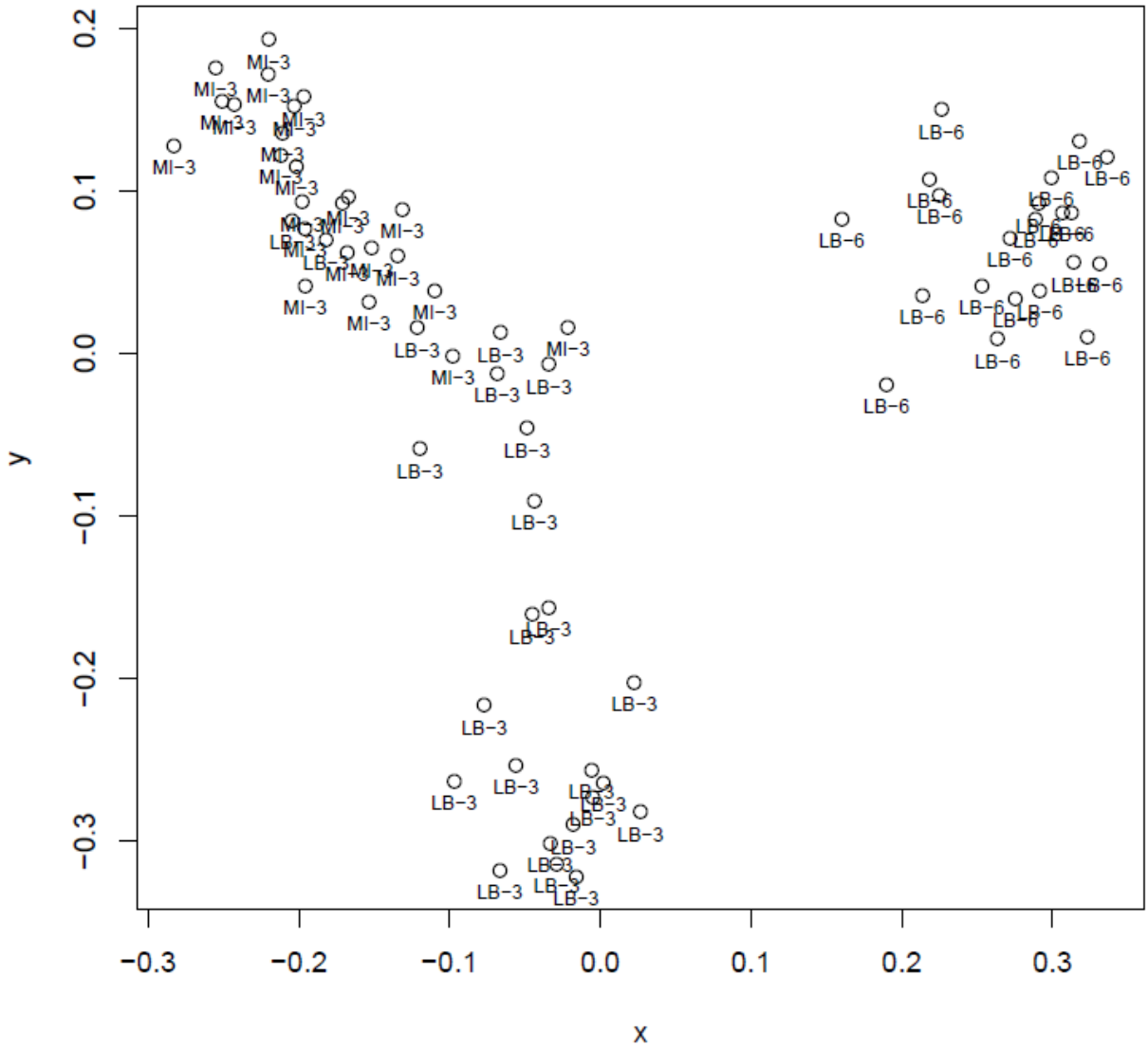


Figures 4a-d: Examples of multidimensional scaling analyses showing genetic grouping by plate. Location code is followed by plate of origin.

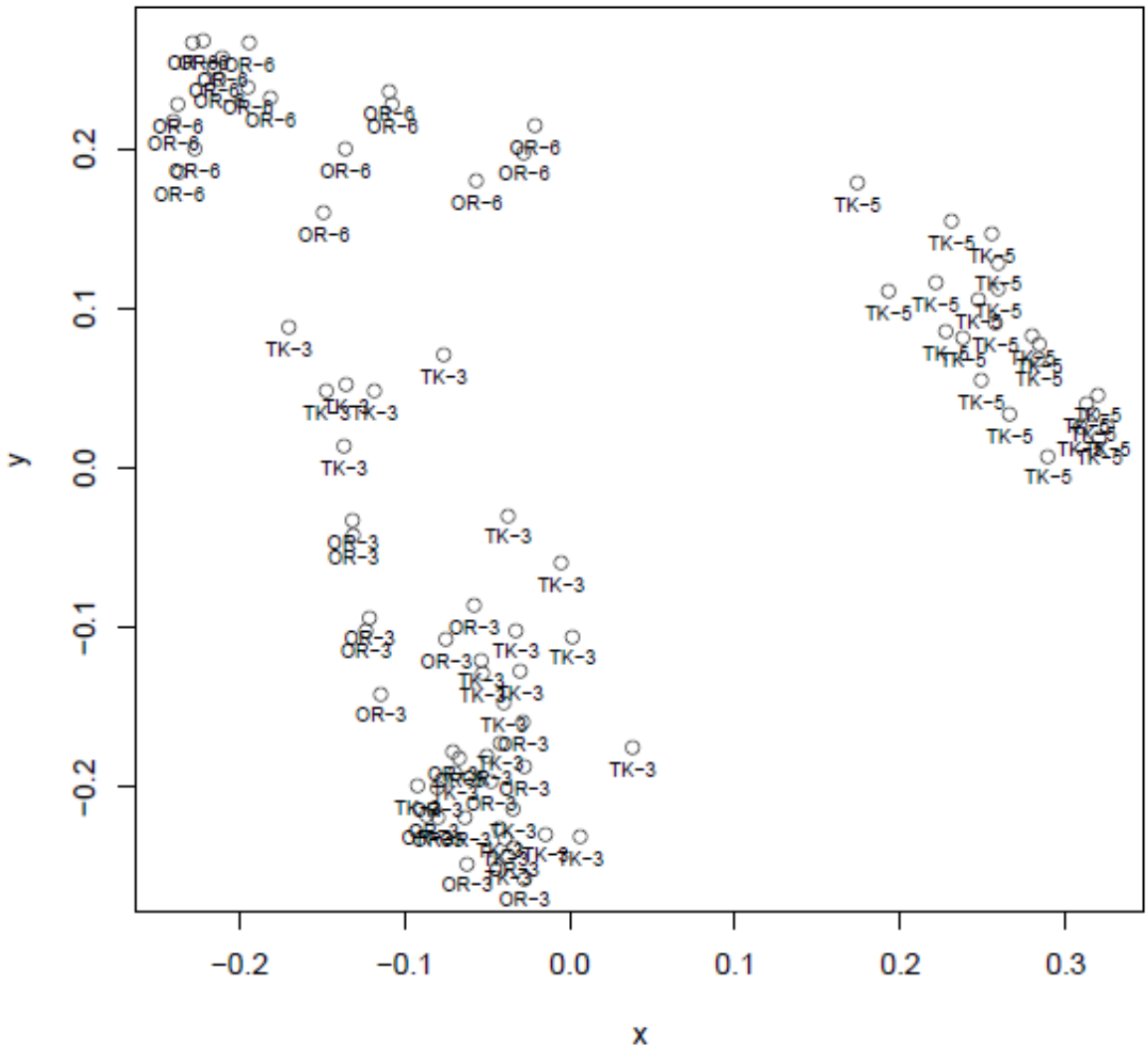
4a. BS-1 = Bowie Seamount from Plate 01, BS-5 = Bowie Seamount from Plate 05, TA-1 =Tasu from Plate 01, TA-5=Tasu from Plate 05. Both BS and TA are 'outside' locations.



4b. LB-3 = Lower Bute Inlet from Plate 03, LB-6= Lower Bute Inlet from Plate 06, MI-3 – Mittlenatch from Plate 03. Both LB and MI are 'inside' locations.



4c. OR-3 = Oregon from Plate 03, OR-6 = Oregon from Plate 06, TK-3 = Top Knot from Plate 03, TK-5 = Top Knot from Plate 05. OR and TK are 'outside' locations.



4d. ES-1 = Esperanza from Plate 01, ES-8 = Esperanza from Plate 08, TI-1 = Texada Island from Plate 01, TI-6 = Texada Island from Plate 06. ES is an 'outside' location, TI is an 'inside' location.

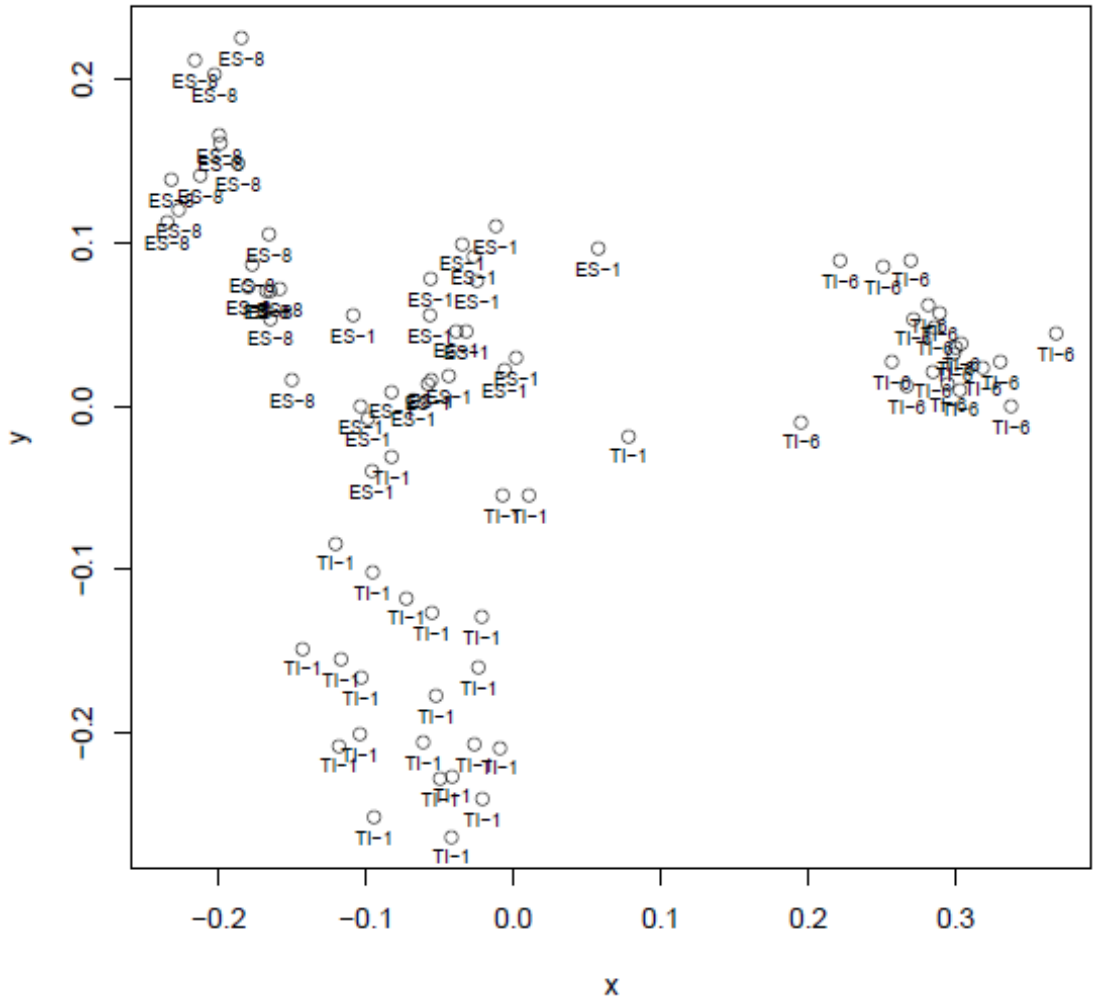


Figure 5: Proportion of initial loci (N=4209) that have either a presence or absence score. There is considerable variation in the amount of relative missing data between plates (especially between Plate 04 and Plate 06). The plate number is located above the corresponding peak.

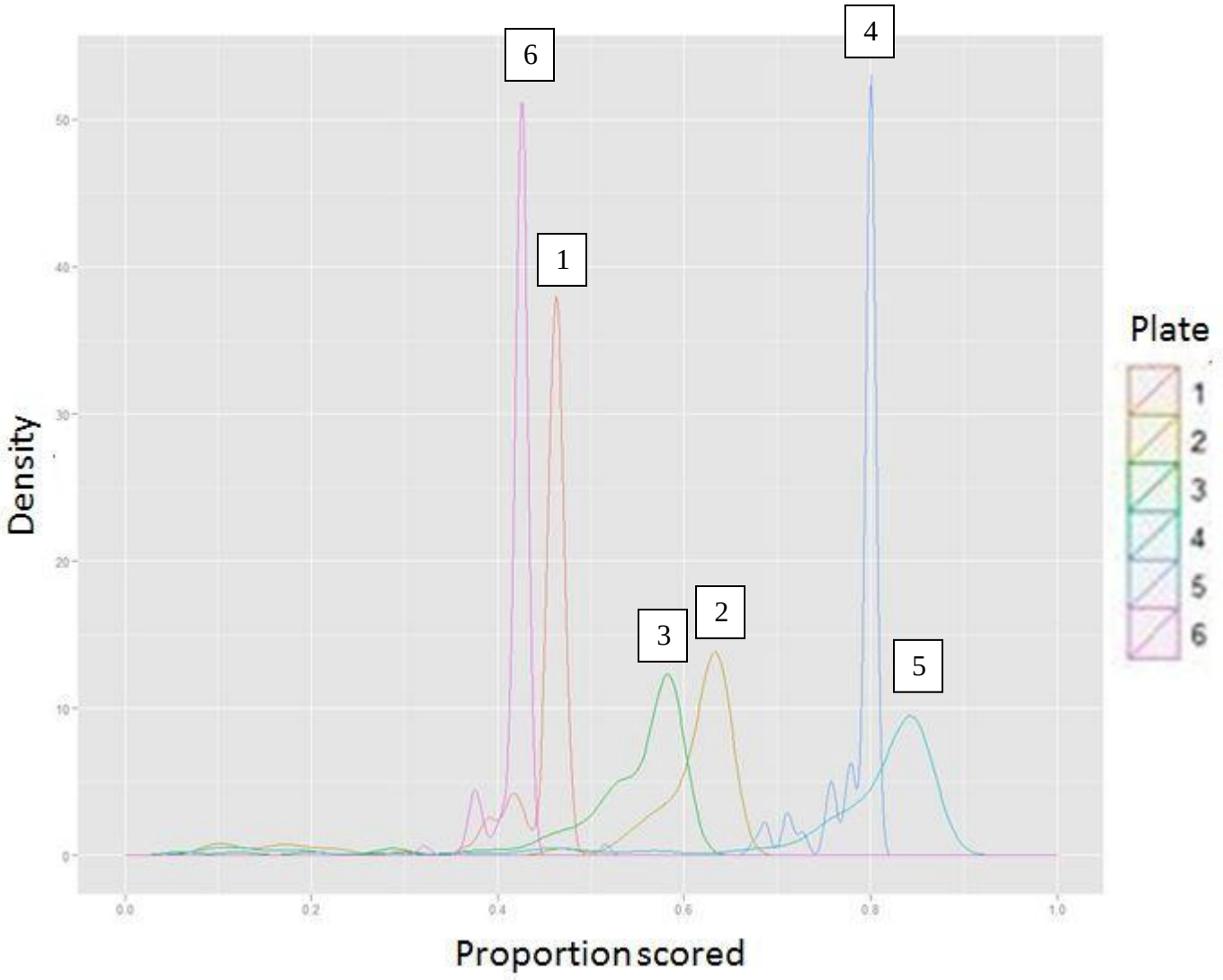


Figure 6: Proportion of fragment presences ("1" scores) shown by plate. On the extremes, Plate 03 has an average fragment presence frequency approximately 10% lower than Plate 01. The plate number is located above the corresponding peak.

