Journal of Heredity 2014:105(6):835–845
doi:10.1093/jhered/esu063
Advance Access publication October 6, 2014

Temporally Isolated Lineages of Pink Salmon Reveal Unique Signatures of Selection on Distinct Pools of Standing Genetic Variation

Morten T. Limborg, Ryan K. Waples, James E. Seeb, and Lisa W. Seeb

From the School of Aquatic and Fishery Sciences, University of Washington, 1122 NE Boat Street, Box 355020, Seattle, WA 98195 (Limborg, Waples, Seeb, Seeb); and the National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark (Limborg).

Address correspondence to M. T. Limborg at the addresses above, or e-mail: moli@uw.edu.

Data deposited at Dryad: doi:10.5061/dryad.pp43m

Abstract

A species' genetic diversity bears the marks of evolutionary processes that have occurred throughout its history. However, robust detection of selection in wild populations is difficult and often impeded by lack of replicate tests. Here, we investigate selection in pink salmon (*Oncorhynchus gorbuscha*) using genome scans coupled with inference from a haploid-assisted linkage map. Pink salmon have a strict 2-year semelparous life history which has resulted in temporally isolated (allochronic) lineages that remain sympatric through sharing of spawning habitats in alternate years. The lineages differ in a range of adaptive traits, suggesting different genetic backgrounds. We used genotyping by sequencing of haploids to generate a high-density linkage map with 7035 loci and screened an existing panel of 8036 loci for signatures of selection. The linkage map enabled identification of novel genomic regions displaying signatures of parallel selection shared between lineages. Furthermore, 24 loci demonstrated divergent selection and differences in genetic diversity between lineages, suggesting that adaptation in the 2 lineages has arisen from different pools of standing genetic variation. Findings have implications for understanding asynchronous population abundances as well as predicting future ecosystem impacts from lineage-specific responses to climate change.

Subject areas: *Molecular adaptation and selection*

Key words: *climate change, haploid, linkage mapping, life history, parallel evolution*

Introduction

Understanding the origin of adaptive genetic variation in wild populations is a fundamental goal in evolutionary biology. Genomes of wild organisms have been shaped by selection throughout their evolutionary history. Sometimes, historical separation has split species into genetically diverged lineages that are often derived from survival in different glacial refugia. Populations from lineages with different evolutionary backgrounds are therefore expected to possess different pools of standing genetic variation and hence differ in their potential to adapt to similar environmental conditions (Barrett and Schluter 2008). Despite this expectation, we only have a narrow idea about the general importance of standing genetic variation from a few case studies in model species (reviewed in Barrett and Schluter 2008).

An ideal, but rare, setting for studying the effect of standing genetic variation in the wild occurs when replicate gene pools (e.g. distinct lineages) independently adapt to identical environments. Divergence driven by isolation over time rather than allopatric separation by geography has been coined allochronic speciation (Alexander and Bigelow 1960). Cases are mainly known from periodical insects characterized by semelparity and fixed longevity (Heliövaara et al. 1994) including aphids (Abbot and Withgott 2004) and moths (Santos et al. 2007), but remain extremely rare in other taxa. When previously isolated allochronic lineages share post-glacial colonization histories of the same environmental gradient, we might expect different evolutionary outcomes, creating a common-garden experiment in the wild. These situations offer exciting evolutionary insights; parallel signatures of selection greatly increase evidence for local adaptation by both lineages, while lineage-specific adaptations

may reflect unique adaptive potentials reflecting different backgrounds of standing genetic variation.

Pink salmon (*Oncorhynchus gorbuscha*) represents an interesting species for studying the importance of standing genetic variation in 2 allochronic sympatric lineages that are spatially overlapping but reproductively isolated. Pink salmon exhibit a unique 2-year, semelparous life history with 2 diverged lineages characterized by a preglacial origin (Zhivotovsky et al. 1994; Brykov et al. 1996; Churikov and Gharrett 2002). Rare, if any, gene flow occurs between these even- and odd-year lineages; yet lineages share spawning and nearshore habitats in alternate years in many locations (Gilbert 1912; Aspinwall 1974; Heard 1991). During refugial isolation, lineages likely underwent independent drift and adaptations to varying environmental conditions resulting in differing genetic backgrounds and evolutionary legacies. Indeed, the 2 lineages differ in a range of biological traits throughout their North American distribution including geographic distribution limits and adaptation to local temperature regimes (Bams 1976; Beacham and Murray 1988; Heard 1991; Churikov and Gharrett 2002). However, despite recent evidence for locally adapted populations (Kovach et al. 2012; Gharrett et al. 2013), we know little about the genomic distribution (architecture) of adaptive genetic variation within versus between lineages of pink salmon.