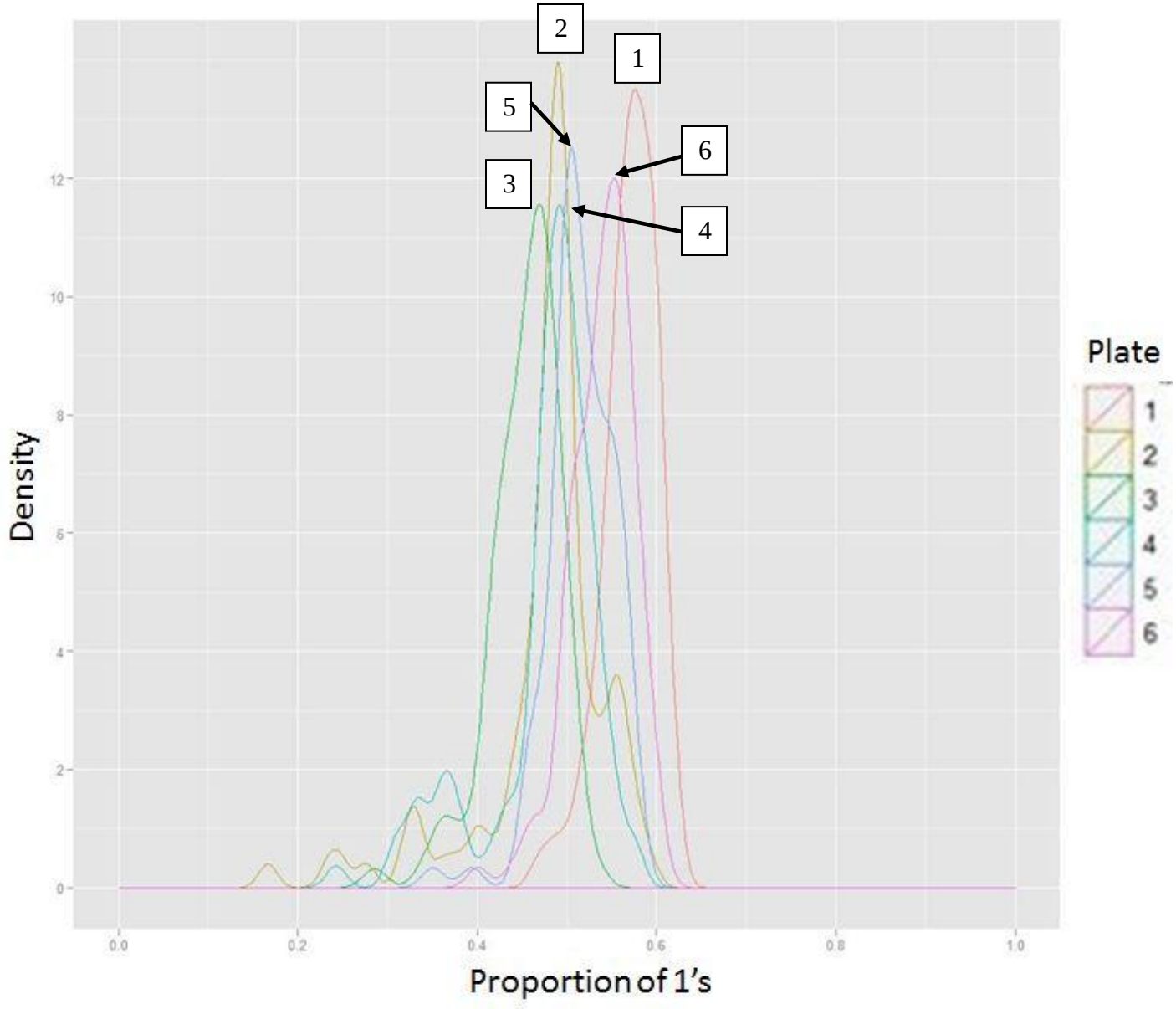
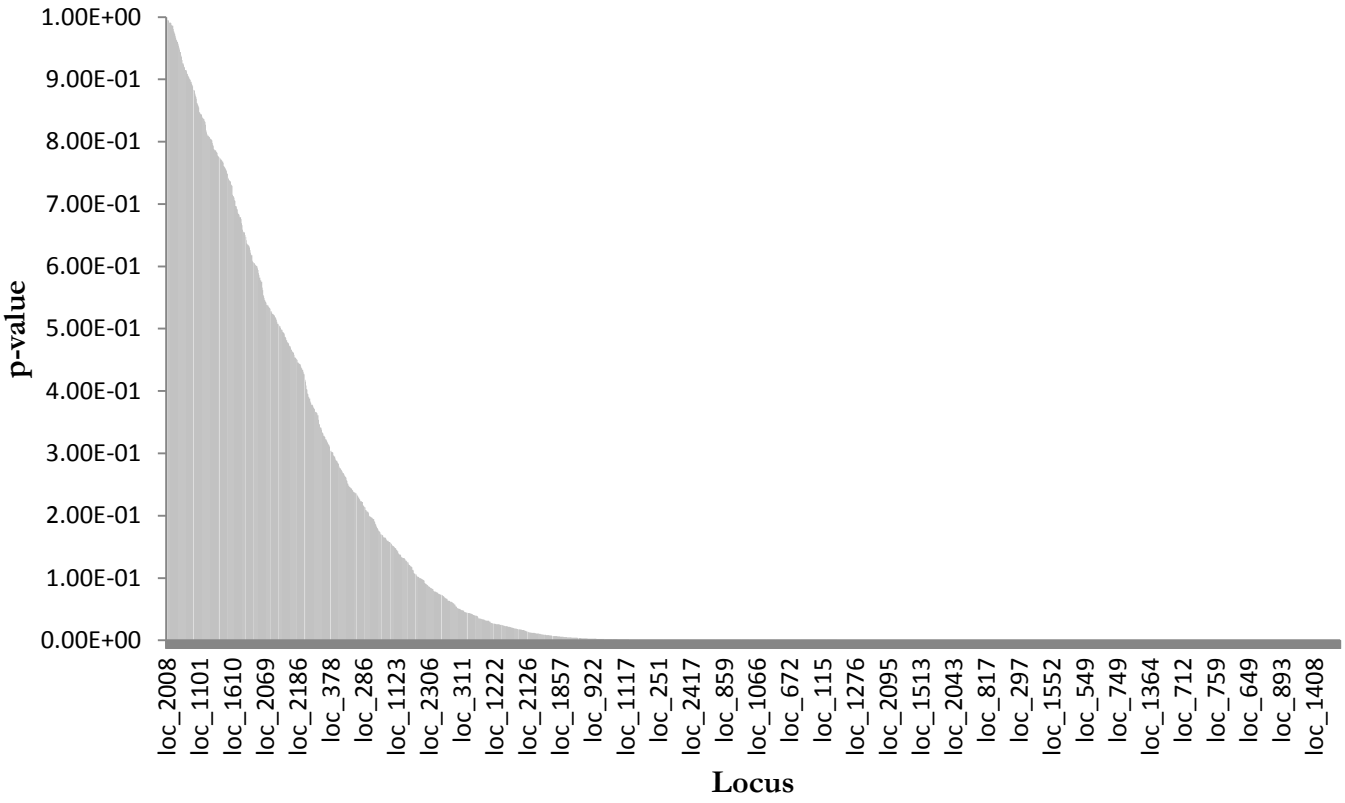


Figure 7: Distribution of p-values after the generalized linear model check for "plate" as a significant predictor of genotype frequency in yelloweye rockfish. Loci are on the x-axis, and corresponding p-values are on the y-axis. (The majority of loci are not listed on the x-axis).



REFERENCES

Allen, R.M. and Marshall, D.J. (2010) The larval legacy: cascading effects of recruit phenotype on post-recruitment interactions. *Oikos*. **119**, 1977-1983.

Allendorf, F. W. and Phelps, S. R. (1981) Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences.* **38**, 1507–1514.

Arrigo, N., Tuszynski, J.W., Ehrich, D., Gerdes, T. and Alvarez N. (2009) Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics.* **10**, doi:10.1186/1471-2105-10-33.

Balloux, F., Brünner, H., Lugon-Moulin, N., Hausser, J., and Goudet, J. (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution* **54**, 1414-1422.

Barrett, R.D.H, Rogers, S.M. and Schluter, D. (2008) Natural selection on a major armor gene in threespine stickleback. *Science.* **322**, 255-257.

Beaumont, M.A. and Balding, D.J. (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13** 969-980.

Benham, J., Jeung, J., Jasieniuk, M., Kanazin, V., and Blake, T. (1999) Genographer: A graphical tool for automated fluorescent AFLP and microsatellite analysis. *Journal of Agricultural Genomics*. **4**:3.

Bensch, S. and Åkesson, M. (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*. **14**, 2899-2914.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., and Taberlet, P. (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology.* **13**, 3261-3273.

Bonin, A., Pompanon, F., and Taberlet, P. (2005) Use of amplified fragment length polymorphism (AFLP) markers in surveys of vertebrate diversity. *Methods in Enzymology*. **395**, 145-161.

Bonin, A., Taberlet, P., Miaud, C., and Pompanon, F.(2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution* **23**, 773-783.

Bonin, A., Ehrich, D. and Manel, S. (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737-3758.

Bradbury, I.R. and Bentzen, P. (2007) Non-linear genetic isolation by distance: implications for dispersal estimation in anadromous and marine fish populations. *Marine Ecology Progress Series.* **340**, 245-257.

Buonaccorsi, V. P., Kimbrell, C.A., Lynn, E.A., and Vetter, R.D. (2002) Population structure of copper rockfish (*Sebastes caurinus*) reflects postglacial colonization and contemporary patterns of larval dispersal. *Canadian Journal of Fisheries and Aquatic Sciences.* **59**, 1374-1384.

Carr, M. and Syms, C. (2006) Recruitment. In: Allen, L.G., Pondella, D.J., and Horn, M.H. (eds) The ecology of marine fishes: California and adjacent waters. University of California Press, Berkeley, 411-427.

Canino, M.F., O'Reilly, P.T., Hauser, L., and Bentzen, P. (2005) Genetic differentiation in walleye Pollock (*Theragra chalcogramma*) in response to selection at the pantophysin (*PanI*) locus. *Canadian Journal of Fisheries and Aquatic Sciences* **62**, 2519-2529.

Case, R.A.J., Hutchinson, W.F., Hauser, L., Van Oosterhout, C. and Carvalho, G.R. (2005) Macro- and microgeographic variation in pantophysin (*Pan I*) allele frequencies in NE Atlantic cod *Gadus morhua*. *Marine Ecology Progress Series*. **301**, 267-278.

Case, R.A.J., Hutchinson, W.F., Hauser, L., Buehler, V., Clemmesen, C., Dahle, G., Kjesbu, O.S., Moksness, E., Ottera, H., Paulsen, H., Svasand, T., Thorsen, A. and Carvalho, G.R. (2006) Association between growth and *PanI* genotype within Atlantic cod full-sibling families. *Transactions of the American Fisheries Society*. **135**, 241-250.

Colbeck, G.J., Turgeon, J., Sirois, P. and Dodson, J.J. (2011) Historical introgression and the role of selective vs. neutral processes in structuring nuclear genetic variation (AFLP) in a circumpolar marine fish, the capelin (*Mallotus villosus*) *Molecular Ecology*. **20**, 1976-1987.

Conover, D.O., Clarke, L.M., Munch, S.B. and Wagner, G.N. (2006) Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology* **69**, 21-47.

Cowen, R.K., Paris, C.B., and Srinivasan, A. (2006) Scaling of connectivity in marine populations. *Science*. **311**, 522-527.

Dower, J., Yelland, D., and Crawford, W. (2001) Physical and biological interactions between Haida eddies and seamounts in the NE Pacific. **PICES 10th Annual Meeting, October 5-13, Victoria, B.C., Canada.**

Earl, D.A. (2011) Structure harvester v0.6.1. Available at http://taylor0.biology.ucla.edu/structureHarvester/ (accessed Summer 2011)

Ehrich, D. (2006) AFLPDAT: a collection of R functions for convenient handling of ALFP data. *Molecular Ecology Notes*. **6**, 603-604.

Estoup, A. Tailliez, C., Cornuet, J. and Solignac, M. (1995) Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution*. **12**, 1074-1084.

Estoup, A., Jarne, P., and Cornuet, J. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*. **11**, 1591-1604.

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

Falush, D., Stephens, M., and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*. **7**, 574–578.

Fevolden, S.E. and Pogson, G.H. (1997) Genetic divergence at the synaptophysin (*Syp I*) locus among Norwegian coastal and north-east Arctic populations of Atlantic cod. *Journal of Fish Biology*. **51**, 895-908.

Foll, M. and Gaggiotti, O. (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics*. **174**, 875–891.

Foll, M. and Gaggiotti, O. (2008) A genome scan method to identify selected loci appropriate for both dominant and co-dominant markers: a Bayesian perspective. *Genetics*. **180**, 977-993.

Freeland, H.J., Crawford, W.R. and Thomson, R.E. (1984) Currents along the pacific coast of Canada. *Atmosphere-Ocean*. **22**, 151-172.

Foll, M., Fischer, M.C., Heckel, G. and Excoffier, L. (2010) Estimating population structure from AFLP amplification intensity. *Molecular Ecology*. **19**, 4638-4647.

Gaggiotti, O., Bekkevold, D., Jørgensen, H., Foll, M. Carvalho, G.R., Andre, C., and Ruzzante D.E. (2009) Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure in natural populations: Atlantic herring as a case study. *Evolution*. **63**, 2939-2951.

Gaggiotti, O. (2010) Bayesian statistical treatment of the fluorescence of AFLP bands leads to accurate genetic structure inference. *Molecular Ecology*. **19**, 4586-4588.

Garant, D.L., Pietsch, T.W., Utter, F.M. and Hauser, L. (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*. **21**, 434-443.

Gaudeul, M., Till-Bottraud, I., Barjon, F. and Manel, S. (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers. *Heredity*. **92**, 508-518.

Gunderson, D.R. and Vetter, R.J. (2006) Temperate Rocky Reef Fishes. In: Kritzer, J.P and Sale, P.F. (eds) Marine Metapopulations. Elsevier Academic Press, San Francisco, 69-117.

Hamilton, S.L., Regetz, J., and Warner, R.R. (2008) Postsettlement survival linked to larval life in a marine fish. *Proceedings of the National Academy of Sciences of the United States*. **105**, 1561-1566.

Harding, L. and Hirvoven, H. (1996) Marine ecological classification for Canada. Preliminary report to Environment Canada. Vancouver, British Columbia.

Harper, J.R., Christian, J., Cross, W.E., Frith, R., Searing, G.F. and Thompson, D. (1993) *A classification of the marine regions of Canada*. Ottawa, Ontario, Canada: Environment Canada.

Hauser, L. and Carvalho, G.R. (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*. **9**, 333-362.

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

Hemmer-Hansen, J., Nielsen, E.E., Frydenberg, J. and Loeschcke, V. (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European Flounder (*Platichthys flesus* L.). *Heredity*. **99**, 592-600.

Herrmann, D., Poncet, B.N., Manel, S., Rioux, D., Gielly, L., Taberlet, P. and Gugerli, F. (2010) Selection criteria for scoring amplified fragment length polymorphisms (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*. **53**, 302-310.

Hickey, B.M. (1979) The California current system – hypotheses and facts. *Progress in Oceanography*. **8**, 191-279.

Hilborn, R., Quinn, T.P., Schindler, D.E., and Rogers, D.E. (2003) Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences of the United States of America.* **100**, 6564-6568.

Hirvonen, H. L., Harding, L. and Landucci, J. (1995) National marine ecological framework for ecosystem monitoring and state-of-the-environment reporting. In Shackel, N. and Woolaston, M. (eds) *MPA and sustainable fisheries*, 117-129. Ottawa, Ontario, Canada: Science and Management of Protected Areas Association.

Holland, B.R., Clarke, A.C. and Meudt, H.M. (2008) Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology*. **57**, 347-366.

Jamieson, G. and Phillips, A. (1988) Occurrence of *Cancer* crab (*C. magister* and *C. oregonenis*) megalopae off the west coast of Vancouver Island, British Columbia. *Fisheries Bulletin, U.S*. **86**, 425-437.

Jenczewski, E., Prosperi, J.M. and Ronfort, J. (1999) Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes. *Molecular Ecology*. **8**, 1317-1330.

Johansson, M.L., Banks, M.A., Glunt, D.D., hassel-Finnegan, H.M., and Buonaccorsi, V.P. (2008) Influence of habitat discontinuity, geographical distance, and oceanography on fine-scale population genetic structure of copper rockfish (*Sebastes caurinus*). *Molecular Ecology*. **17**, 3051-3061.

Jones, G.P., Milicich, M.J., Emslie, M.J. and Lunow, C. (1999) Self-recruitment in a coral reef fish population. *Nature*. **402**, 802-804.

Jónsdóttir, I.G., Marteinsdóttir, G. and Pampoulie, C. (2008) Relation of growth and condition with the *Pan I* locus in Atlantic cod (*Gadus morhua* L.) around Iceland. *Marine Biology*. **154**, 867-874.

Joost, S., Bonin, A., Bruford, M.W., Després, L., Conord, C., Erhardt, G., and Taberlet, P. (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*. **16**, 3955-3969.

Kinlan, B.P. and Gaines, S.D. (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*. **84**, 2007-2020.