Inference about the genomic architecture of adaptive traits in non–model organisms has been greatly facilitated by the increased accessibility of genome-wide data sets. One intriguing, and increasingly popular, way to study the genomic architecture of adaptively important traits is to consider genome scans along linkage maps (Gagnaire et al. 2013a; Tsumura et al. 2012; Hemmer-Hansen et al. 2013, among many). Briefly, genome scans simultaneously analyze multiple populations to detect individual loci (outliers) exhibiting increased differentiation (e.g. F_{ST}) compared with the level expected for loci only affected by neutral processes such as genetic drift and gene-flow (Lewontin and Krakauer 1973). While interpretation of genome scans has suffered from nonnegligible rates of false positive results (Narum and Hess 2011; De Mita et al. 2013), the ability to map loci to specific regions adds a genomic perspective greatly increasing the value of genome scans. Increased support can then be given to outlier loci mapping to the same regions and consideration of the entire map increase our understanding of the genomic architecture underlying adaptively important variation.

The overarching objective of this study is to provide the first description of the genomic architecture underlying adaptive differences within and between lineages of pink salmon. First, we make use of a unique mapping resource, gynogenetic haploids, for generating a high-density linkage map including 7035 loci. We then use this map to evaluate population data from Seeb et al. (2014) who characterized over 8000 loci for 3 pairs of spatially overlapping even- and odd-year pink salmon populations (Figure 1a). We illustrate how the linkage map facilitates identification of genomic regions with novel signatures of parallel selection between lineages, parallel patterns that were not detected with genome scans alone. Second, we studied potential differences in genetic background by looking for signatures of divergent selection between lineages within each of the population pairs. We identified 24 loci showing signatures of divergent selection as well as varying levels of diversity between lineages. These loci mapped to multiple genomic regions and likely represent different genetic backgrounds and may include genes related to adaptive differences between lineages. We discuss potential ecosystem consequences of lineage-specific responses to future climate change based on these results.

Methods

Haploid Linkage Map

We collected eggs and sperm from 2 male and 2 female pink salmon from the odd-year population at the Hoodsport Hatchery, Hoodsport, Washington (USA) to produce 2 unrelated single pair matings. Fin clips for DNA analyses were taken from adults and stored in alcohol at room temperature.

Two haploid families (X01 and X05) were generated, and embryos harvested following University of Washington Institutional Animal Care and Use Committee protocol 4229-01. Embryo development was activated using UV-irradiated sperm to fertilize eggs following the methods described in Seeb and Seeb (1986). Families were incubated individually at the University of Washington Hatchery for 50 days and removed to alcohol just prior to hatch. DNA was extracted from parents ($n = 4$) and offspring ($n = 192$) using DNeasy-96 kits (Qiagen, Valencia, CA), and concentrations were subsequently standardized using the Quant-iT PicoGreen dsDNA Assay (Life Technologies, Carlsbad, CA) on a Victor D plate reader (PerkinElmer, Waltham, MA).

Adults and 96 embryos from each family were initially genotyped for 19 single nucleotide polymorphism (SNP) loci with 5'nuclease assays (Seeb et al. 2009) to confirm that progeny were haploid. Only 3 embryos expressed paternal alleles; those 3 were excluded from further analyses.

Restriction-site associated DNA (RAD) sequencing libraries were prepared and sequenced following existing protocols (Etter et al. 2011; Everett et al. 2012). Genomic DNA was digested using the restriction enzyme *SbfI,* and each individual was barcoded with 6 bp long adaptors differing by at least 2 nucleotides following Miller et al. (2012). Libraries were assessed for DNA quality and sequenced on an Illumina HiSeq2000 sequencer producing 101 bp single-end reads. After sequencing, raw unfiltered sequences for each individual in both families were deposited in the NCBI short read archive (Accession number: SRP035433).