Krutovskii, K.V., Erofeeva, S.Y., Aagaard, J.E. and Strauss, S.H. (1999) Simulation of effects of dominance on estimates of population genetic diversity and differentiation. *Journal of Heredity.***90**, 499-502.

Latta, R.G. (2003) Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytologist*. **161**, 51-58.

Leinonen, T. O'Hara, R.B., Cano, J.M., Merilä, J. (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*. **21**, 1-17.

Lenfant, P. and Planes, S. (2002) Temporal genetic changes between cohorts in a natural population of a marine fish, *Diplodus sargus*. *Biological Journal of the Linnean Society* **76**, 9-20.
Lenormand, T. (2002) Gene flow and the limits to natural selection. *TRENDS in Ecology & Evolution*. **17**, 183-189.

Levin, L.A. (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*. **46**, 282-297.

Love, M.S., Yoklavich, M.M. and Thorsteinson, L. (2002) The Rockfishes of the Northeast Pacific. *University of California Press*. Berkley and Los Angeles, California.

Lynch, M. (1996) A quantitative genetic perspective on conservation issues. In: *Conservation Genetics: Case Studies from nature* (eds J.C. Avise and J.L. Hamrick). Kluwer Academic Publications, Norwell, MA, pp. 471-501.

Mantel, N.A. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*. **27**, 209-220.

Mariette, S., Chagné, D., Lézier, C., Pastuszka, P., Raffin, A., Plomion, C., and Kremer, A. (2001) Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity*, **86**, 469-479.

Marshall, C.T. and Browman, H.I. (2007) Disentangling the causes of maturation trends in exploited fish populations. *Marine Ecology Progress Series* **335**, 249-251.

Marshall, D.J., Monro, K., Bode, M., Keough, M.J., and Swearer, S. (2010) Phenotype-environment mismatches reduce connectivity in the sea. *Ecology Letters*. **13**, 128-140.

McCusker, M.R. and Bentzen, P. (2010) Historical influences dominate the population genetic structure of a sedentary marine fish, Atlantic wolfish (*Anarhichas lupus*), across the North Atlantic Ocean. *Molecular Ecology*, **19** 4228-4241.

McCusker, M.R. and Bentzen, P. (2011) Limited population structure in Northern and Spotted Wolffishes (*Anarhichas denticulatus* and *A. minor*) despite low apparent dispersal potential. *Marine Biology*, **158**, 1869-1878.

McKay, J.K. and Latta, R.G. (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*. **17**, 285-291.

Merilä, J. and Crnokrak, P. (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*. **14**, 892-903.

Meudt, H.M., and Clarke, A.C. (2007) Almost forgotten or latest practice: AFLP applications, analyses and advances. *TRENDS in Plant Science*. **12**, 106-117.

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

Mitarai, S., Siegel, D.A. and Winters, K.B. (2008) A numerical study of stochastic larval settlement in the California Current system. *Journal of Marine Systems*. **69**, 295-309.

Moen, T., Hayes, B., Nilsen, F., Delghandi, M., Fjalestad K.T., Fevolden, S., Berg, P.R. and Lien, S. (2008) Identification and characterisation of novel SNP markers in Atlantic cod: Evidence for directional selection. *BMC Genetics*. **9**, doi:10.1186/1471-2156-9-18

Naish, K.A., Hard, J.J. (2008) Bridging the gap between the genotype and the phenotype: linking genetic variation, selection and adaptation in fishes. *Fish and Fisheries* **9**, 396-422.

Nielsen, E.E., Hansen, M.M. and Ruzzante, D.E. (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology*. **12**, 1497-1508.

Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F. and Bekkevold, D. (2009a) Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology*. **18**, 3128-3150.

Nielsen, E.E., Hemmer-Hansen, J., Poulsen, N.A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G., Ogden, R. and Carvalho, G.R. (2009b) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (Gadus morhua). *BMC Evolutionary Biology*. **9**, doi:10.1186/1471-2148-9-276.

Nosil, P., Funk, D.J. and Ortiz-Barrientos, D. (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*.**18**, 375-402.

Nybom, H. (2004) Comparison of different nuclear DNA markers for estimating intraspecific diversity in plants. *Molecular Ecology*. **13**, 1143-1155.

Palumbi, S.R. (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, **Supp**, S146-S158.

Pampoulie, C., Ruzzante, D.E., Chosson, V., Jörundsdóttir, T.D., Taylor, L, Thorsteinsson, V., Danielsdóttir, A.K., Marteinsdóttir, G. (2006) The genetic structure of Atlantic cod (*Gadus Morhua*) around Iceland: insight from microsatellites, the *PanI* locus, and tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences*. **63** (12) 2660-2674.

Paris, M., Boyer, S., Bonin, A., Collado, A., David J., and Despres, L. (2010) Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. *Molecular Ecology*. **19**, 325-337.

Paupy, C. Orsoni, A., Mousson, L. and Huber, K. (2004) Comparisons of amplified fragment length polymorphism (AFLP), microsatellite, and isoenzyme markers: Population genetics of Aedes aegypti (Diptera: Culicidae) from Phnom Penh (Cambodia). *Journal of Medical Entomology*. **41**, 664-671.

Parker, S.J., Berkeley, S.A., Golden, J.T., Gunderson, D.R., Heifetz, J., Hixon, M.A., Larson, R., Leaman, B.M., Love, M.S., Musick, J.A., O'Connell, V.M., Ralston, S., Weeks, H.J., Yoklavich, M.M. (2000) Management of Pacific Rockfishes. *Fisheries*, **25**, 22-29.

Planes, S. and Lenfant, P. (2002) Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Molecular Ecology*. **11**, 1515-1524.

Pogson, G.H., Mesa, K.A., and Boutilier, R.G. (1995) Genetic population structure and gene flow in the Atlantic cod, Gadus morhua: A comparison of allozyme and nuclear RFLP loci. *Genetics*. **129**, 375-385.

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

Pogson, G.H. (2001) Nucleotide polymorphism and natural selection at the pantophysin (Pan I) locus in the Atlantic cod, Gadus morhua (L.). *Genetics*, **157**, 317-330.

Pompanon, F., Bonin, A., Bellemain, E., and Taberlet, P. (2005) Genotyping errors: causes, consequences and solutions. *Nature Genetics*. **6**, 847-859.

Powers, D.A. and Schulte, P.M. (1998) Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: A multidisciplinary approach to address the million-year saga of a small fish. *Journal of Experimental Zoology*. **282**, 71-94.

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

Roberts, C.M., Branch, G., Bustamante, R.H., Castilla, J.C., Dugan, J., Halpern, B.S., Lafferty, K.D., Leslie, H., Lubchenco, J., McArdle, D., Ruckelshaus, M. and Warner, R.R. (2003) Application of ecological criteria in selecting marine reserves and developing reserve networks. *Ecological Applications*, supplement. **13**, S215-S228.

Rocha-Olivares, A., and Vetter, R.D. (1999) Effects of oceanographic circulation on the gene flow, genetic structure, and phylogeography of the rosethorn rockfish (*Sebastes helvomaculatus*). *Canadian Journal of Fisheries and Aquatic Sciences,* **56**, 803-813.

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

Scheltema, R.S. (1986) On dispersal and planktonic larvae of benthic invertebrates: an eclectic overview and summary of problems. *Bulletin of Marine Science*. **39**, 290-322.

Shanks, A.L., Largier, J., Brink, L., Brubaker, J. and Hooff, R. (2000) Demonstration of the onshore transport of larval invertebrates by the shoreward movement of an upwelling front. *Limnology and Oceanography*. **45**, 230-236.
Slatkin, M. (1973) Gene flow and selection in a cline. *Genetics*. **75**, 733-756.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science*. **236**, 787-792.

Storz, J.F. (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*. **14**, 671-688.

Swearer, S.E., Shima, J.S., Hellberg, M.E., Thorrold, S.R., Jones, G.P., Robertson, D.R., Morgan, S.G., Selkoe, K.A., Ruiz, G.M. and Warner, R.R. (2002) Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science.* **70**, 251-271.

Thomson, R.E. (1981) Oceanography of the British Columbia Coast. *Can. Spec. Publ. Fish. Aquat. Sci.* **56**, 291 p.

Thorson, G. (1950) Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews.* **25**, 1-45.

Vekemans, X. (2002) AFLP-SURV version 1.0. Distributed by author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Vigliola, D., Doherty, P.J., Meekan, M.G., Drown, D.M., Jones, M.E. and Barber, P.H. (2007) Genetic identity determines risk of post-settlement mortality of a marine fish. *Ecology*, **88**, 1263-1277.

Vos, P. Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research.* **23**, 4407-4414.

Waples, R.S. (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**, 438-450.

Ward, R.D., Woodwark, M. and Skibinski, D.O.F. (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fish Biology* **44**, 213-232.

Wainwright, P., J. Harper, G. F. Searing, and Howes, D.E. (1995) Ecological classification within British Columbia's marine regions. Victoria, British Columbia, Canada: Land Use Coordination Office.

Weber, J.L. and Wong, C. (1993) Mutation of human short tandem repeats. *Human Molecular Genetics.* **2**, 1123-1128.

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

Whitlock, R., Hipperson, H., Mannarelli, M., Butlin, R.K. and Burke T. (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources.* **8**, 725-735.

Woodhead, M., Russell, J., Squirrell J., Hollingsworth, M., Mackenzie, K., Gibby, M. and Powell, W. (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Molecular Ecology*. **14**, 1681-1695.
Wright, S. (1931) Evolution in Mendelian populations. *Genetics*. **16**, 97-159.

Wright, S. (1943) Isolation by distance. *Genetics*. **28**, 114-138.

Yamanaka, K.L., Lacko, L.C., Withler, R., Grandin, C., Lochead, J.K., Martin, J.C., Olsen, N. and Wallace, S.S. (2006) A review of yelloweye rockfish (*Sebastes ruberrimus)* along the Pacific coast of Canada: biology, distribution and abundance trends. *Canadian Science Advisory Secretariat.* Research Document 2006/076.

Yamanaka, K.L. and Logon, G. (2010) Developing British Columbia's inshore rockfish conservation strategy. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science.* **2**, 28-46.

Zacharias, M.A., Howes, D.E., Harper, J.R. and Wainwright, P. (1997) The British Columbia marine ecosystem classification: Rationale, development, and verification. *Coastal Management*. **26**, 105-124.

Zhivotovsky, L.A. (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*. **8**, 907-913.

APPENDICES

Appendix A: AFLPs and the AFLPSCORE scoring procedure

AFLP methods

Vos *et al.* (1995) describe the development of genetic markers though the selective amplification of restriction fragments. They termed this type of genetic marker, amplified fragment length polymorphisms (known as AFLPs), which are a relatively inexpensive and rapid way to accumulate markers randomly dispersed throughout the genome. AFLPs are generated by digesting genomic DNA with restriction enzymes (traditionally EcoRI and MseI) and ligating short oligonucleotide adapters (~20 base pairs long) to the restriction cut sites. These adapters (whose sequences are known) are then used to generate primers for two successive rounds of PCR. The first round, called the preselective amplification, uses a primer with an additional base that extends past the adapter, theoretically reducing the number of fragments that will amplify by a factor of four. The final PCR is known as the selective amplification, and primers that contain an additional two selective bases are used, further reducing the number of DNA fragments that will amplify by a factor of 16. The final PCR products are separated via electrophoresis.

Each fragment represents a locus, and an individual is scored at each locus for the "presence" or "absence" of that particular fragment. Many loci can be generated through the use of different selective primer combinations. For example, a primer pair might consist of one primer that anneals to the EcoRI cut site, ending with E-**AAC** (the AAC represents the three selective bases that extend past the E adapter) and a primer annealing to MseI cut site, ending with M-**CAA**. Every base pair change in the E or M selective bases represents a new primer combination and theoretically amplifies a new set of fragments. If we keep the M primer constant (M–CAA) and change the E-ACC to E-ACT, we will amplify a new set of fragments that are complementary to our new E primer (E-ACT complementary to TGA). Thus, a much greater number of loci can be generated by utilizing primers with different selective bases.

The AFLPSCORE approach

Differences in overall fluorescence intensities between electropherogram profiles can lead to incorrect genotyping because the fluorescence intensity thresholds used may not be appropriate for profiles with consistently low peak heights. Furthermore, distinguishing loci with consistently low peak heights from background noise can be challenging and lead to genotyping errors. Therefore, the first stage in the AFLPSCORE analysis is the normalization of the raw peak-height data. The total fluorescence of each electropherogram profile (i) is calculated as the sum of every peak-height within the profile. The median value of profile intensity (m) is calculated across all the samples in the project, and a normalization factor (m/i) is calculated for each individual. All the raw peak-heights for each individual are multiplied by that individual's unique normalization factor, creating a new normalized data matrix used to assign AFLP phenotypes. This normalization procedure helps to reduce the variation in PCR amplification between individuals, which can result in assigning incorrect genotypes.

AFLPSCORE uses two thresholds (set by the user) that generate the final bin set and the phenotype scores for each individual: the locus selection threshold and the phenotype-calling threshold. For each locus, the mean peak height of all peaks is calculated from the normalized data. Loci whose mean peak height is low may be more difficult to score accurately, due to individual drop-out at low intensity and the non-detection of peaks. The locus selection threshold is used as a cutoff point applied to the mean peak height data, such that loci with mean peak heights lower than the threshold are discarded. Once the final bin set is established, the phenotype-calling threshold is used to assign the presence/absence scores. An absolute or relative threshold value may be used as the cutoff. For example, an absolute cutoff of 100 rfus results in individuals with a peak height less than 100 rfus getting an absence score, while individuals with a peak height greater than 100 rfus gets assigned a presence score. Likewise, the relative scoring threshold uses a set percentage of the mean peak height as the cutoff. A relative threshold of 10% results in absence scores assigned when an individual's peak height for a locus is lower than 10% of the mean peak height.

An important component of AFLPSCORE is the ability to calculate error rates among a set of replicated individuals, which can then be used to fine-tune the locus selection and phenotype-calling thresholds to maximize the scoring accuracy. A suite of different locus selection and phenotype-calling thresholds are specified for the user and the resulting error rates are calculated for each pair of thresholds. A threshold combination that maximizes the number of loci retained with an acceptable error rate is determined and used for the final scoring analysis.

Appendix B: Discussion of plate effects

A plate effect exists when individuals grouped on the same PCR plate appear most genetically similar with each other, regardless of sample location or other geographical or biological differences between samples. For example, if a plate effect is present, individuals from different locations amplified together on the same PCR plate would appear more similar than individuals from the same locations but run on a different plate. These effects may arise from variation in any of the laboratory steps (digestion, ligation, amplification), and variation in the migration of fragments in the capillary of the genotyping machine between runs. Despite a high potential for plate effects to exist, there is surprisingly little written about dealing with these biases in the literature. A few AFLP scoring programs indirectly deal with this effect by reducing variation in amplification intensities (such as GENOGRAPHER (Behham et al. 2009); AFLPSCORE; RAWGENO (Arrigo et al. 2009)). However, simply accounting for variation in amplification intensity does not remedy other potential sources of plate effects, such as variation in background noise or the preferential amplification of fragments between plates.

While a robust, well established AFLP protocol is the best way to ensure high quality electropherograms and a minimization of plate effects, it may be difficult to detect the presence of plate effects before a final dataset is generated, as downstream analyses (such as multidimensional scaling analyses or assignment methods) may be the only way to detect such biases. Furthermore, as DNA sequencing costs continue to decrease, population genetic studies are including a greater number of individuals and loci, and requiring a greater number of PCR plates. The increase in PCR plates increases the number of plate comparisons and potential for a plate effects.

Plate effects within this study

Plate effects were identified in multidimensional scaling analyses (Figure 4a-d) with a dataset generated in AFLPSCORE. The most likely culprit responsible for the plate effect is variation in the number of fragments amplified between plates, which could have resulted from variable success in the digestion, ligation and/or preselective amplification laboratory steps. This can be observed when the proportion of

bins that have scores and the proportion of bins with a presence score are plotted by plate (Figure 6 and Figure 7). I tested the significance of plate as a predictor of genotype frequency for each locus with a generalized linear model to identify loci whose amplification was restricted to individuals from the same plate. The majority of loci (~75%) were shown to have this significant association (p-value >0.05, Figure 8) and were dropped from the dataset. Unfortunately, new multidimensional scaling analyses on the filtered dataset still revealed strong genetic similarity between individuals from the same PCR plate.

Final scoring strategy

Unable to eliminate the plate effect, I decided to remove all between plate comparisons and treat each plate as its own independent project. The two major drawbacks to this approach are (1) reduction in sample size for each sampling location and (2) inability to analyze and compare all sample locations together. My project includes a total of seven plates, six of which contain four sample locations with an average of 23 individuals per location and one plate with two locations containing 46 and 37 individuals. While this approach is certainly not preferable, the strength of the plate effect made distinguishing any biological signal from the plate signal impossible.

After the final bin set and AFLP phenotype scores were determined in AFLPSCORE, 46 individuals (6.8% of total sample) contained more than 15% missing data and were dropped from the project. Individuals were sorted by plate and binary tables were generated for each plate. The bin set for each plate was filtered a final time, and only bins with a presence frequency greater than 5% were retained.

Appendix C: Individual plate results

Plate 01 (BS, ES, TA, TI):

Table A.1. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. BS=Bowie Seamount, ES=Esperanza, TA=Tasu, TI=Texada Island. Outside sites are **bolded**.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
BS	23	859	83.8	0.264 (0.005)	18	100	0.374 (0.028)
ES	22	796	77.7	0.264 (0.005)	14	77.8	0.242 (0.038)
TA	23	710	69.3	0.223 (0.005)	13	72.2	0.233 (0.044)
TI	23	839	81.9	0.270 (0.005)	1	5.6	0.031 (0.011)

Table A.2. Pairwise F_{ST} values in yelloweye rockfish calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.1.

	BS	ES	TA	TI
BS	---	0.3746	0.3704	0.5525
ES	0.0723	---	0.0369	0.8349
TA	0.0870	0.0454	---	0.8405
TI	0.0769	0.1324	0.1487	---

Table A.3. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	BS-ES-TA vs. TI	outside vs. inside clusters
Outlier	BS vs. ES vs. TA vs. TI	clusters associated with Ecosections

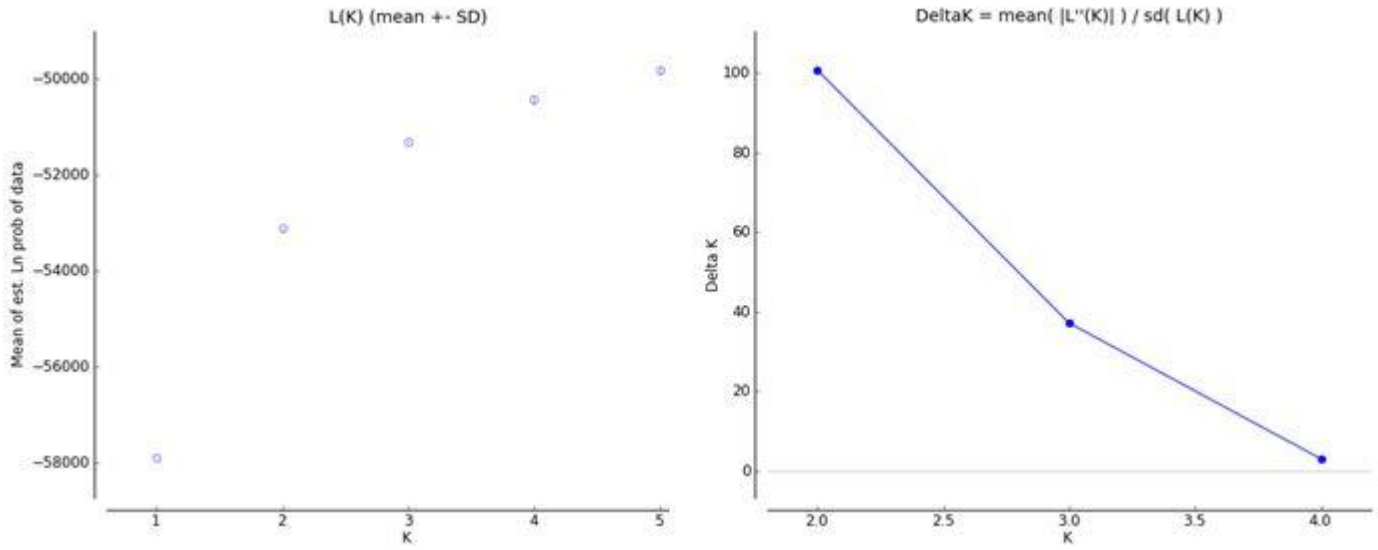
Results-

For both the neural and outlier datasets three clusters are observed. Two of the four Ecosections (BS and TI) cluster independently, while ES and TA cluster together. The clustering pattern is consistent with expectations for the outlier loci for two of the four representative Ecosections, and consistent with expectations for the neutral loci in that ES and TA cluster together.

Plate 01 continued

Figure A.1. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral dataset



Outlier dataset

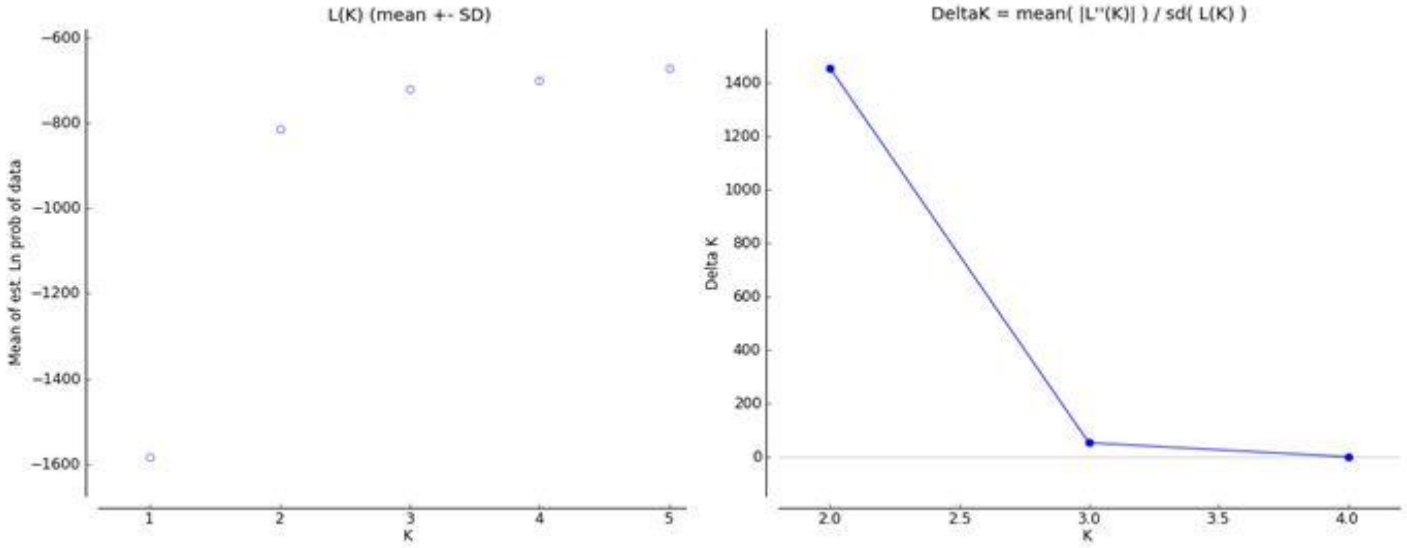
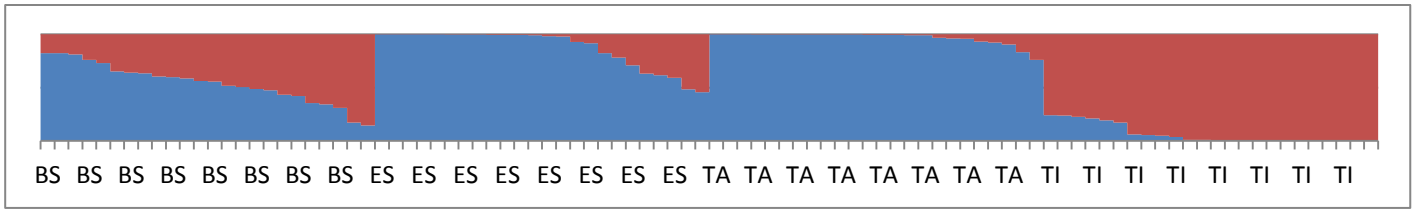


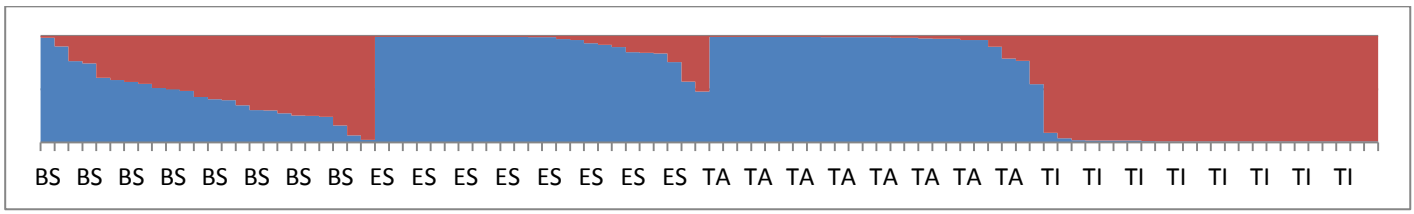
Plate 01 continued

Figure A.2. STRUCTURE output for Plate 01 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1025 neutral loci and 18 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a differnet color (red, blue, and purple). BS=Bowie Seamount, ES=Esperanza, TA=Tasu, TI=Texada Island. BS, ES and TA correspond to 'outside' locations, while TI corresponds to an 'inside' location.

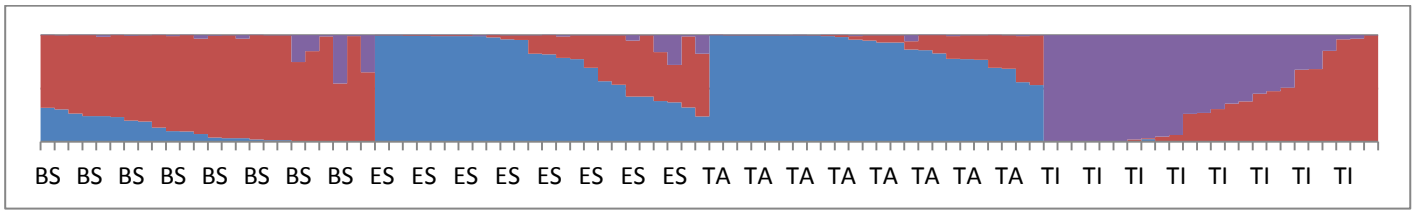
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

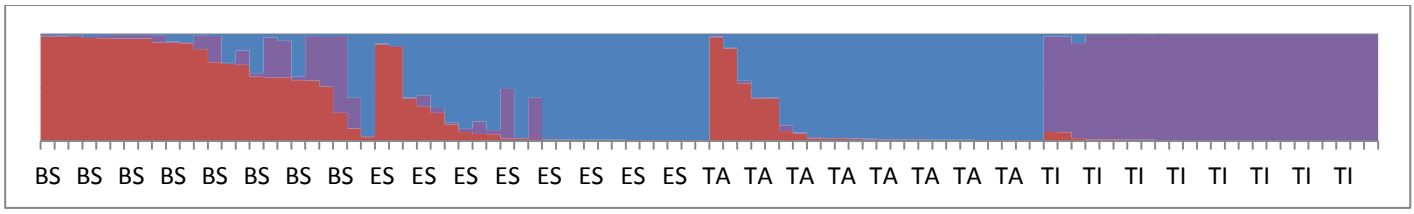


Plate 01 continued

Figure A.3. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. BS=Bowie Seamount, ES=Esperanza, TA=Tasu, TI=Texada Island.

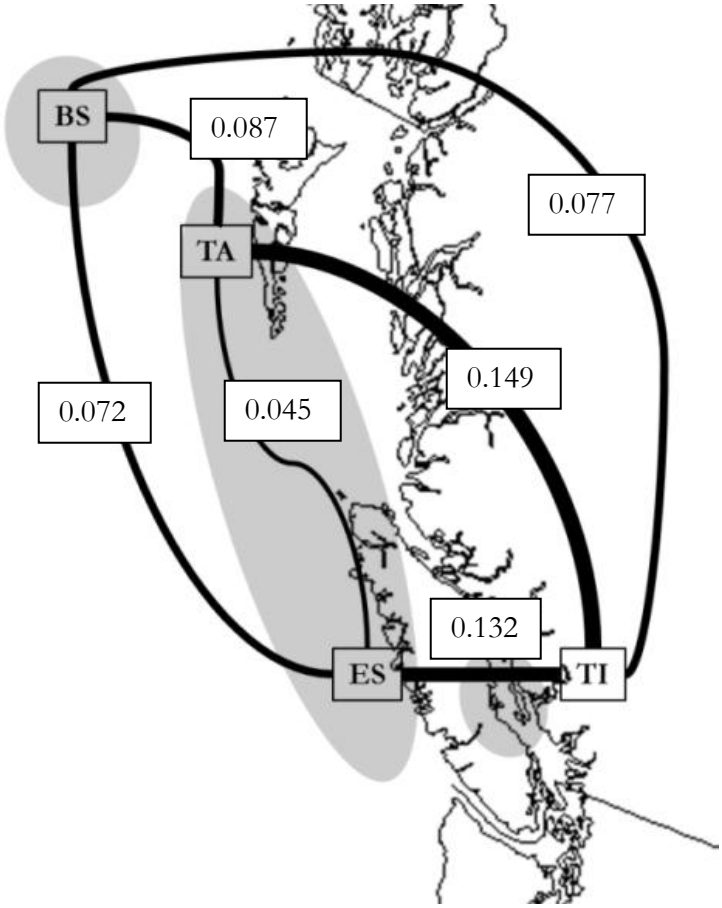


Plate 02 (CI, GA, GC, WA):

Table A.4. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. CI=Calvert Island, GA=Gabriola, GC=Gordon Channel, WA=Washington. Outside sites are **bolded**.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
CI	21	943	78.8	0.257 (0.005)	5	31.3	0.127 (0.037)
GA	16	798	66.7	0.229 (0.005)	10	62.5	0.183 (0.043)
GC	10	794	66.3	0.215 (0.005)	15	93.8	0.330 (0.043)
WA	16	762	63.7	0.209 (0.005)	4	25.0	0.105 (0.030)

Table A.5. Pairwise F_{ST} in yelloweye rockfish values calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.4.

	CI	GA	GC	WA
CI	---	0.8108	0.4450	0.8642
GA	0.1658	---	0.5517	0.1371
GC	0.1095	0.1048	---	0.6196
WA	0.1938	0.0388	0.1262	---

Table A.6. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	CI-GC-WA vs. TA	outside vs. inside clusters
Outlier	GA vs. WA vs. CI-GC	GA and WA belong to different Ecosections; CI and GC are in different Ecosections despite being close in proximity

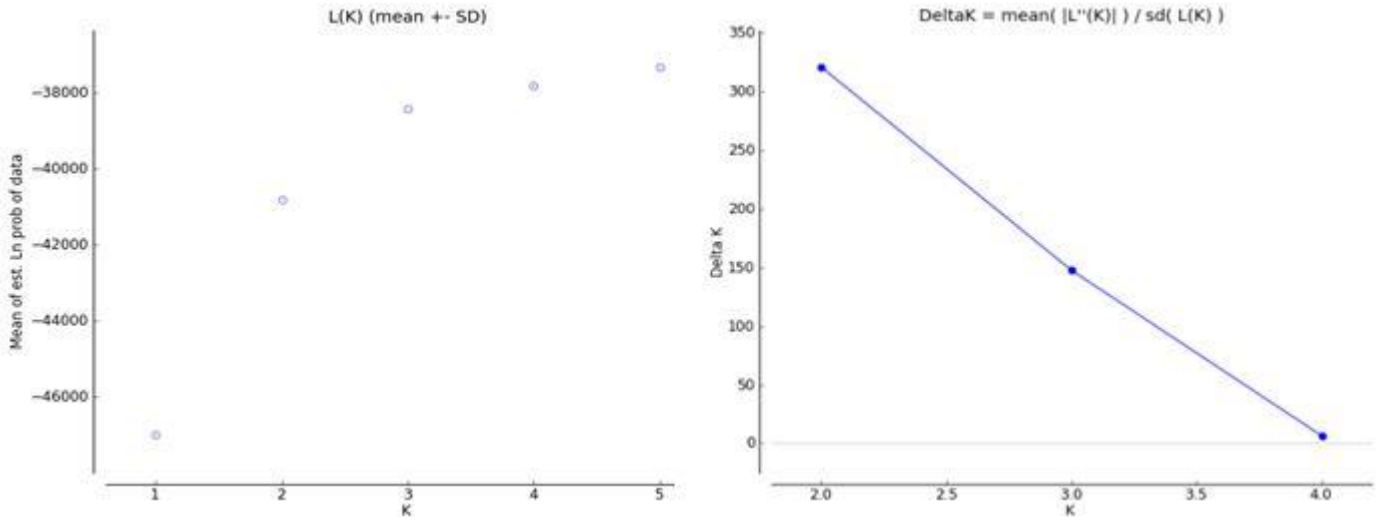
Results-

The clustering pattern does not meet either neutral or outlier expectations. A more southern cluster (GA and WA) is observed, while CI and GC individuals generally assign to the same cluster. Increasing K to three introduces a GC cluster, which is more apparent in the outlier dataset.