Raw sequence data were quality filtered and used for detection of polymorphisms with the software package *Stacks* v.0.9996 (Catchen et al. 2011). First, we used the *Stacks* program *process_radtags* to trim the terminal nucleotide (which suffered consistently poor quality among all sequencing lanes), de-multiplex individuals, and remove low-quality reads. Remaining 94 bp reads were assembled into matching stacks and used to detect polymorphic loci with *ustacks*. We enabled the deleveraging and removal algorithms to discard highly repetitive and over merged "lumberjack" stacks likely to represent paralogous sequence assembly (Catchen et al. 2011). We built family-specific catalogs of variation using *cstacks*. Offspring from each cross

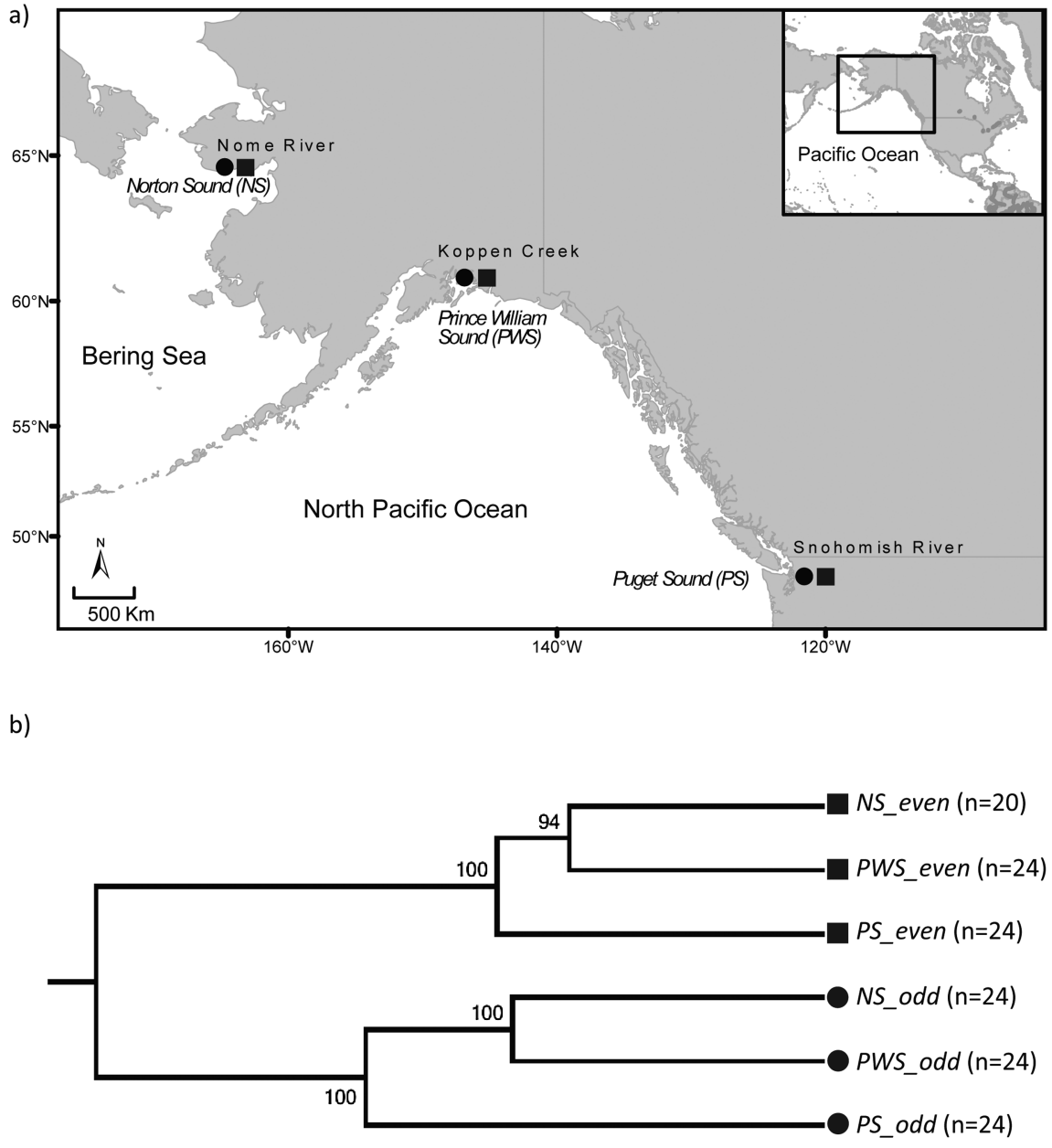

Figure 1. Population information. (**a**) Map showing sampling locations. (**b**) UPGMA tree based on 7820 neutrally behaving loci (marker set i) using Nei's D_A genetic distance. Percent bootstrap support is shown for each branch. Population abbreviations and sample sizes (*n*) are shown for even (squares) and odd (circles) lineage populations.

were matched to their respective maternal catalogs. Within each family, we discarded loci scored in less than 80% of the offspring. In fulfillment of data archiving guidelines (Baker 2013), we have deposited the primary genotype data underlying these analyses with Dryad.

We constructed a joint *Stacks* catalog from the 2 female parents in order to establish the correspondence of loci between the 2 separate families. Parental RAD tags merged in this process were determined to be the same. Non-matching RAD tags were kept as family-specific loci. RAD tags with conflicting matching patterns were excluded.

We used the R/qtl package (Broman et al. 2003) developed for R (R Core Team 2013) to construct linkage maps, using the Kosambi mapping function (Kosambi 1943), for each family. We omitted loci exhibiting significant segregation distortion ($\alpha = 0.05$) after Bonferroni correction (Rice 1989) including 56 loci from cross X01 and 43 loci from cross X05. Groups of markers with identical genotype data (i.e. markers not separated by recombination events in our crosses) were identified, and then all but 1 marker from each group was removed. The removed markers were subsequently added back to the final map at the same position as the retained marker from

that group. We constructed linkage groups using a minimum logarithm of odds (LOD) score of 6 and a maximum recombination frequency of 0.35. Markers were initially ordered using *orderMarkers* followed by visual inspection of recombination frequency and LOD score plots. When necessary, marker order was determined using the *ripple* and *switch.order* functions considering sliding windows of length 7–10 markers.

To construct a more robust consensus map, we further estimated independent linkage maps using MSTMap (Wu et al. 2008) applying the Kosambi map function and a grouping threshold of $P < 10^{-6}$. MSTMap uses a graph-based minimum spanning tree approach to linkage map construction (Wu et al. 2008). Finally, we generated a consensus female linkage map for pink salmon by merging the 4 equally weighted individual maps (Maps generated with both R/qtl and MSTMap for both families) using MergeMap (Wu et al. 2011).

All loci on the consensus map were annotated following the procedure described in Seeb et al. (2014). Twelve individuals were selected for paired-end sequencing to assemble contigs using CAP3 (Huang and Madan 1999). Loci from our map were then aligned to all contigs, and exact matches were used to assign annotation results derived from the longer contig sequences.

Lastly, we used *Bowtie2* 2.0.2 (Langmead and Salzberg 2012) to align RAD tag sequences (94 bp) for both alleles from the 8036 polymorphic loci presented in Seeb et al. (2014) against a reference of locus sequences containing both allelic haplotypes of the 7035 loci on the linkage map created in this study. We used the "end-to-end" option with a maximum number of reported alignments of 3. For loci occurring in both data sets, we expect the 2 alternate alleles in the population data set to return the 2 most significant alignments, whereas a third alignment allowed us to detect cases without a one-to-one correspondence. Loci from the population data returning a match to only a single locus on the map were considered identical and allowed placement of loci derived from Seeb et al. (2014) on our linkage groups.

Population Genomics

Downstream analyses were based on genotypic data and populations presented in detail in Seeb et al. (2014). Seeb et al. (2014) analyzed 140 individuals from 6 populations of pink salmon representing both even- and odd-year lineages in each of 3 rivers spanning the species' latitudinal distribution in North America (Figure 1a).

Many different methods for detecting genetic signatures of selection exist (reviewed in Nielsen 2005; Barrett and Hoekstra 2011). Methods applying estimates of genetic differentiation (e.g. F_{ST}) have been shown to outperform approaches based around the site frequency spectrum for population-based analyses (Thornton and Jensen 2007; De Mita et al. 2013). Here, we used a Bayesian differentiation-based method considering population-specific F_{ST} estimates as implemented in BayeScan 2.1 (Foll and Gaggiotti 2008) to identify outlier loci. While BayeScan may suffer reduced power to detect true outliers relative to other methods, it has repeatedly been shown to outperform these in terms of producing a low rate of false positive outliers (Narum and Hess 2011; De Mita et al. 2013). Here we opt for a conservative approach, considering only BayeScan for detecting outliers. While Seeb et al. (2014) also performed an environmental correlation-based outlier approach, the vast majority of outliers were detected by BayeScan. Thus, here we accumulate outlier support among multiple testing schemes (rather than among methods) allowing replicated inference about signatures of selection (cf., Zueva et al. 2014).