Plate 02 continued

Figure A.4. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral dataset



Outlier dataset

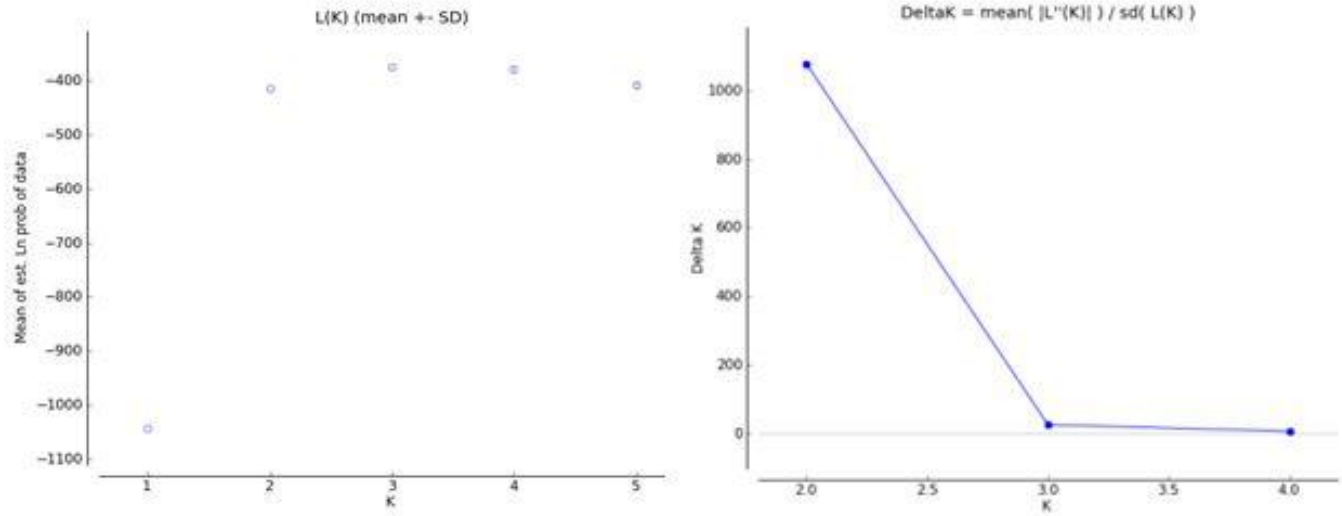
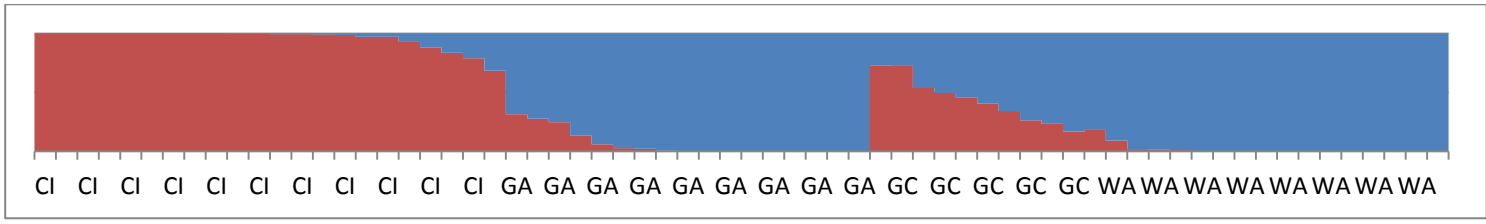


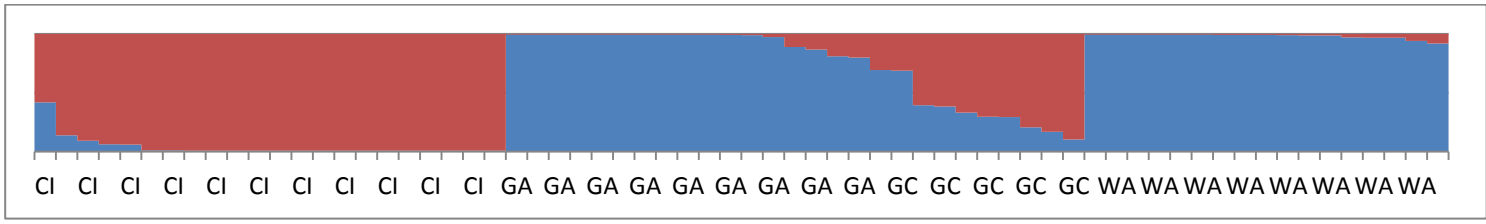
Plate 02 continued

Figure A.5. STRUCTURE output for Plate 02 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1197 neutral loci and 16 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a differnet color (red, blue, and purple). CI=Calvert Island, GA=Gabriola, GC=Gordon Channel, WA=Washington. CI, GC, and WA correspond to 'outside locations' while GA corresponds to an 'inside' location.

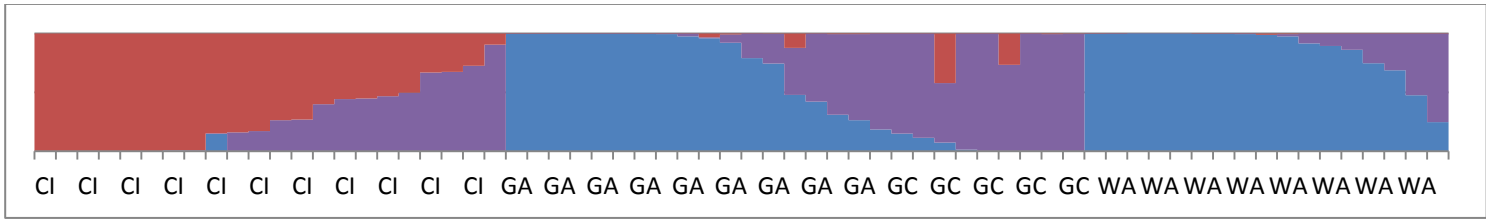
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

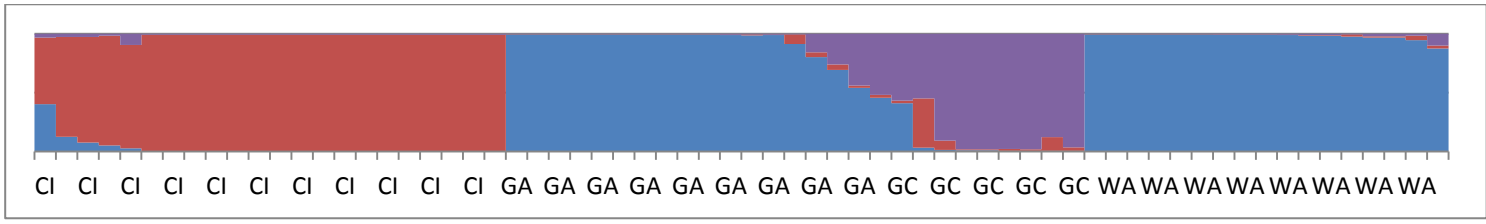


Plate 02 continued

Figure A.6. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. CI=Calvert Island, GA=Gabriola, GC=Gordon Channel, WA=Washington.

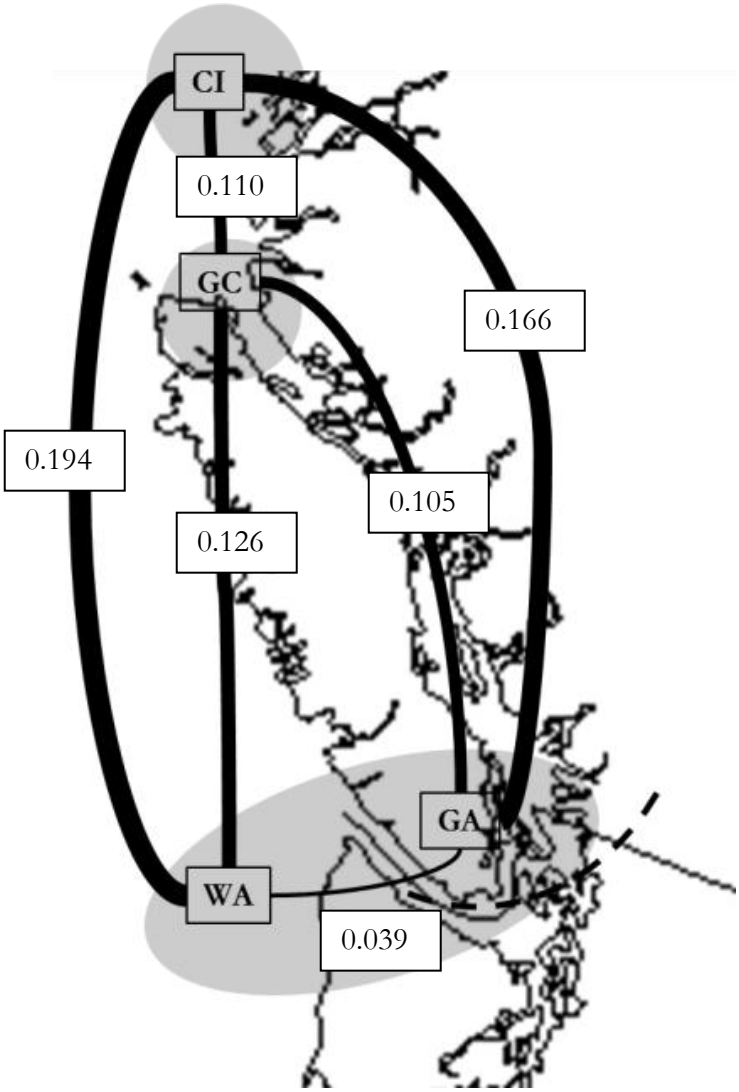


Plate 03 (LB, MI, OR, TK):

Table A.7. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. LB=Lower Bute Inlet, MI=Mittlenatch, OR=Oregon, TK=Top Knot. Outside sites are **bolded**.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
LB	23	1146	78.8	0.241 (0.004)	5	83.3	0.297 (0.070)
MI	24	994	68.3	0.238 (0.004)	2	33.3	0.132 (0.076)
OR	23	811	55.7	0.202 (0.005)	6	100	0.246 (0.059)
TK	23	903	62.1	0.218 (0.005)	6	100	0.439 (0.024)

Table A.8. Pairwise F_{ST} values in yelloweye rockfish calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.7.

	LB	MI	OR	TK
LB	---	0.6640	0.2867	0.1139
MI	0.0460	---	0.7061	0.4065
OR	0.0460	0.0649	---	0.2479
TK	0.0470	0.0371	0.0368	---

Table A.9. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	OR-TK vs. LB-MI	outside vs. inside clusters
Outlier	OR vs. TK vs. LB-MI	clusters associated with Ecosections

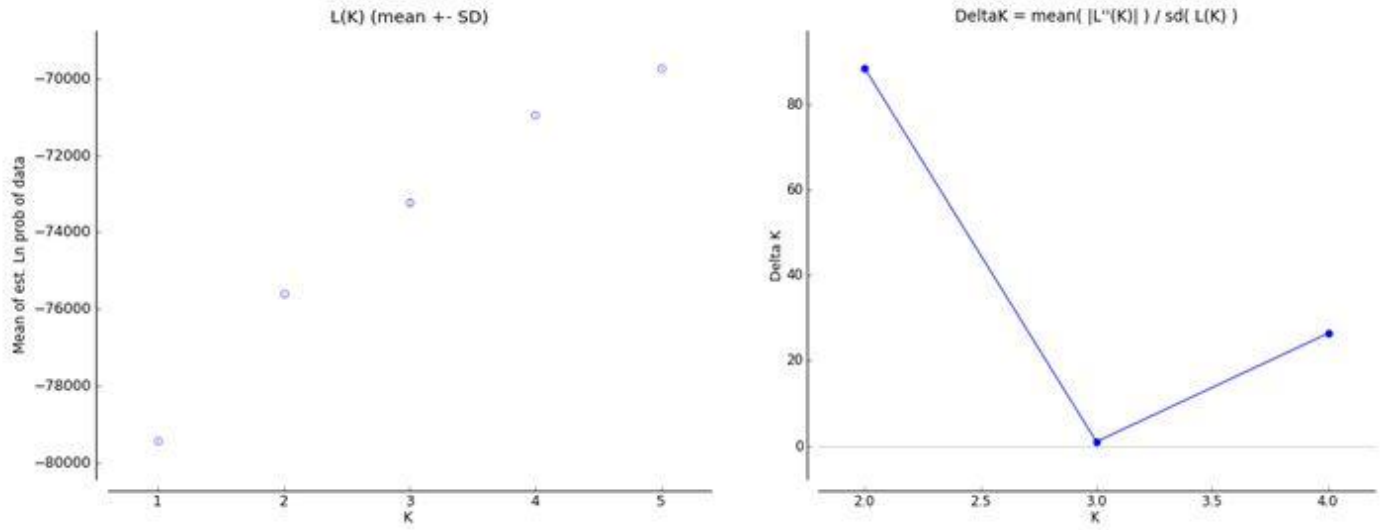
Results-

No clustering patterns matched the neutral or outlier expectations. With a *K* of 3, MI and OR individuals assigned strongly to different clusters. However, LB individuals did not cluster with MI individuals and showed mixed assignment to the OR and third cluster. TK individuals showed mixed assignment to all three clusters.

Plate 03 continued

Figure A.7. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral dataset



Outlier dataset

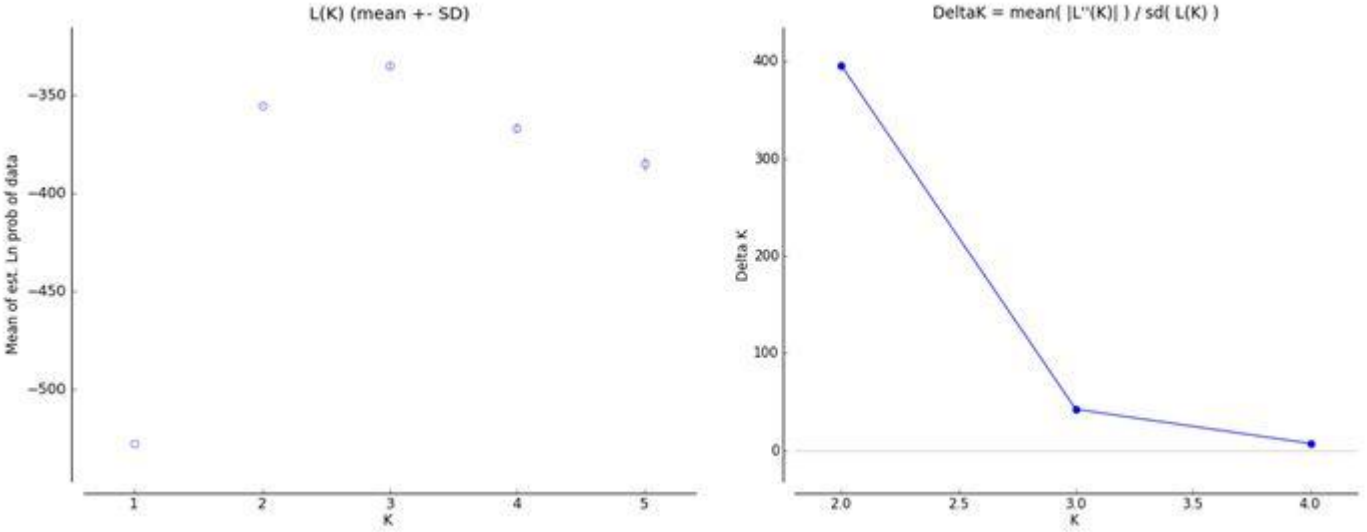
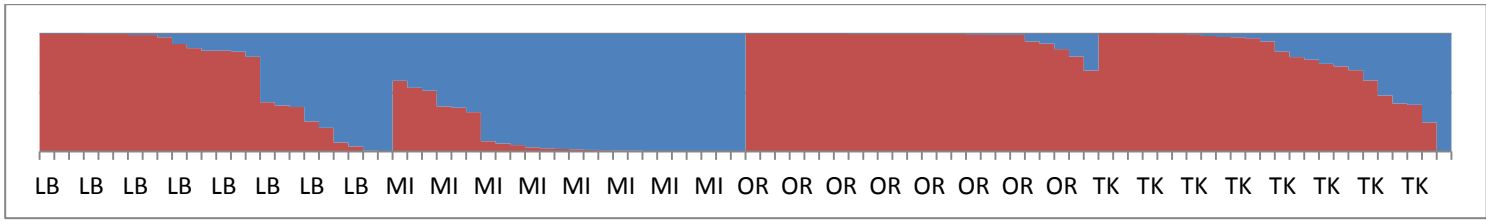


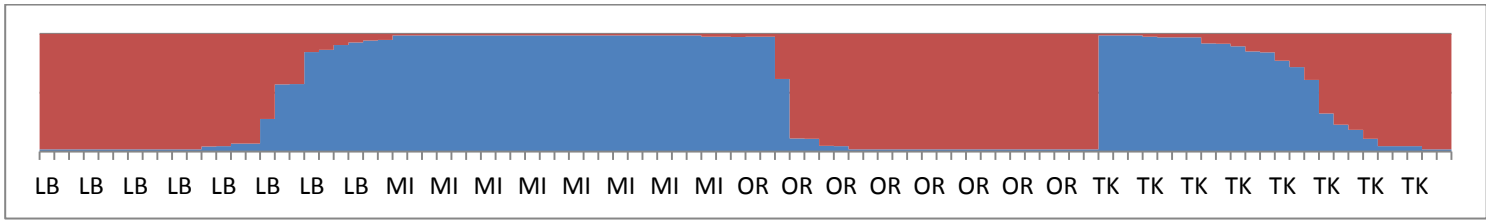
Plate 03 continued

Figure A.8. STRUCTURE output for Plate 03 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1455 neutral loci and 5 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a different color (red, blue, and purple). LB=Lower Bute Inlet, MI=Mittlenatch, OR=Oregon, TK=Top Knot. OR and TK correspond to 'outside' locations, while LB and MI correspond to 'inside' locations.

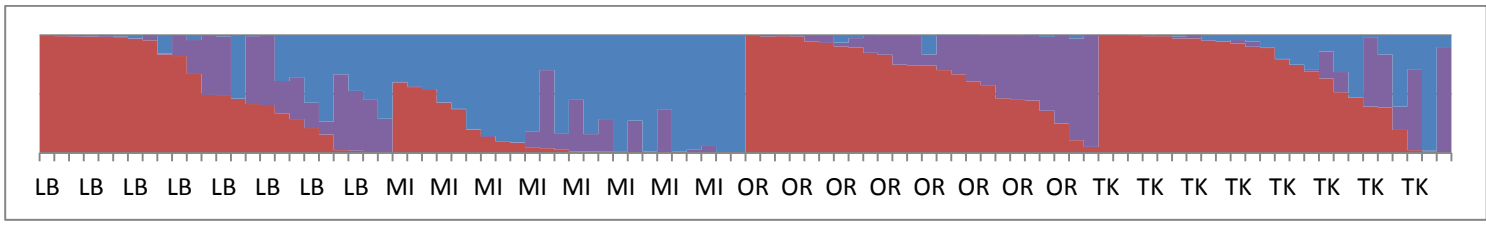
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

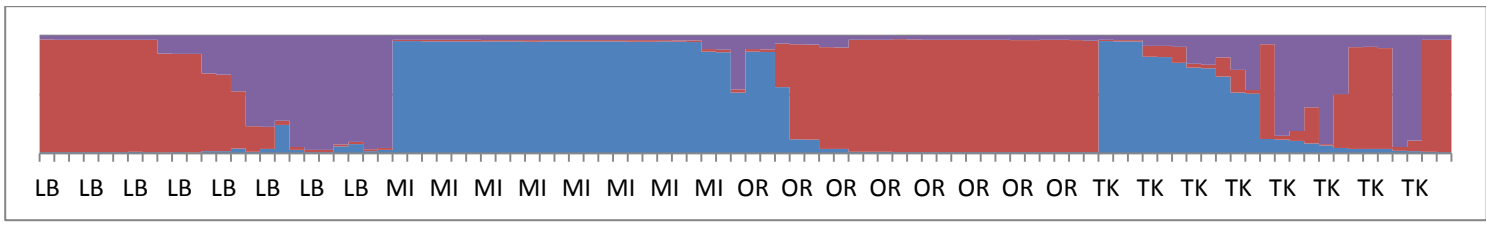


Plate 03 continued

Figure A.9. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. LB=Lower Bute Inlet, MI=Mittlenatch, OR=Oregon, TK=Top Knot.

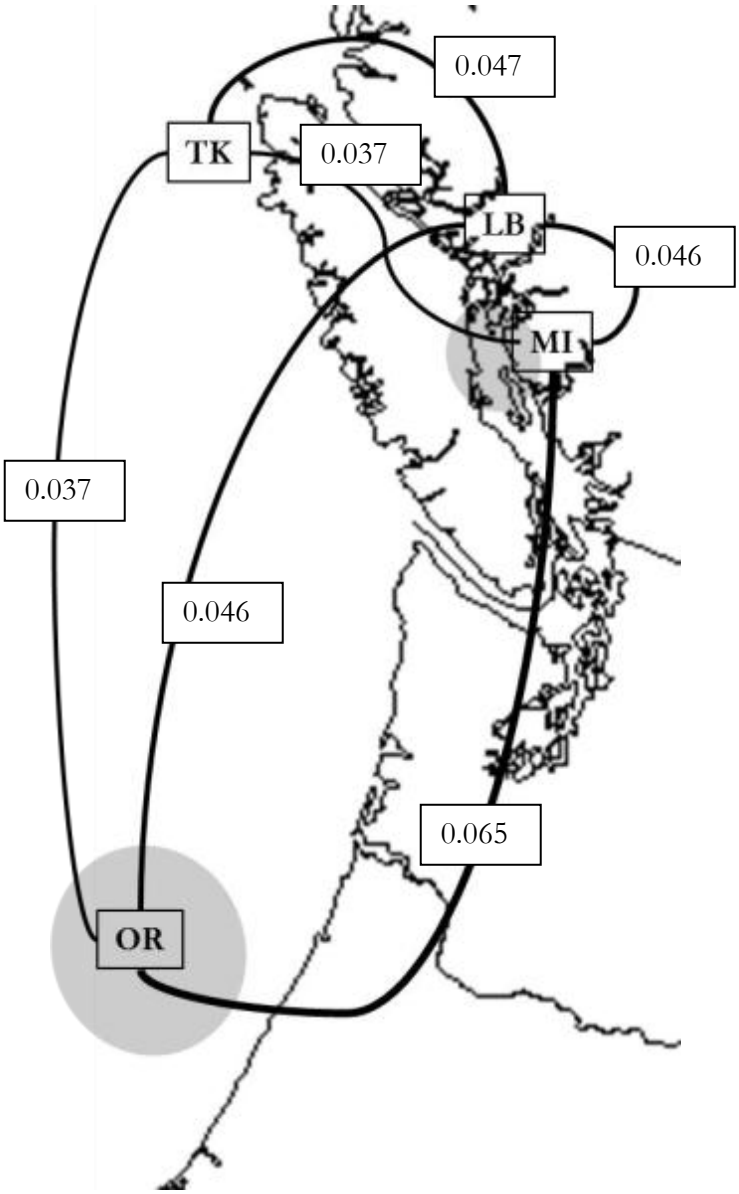


Plate 04 (AK and CJ):

Table A.10. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. AK=Alaska, CJ=Cape St. James.

Location	n	Neutral dataset		
		#loc_P	PLP	Hj (S.E.)
AK	35	1347	85.5	0.286 (0.004)
CJ	43	1081	68.6	0.228 (0.004)

Pairwise F_{ST} value in yelloweye rockfish for neutral dataset: 0.1254

Table A.11. STRUCTURE expectations based on neutral loci.

Dataset	Expected genetic clusters	Rationale
Neutral	AK-CJ	Outside sites cluster together

Results-

The clustering pattern is not consistent with expectations. Two clusters are observed.

Figure A.10. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for the neutral dataset in yelloweye rockfish.

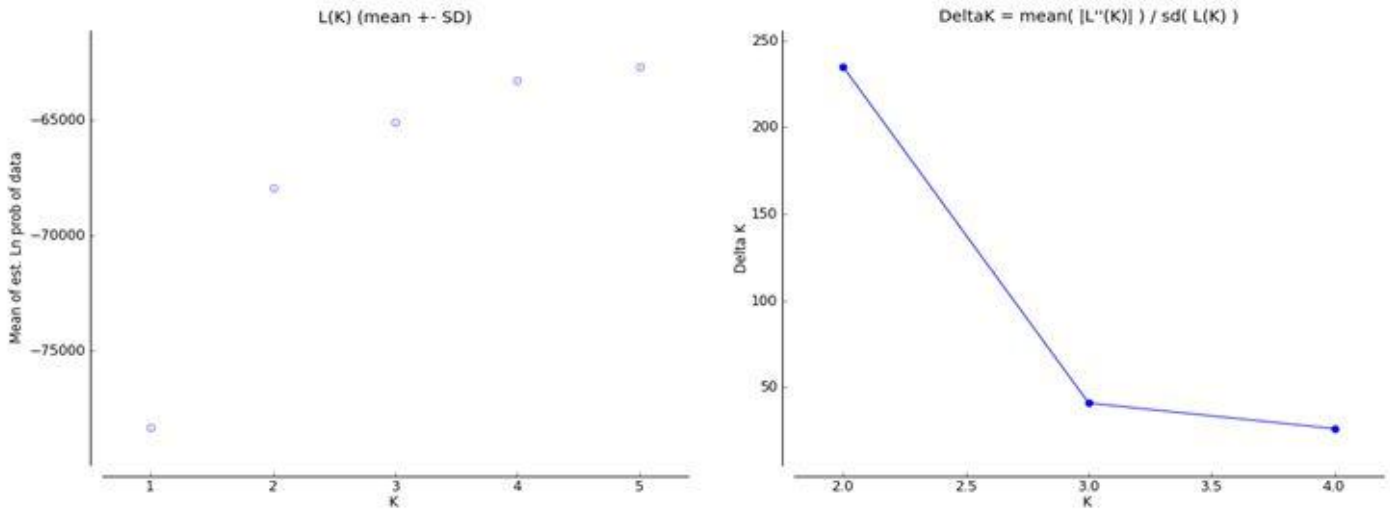
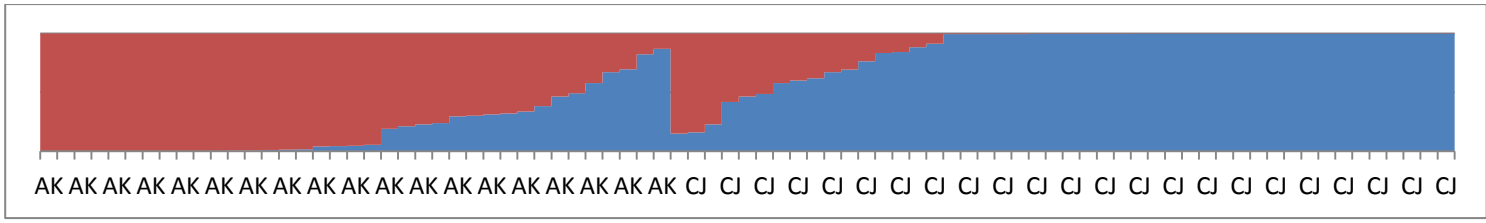


Plate 04 continued

Figure A.11. STRUCTURE output for Plate 04 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1576 neutral AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a different color (red, blue, and purple). AK=Alaska, CJ=Cape St. James.

K=2, neutral



K=3, neutral

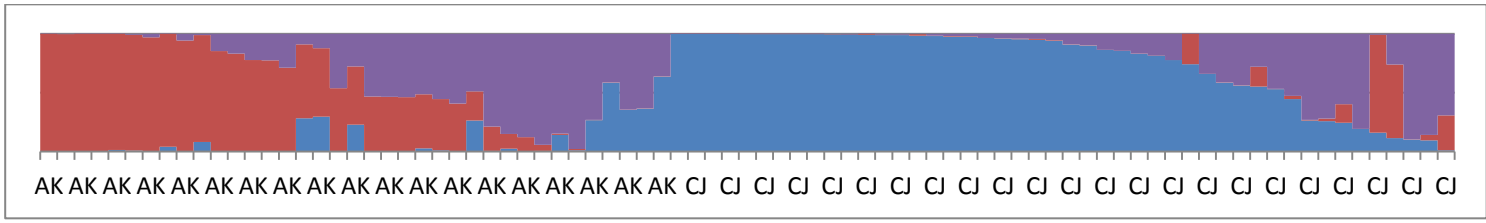


Figure A.12. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. AK=Alaska, CJ=Cape St. James.

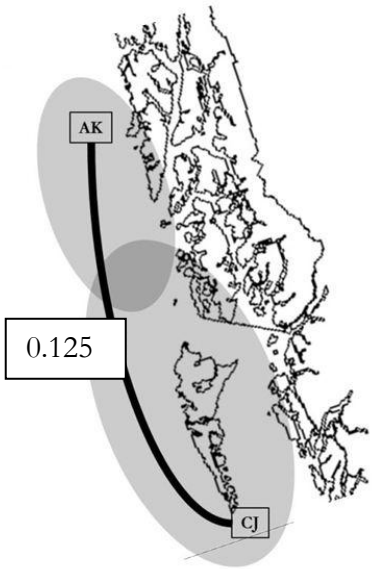


Plate 05 (BS, TA, TK, TR):

Table A.12. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. BS=Bowie Seamount, TA=Tasu, TK=Top Knot, TR=Triangle. Outside sites are **bolded**.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
BS	24	962	76.3	0.263 (0.005)	3	100	0.179 (0.040)
TA	24	841	66.7	0.221 (0.005)	2	66.7	0.161 (0.044)
TK	24	1029	81.6	0.274 (0.005)	3	100	0.367 (0.081)
TR	24	900	71.4	0.240 (0.005)	2	66.7	0.181 (0.091)

Table A.13. Pairwise F_{ST} values in yelloweye rockfish calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.12.

	BS	TA	TK	TR
BS	---	0.7953	0.2515	0.7784
TA	0.1923	---	0.5649	0.0000
TK	0.0982	0.0720	---	0.5290
TR	0.2116	0.0282	0.0728	---

Table A.14. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	All assign to same cluster	outside vs. inside clusters
Outlier	BS vs. TA vs. TK-TR	clusters associated with Ecosections

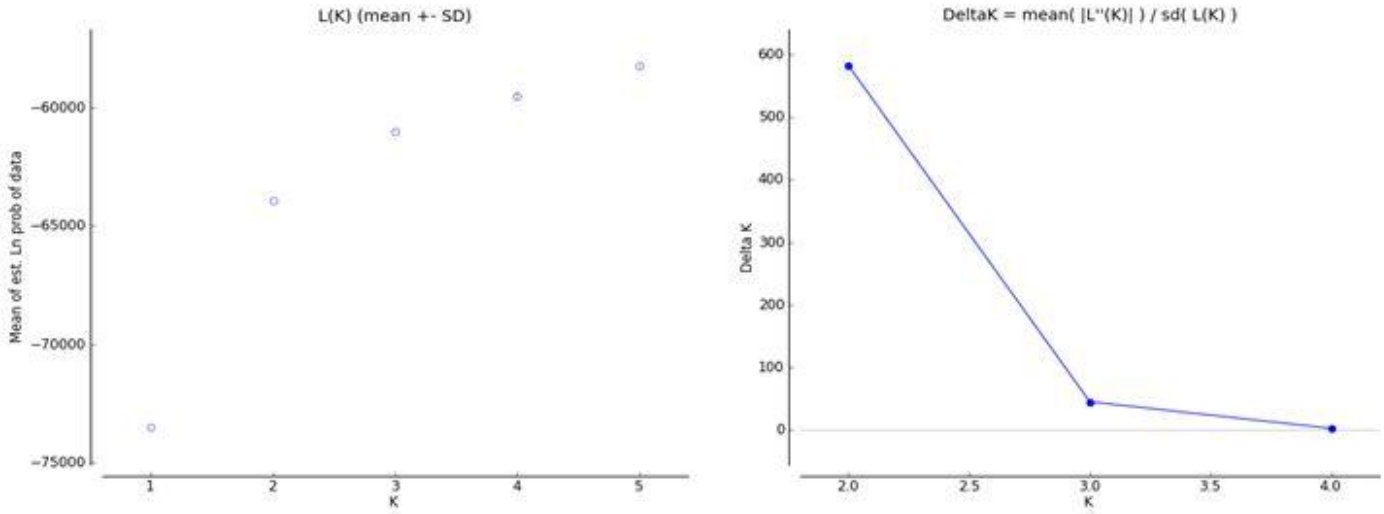
Results-

Overall, three distinct genetic clusters were identified. However, there are varying proportions of individuals within each sampling location that assigned to one of the distinct clusters. Most BS individuals assigned to a unique BS cluster. Additionally, most TA and TR individuals assigned to a second cluster. The third cluster contained individuals with high assignment from all four sampling locations. The lack of high assignment in the outlier dataset for the *K* of 3 is likely due to the limited number of loci (N=3) used.

Plate 05 continued

Figure A.13. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral dataset



Outlier dataset

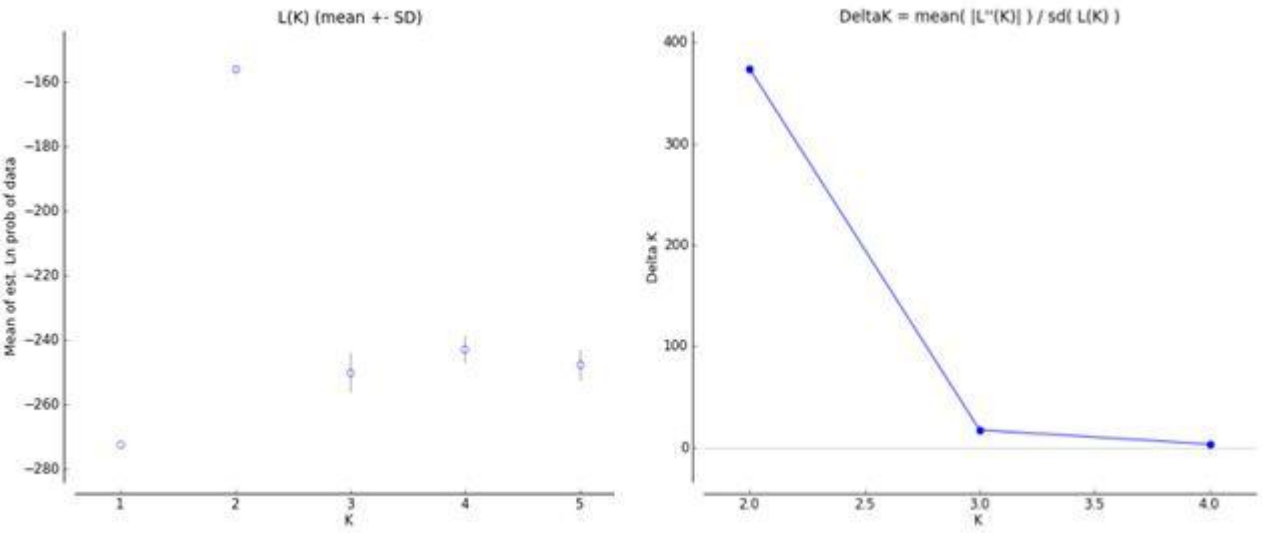
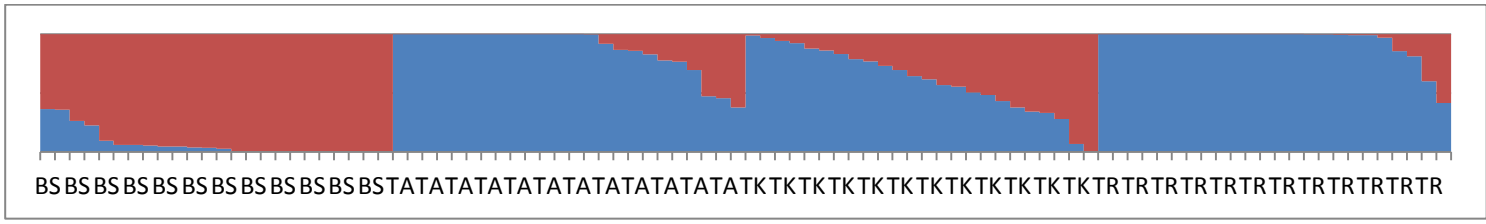


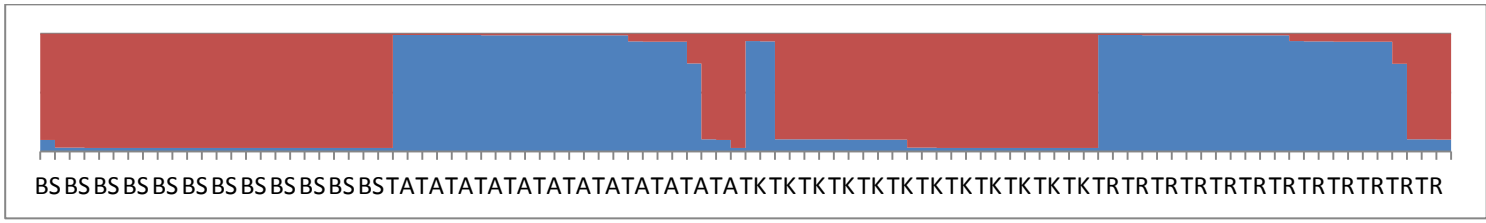
Plate 05 continued

Figure A.14. STRUCTURE output for Plate 05 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1261 neutral loci and 3 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a different color (red, blue, and purple). BS=Bowie Seamount, TA=Tasu, TK=Top Knot, TR=Triangle.

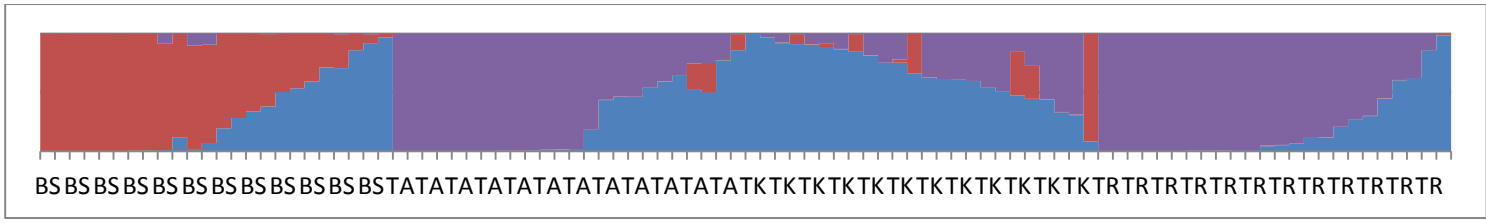
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

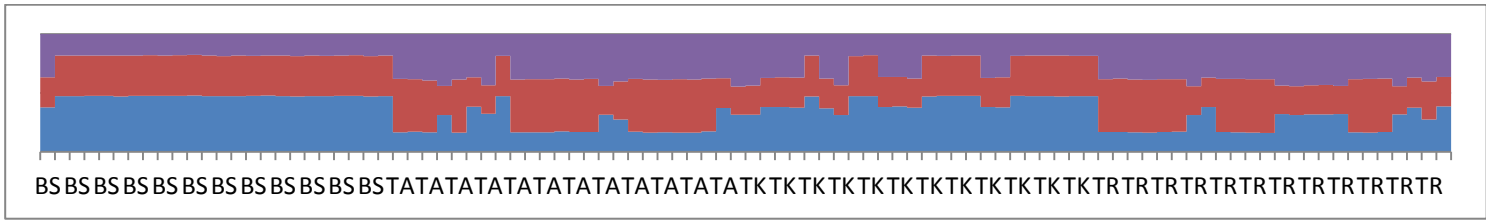


Plate 05 continued

Figure A.15. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. BS=Bowie Seamount, TA=Tasu, TK=Top Knot, TR=Triangle.

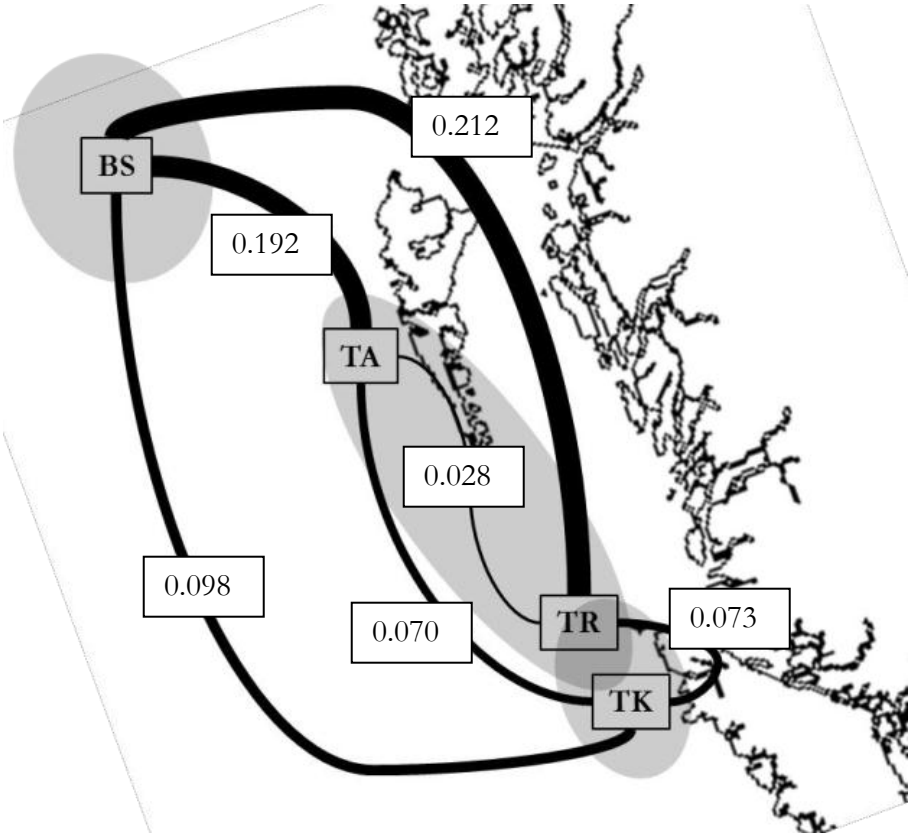


Plate 06 (CI, LB, OR, TI):

Table A.15. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. CI=Calvert Island, LB=Lower Bute Inlet, OR=Oregon, TI= Texada Island. Outside sites are **bolded**.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
CI	24	784	84.5	0.256 (0.005)	12	100	0.375 (0.036)
LB	23	766	82.5	0.264 (0.005)	11	91.7	0.395 (0.036)
OR	23	643	69.3	0.263 (0.006)	0	0	0.013 (0.007)
TI	24	641	69.1	0.235 (0.005)	10	83.3	0.281 (0.047)

Table A.16. Pairwise F_{ST} values in yelloweye rockfish calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.15.

	CI	LB	OR	TI
CI	---	0.2702	0.7125	0.1211
LB	0.0694	---	0.4248	0.3850
OR	0.1016	0.0655	---	0.8174
TI	0.0338	0.0440	0.1140	---

Table A.17. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	OR-CI vs. LB-TI	outside vs. inside clusters
Outlier	Four clusters	clusters associated with Ecosections

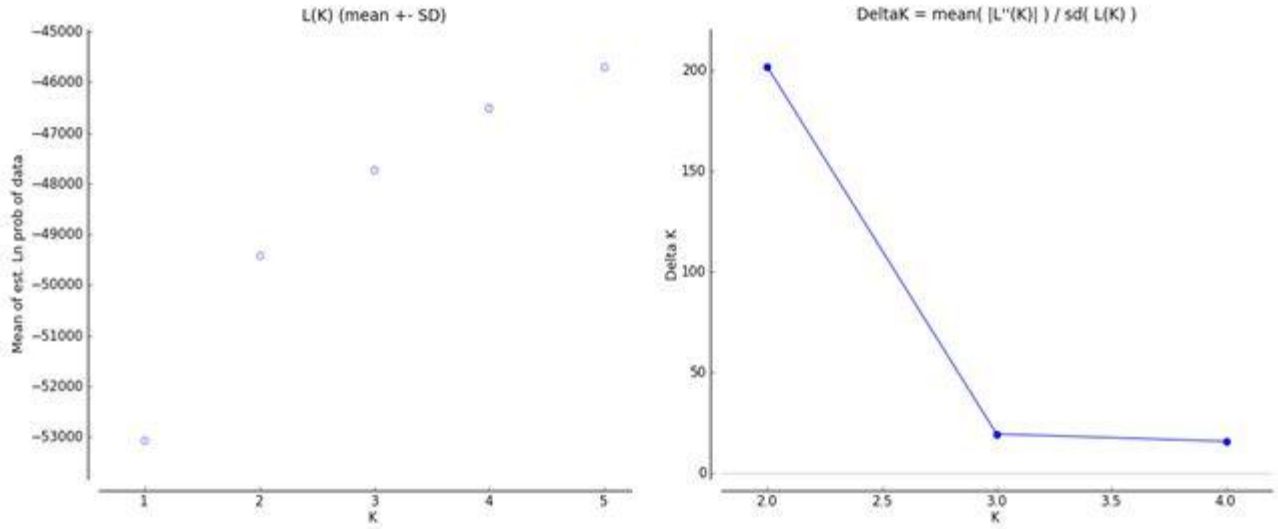
Results-

The neutral and outlier expectations were not met. A *K* of 3 shows the majority of CI and TI individuals assigning to same cluster, OR individuals showing high assignment to a separate cluster and LB individuals showing mixed assignment to the OR cluster and a third cluster. Overall, CI and TI individuals cluster together, OR individuals cluster separately, and LB individuals have mixed assignment to two clusters.

Plate 06 continued

Figure A.16. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral dataset



Outlier dataset

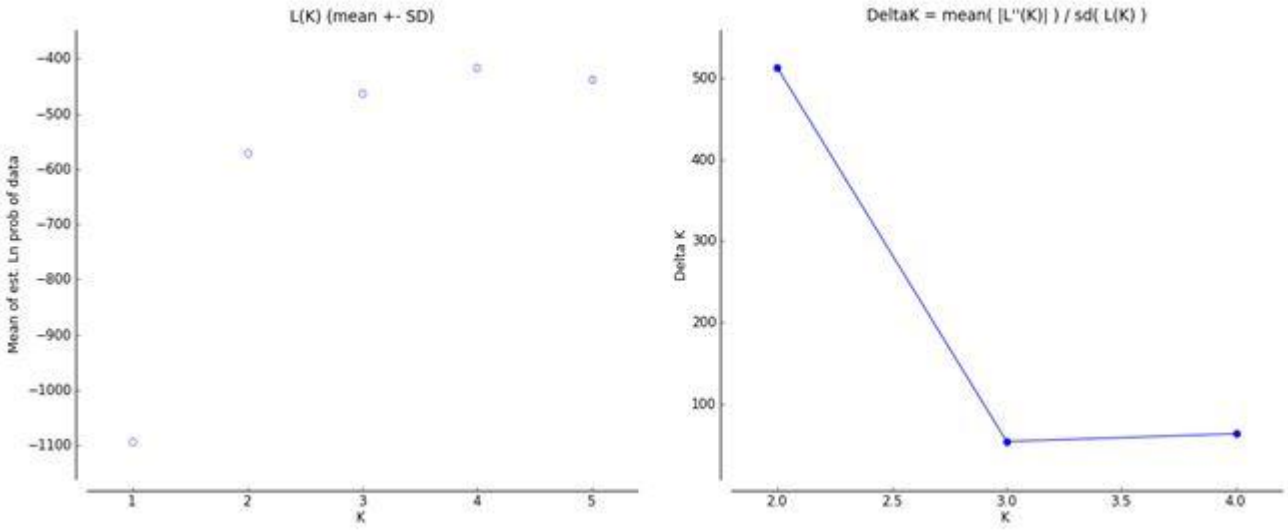
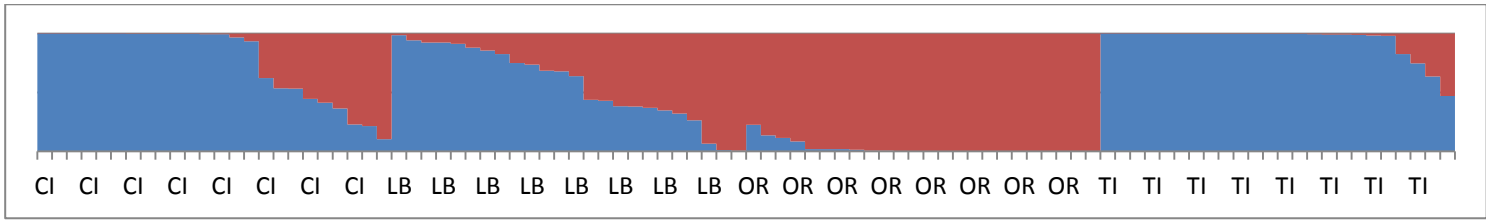


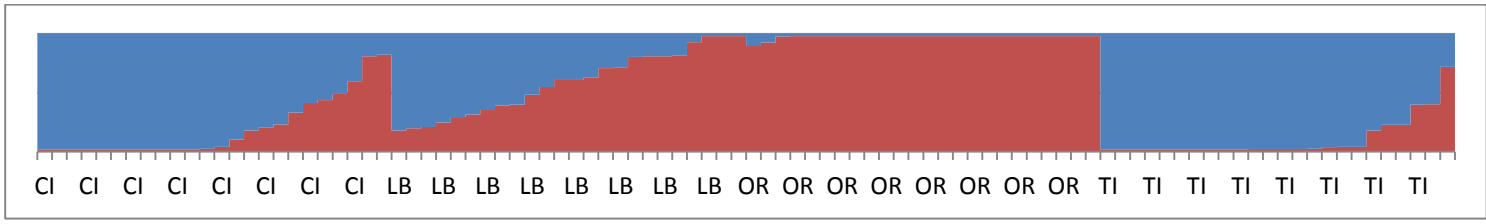
Plate 06 continued

Figure A.17. STRUCTURE output for Plate 06 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 928 neutral loci and 12 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a different color (red, blue, and purple). CI=Calvert Island, LB=Lower Bute Inlet, OR=Oregon, TI= Texada Island. OR and CI correspond to 'outside' locations, while LB and TI correspond to 'inside' locations.

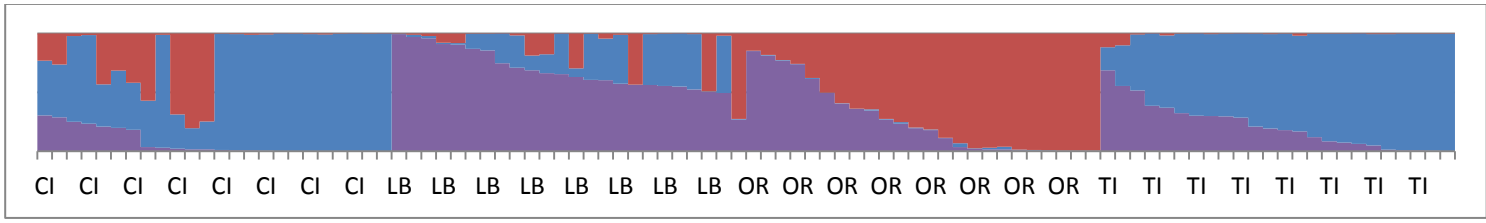
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

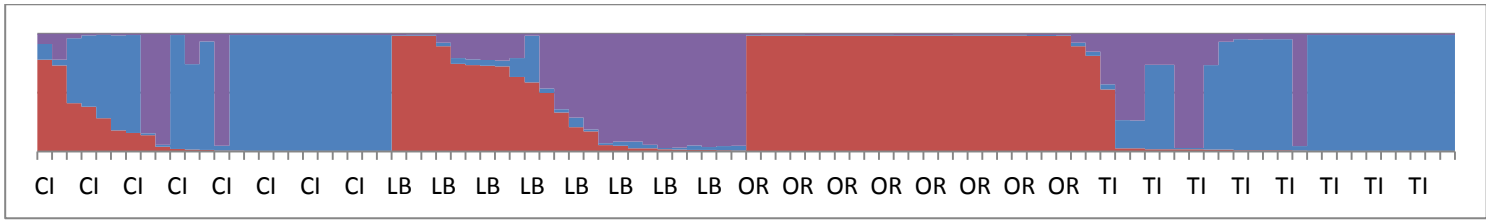


Plate 06 continued

Figure A.18. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. CI=Calvert Island, LB=Lower Bute Inlet, OR=Oregon, TI= Texada Island.

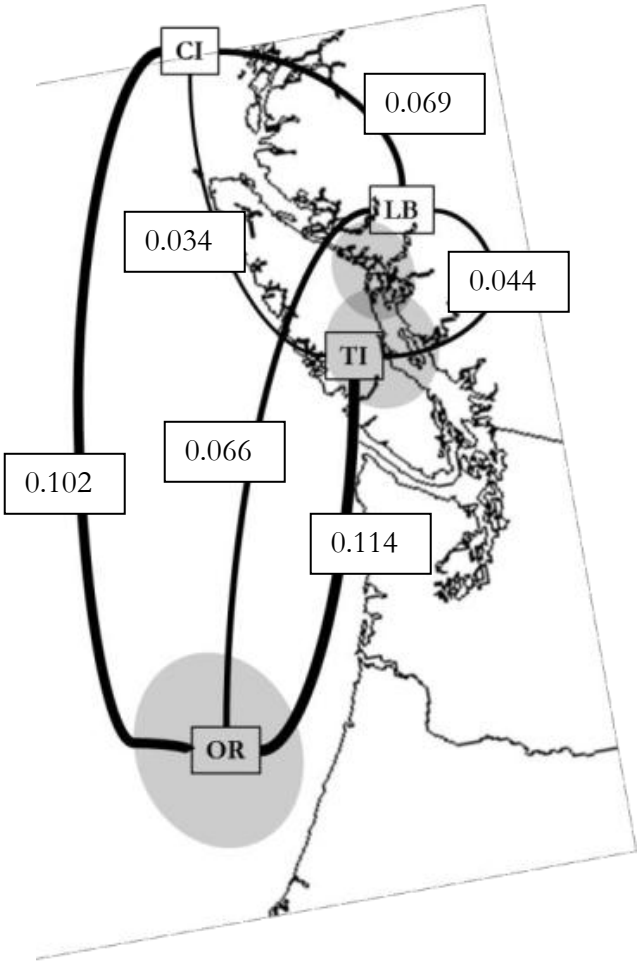


Plate 07 (ES, GC, TR, WA):

Table A.18. Number (#loc_P) and proportion of polymorphic loci (PLP) and Nei's gene diversity (Hj) calculated in AFLP-SURV in yelloweye rockfish. ES=Esperanza, GC=Gordon Channel, TR=Triangle, WA=Washington.

Location	n	Neutral dataset			Outlier dataset		
		#loc_P	PLP	Hj (S.E.)	#loc_P	PLP	Hj (S.E.)
ES	21	974	77.8	0.251 (0.005)	2	20.0	0.066 (0.033)
GC	23	902	72.0	0.240 (0.005)	9	90.0	0.349 (0.049)
TR	22	1011	80.8	0.248 (0.004)	10	100	0.335 (0.047)
WA	22	942	75.2	0.247 (0.005)	6	60.0	0.224 (0.066)

Table A.19. Pairwise F_{ST} values in yelloweye rockfish calculated in AFLP-SURV. Neutral dataset is below diagonal, outlier dataset is above diagonal. Population codes are shown in Table A.18.

	ES	GC	TR	WA
ES	---	0.6932	0.7151	0.3395
GC	0.1404	---	0.0000	0.4743
TR	0.1286	0.0416	---	0.5147
WA	0.0628	0.0657	0.0860	---

Table A.20. STRUCTURE expectations based on neutral loci and outlier loci.

Dataset	Expected genetic clusters	Rationale
Neutral	GC vs. ES-TR-WA	GC straddles inside-outside divide, other sites are all outside
Outlier	ES-WA-TR vs. GC; or ES-TR vs. GC vs. WA cluster.	clusters associated with Ecosections

Results-

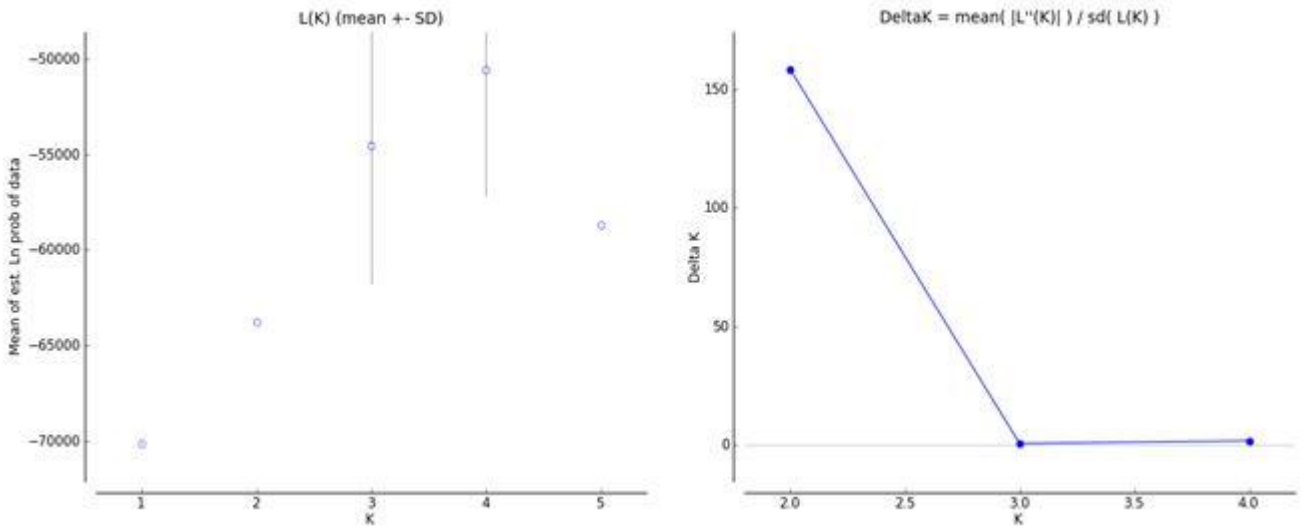
The clustering patterns do not meet expectations. Neutral loci show high assignment of ES individuals to a unique cluster. GC and TR individuals mostly assign to a second cluster, and WA individuals to a third cluster. However, GC, TR, and WA individuals show overall lower assignment to

Plate 07 continued

their respective clusters. Outlier loci further bolster the unique ES and GC/TR clusters, while WA individuals show mixed assignment.

Figure A.19. Log probabilities for each *K*-value and log deltaK value calculated by STRUCTURE for both the neutral and outlier datasets in yelloweye rockfish.

Neutral loci



Outlier loci

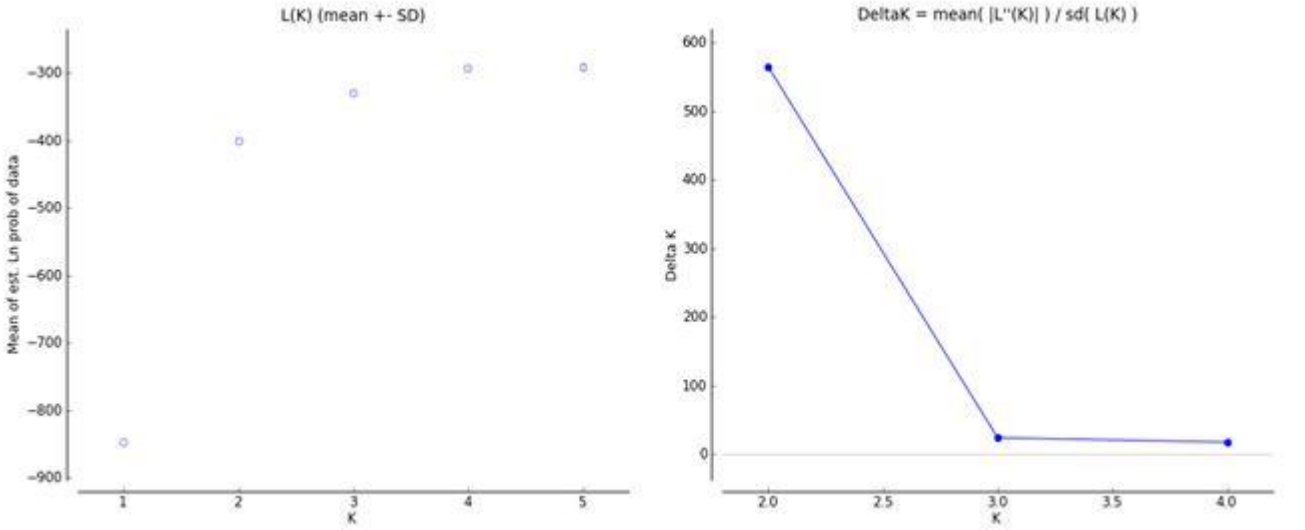
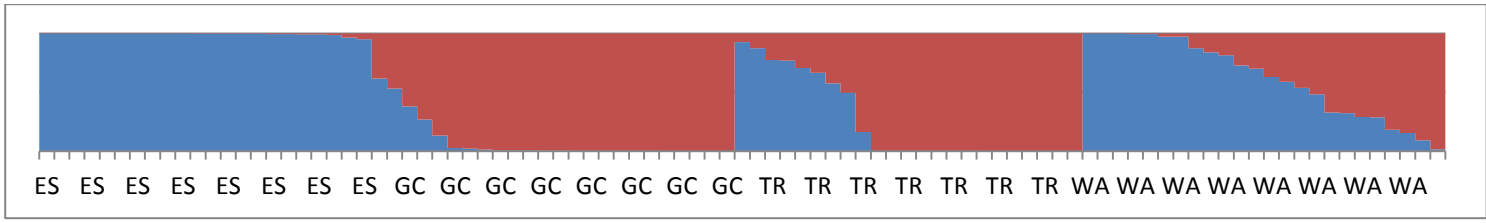


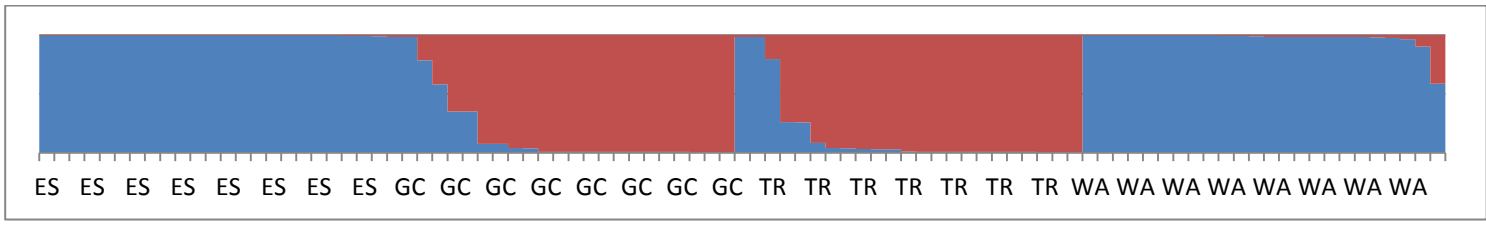
Plate 07 continued

Figure A.20. STRUCTURE output for Plate 07 in yelloweye rockfish. Each fish is represented by a thin vertical line, the height of which (maximum = 1.0) represents the proportion of its genome as assayed by 1252 neutral loci and 10 outlier AFLP loci represented by one of two (*K*=2) or three (*K*=3) genetic groups, each of which is represented by a different color (red, blue, and purple). ES=Esperanza, GC=Gordon Channel, TR=Triangle, WA=Washington.

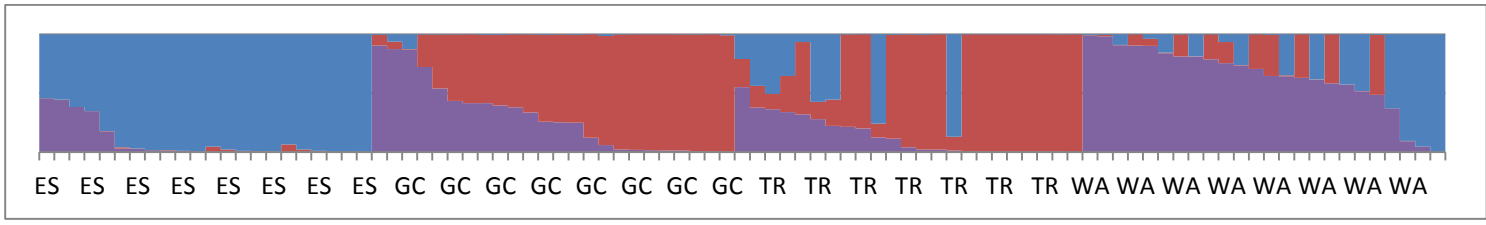
K=2, neutral



K=2, outlier



K=3, neutral



K=3, outlier

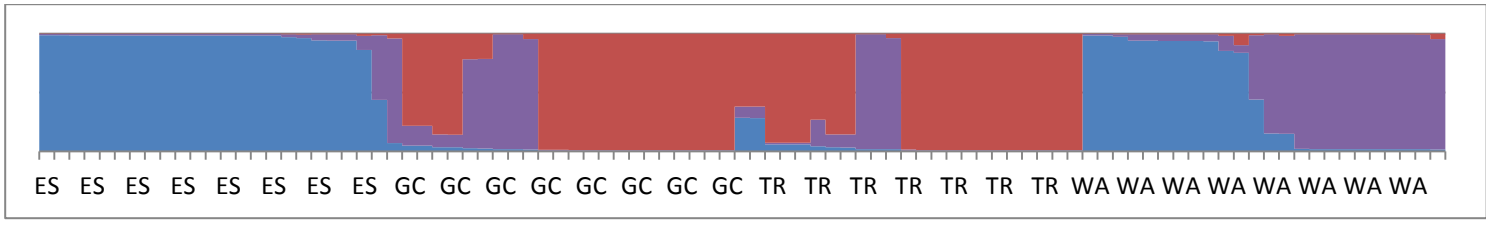


Plate 07 continued

Figure A.21. Pairwise F_{ST} values between the different localities sampled for yelloweye rockfish based on neutral loci are represented by the lines connected the different sample location codes. The thicker the line, the larger the pairwise F_{ST} value is. The grey ovals represent the genetic clusters identified in the STRUCTURE analyses. ES=Esperanza, GC=Gordon Channel, TR=Triangle, WA=Washington.