We performed a total of 6 different genome scans. First we considered the 3 genome scans presented in Seeb et al. (2014): a) all 6 populations, b) the 3 populations within the even-year lineage, and c) the 3 populations within the odd-year lineage (Table 1). We then performed 3 additional genome scans (d–f) to identify potential adaptive differences between the 2 lineages, 1 for each population pair sampled within 3 sampling sites (Table 1). We ran 50 000 iterations with other settings left as defaults in all individual tests. For each test we considered high F_{ST} outliers with a q value < 0.05 as candidates for divergent selection. The combined outlier status for each marker over all 6 genome scans was used to define 3 marker sets for subsequent analyses: i) neutral markers not exhibiting outlier behavior in any of the 6 genome scans; ii) preliminary candidates for parallel selection, defined as outliers within the even- and/or the odd-year

Table 1 Overview of the 6 genome scans including number of outliers detected in each test as well as number of outliers unique to that test. Overlap of outliers among tests is illustrated in Figure 2.

Genome scan	Populations included	No. of outliers	No. unique outliers
a)	All 6	164	101
b)	Norton Sound_even Prince William Sound_even Puget Sound_even	47	39
c)	Norton Sound_odd Prince William Sound_odd Puget Sound_odd	54	33
d)	Norton Sound_even Norton Sound_odd	16	7
e)	Prince William Sound_even Prince William Sound_odd	18	7
f)	Puget Sound_even Puget Sound_odd	11	6

lineage (tests b, c) while not exhibiting outlier status between lineages (tests d, e, f); and iii) candidates for adaptive divergence between lineages including outliers in either 1, 2, or all 3 pairwise genome scans between sympatric populations of the even- and odd-year lineages (tests d, e, f) and not confounded by simultaneous outlier status within lineages (tests b, c). We excluded ambiguous loci defined as markers only detected when analyzing all 6 populations or outliers confounded by detection both within and between lineages. Outliers from marker set (ii) that were detected in both lineages were considered strong candidates for parallel selection following Seeb et al. (2014). We further considered the linkage map to detect potential new signatures of parallel selection if outliers from marker set (ii) only detected in 1 lineage colocated with outliers specific to the other lineage. Likewise, markers showing repeated outlier patterns between lineages at all 3 sites were deemed strong candidates reflecting divergent selection between lineages. The genome-wide distribution of outlier loci reflecting selection within and between lineages was plotted on the linkage map.

Loci exhibiting outlier behavior within lineages (marker set ii) should have increased differentiation among populations within them (F_{SC}) compared with the other marker sets. Likewise, outliers between lineages (marker set iii) were expected to show increased differentiation between population pairs from different lineages (F_{CT}). We used this expectation to further evaluate outlier status by also considering locus-specific differentiation estimates following an alternative hierarchical analysis of molecular variance (AMOVA) model (Excoffier et al. 1992). We grouped populations within lineages and performed a hierarchical locus-by-locus AMOVA with 10 000 permutations in Arlequin 3.5 (Excoffier and Lischer 2010). This was done independently for the 3 marker sets defined above (i, ii, iii) for comparison with BayeScan results.

We generated an UPGMA tree based on the neutral marker set (i) using Nei's D_A genetic distance (Nei et al. 1983) and 10 000 bootstraps in POPTREE2 (Takezaki et al. 2010) to illustrate the evolutionary relationship among populations and lineages.

Lastly, we illustrate potential differences in standing genetic variation for all 3 marker sets among populations and lineages. For each marker set we calculated average observed heterozygosity (H_O) within all 6 populations using Genepop 4.2 (Rousset 2008). We then performed a standard analysis of variance (ANOVA) to test for differences in heterozygosity. If significant differences occurred, pairwise comparisons among all possible population pairs were performed using the Tukey test in R (R Core Team 2013). The large number of replications (i.e. number of markers, $n = 7820$; see results) in the neutral marker set may lead to inflated statistical power resulting in statistically significant, but biologically non-significant, differences. Hence, we performed significance testing for the neutral marker set by randomly sampling 80 markers without replacement. Resampling 80 loci resulted in sample sizes (i.e. replications) comparable to the number of markers in the 2 outlier marker sets. Significant results should therefore be more easily compared among all 3 marker sets. This procedure was repeated 1000 times, and comparisons with more than 95% significant replicates were considered significantly different.

Results

Haploid Linkage Map

We created a linkage map allowing unprecedented resolution for identifying genomic regions of interest in pink salmon (cf., Lindner et al. 2000; Matsuoka et al. 2004). We retained 45–58% of raw reads from 6 lanes of sequencing for SNP detection and genotyping (Supplementary Figure S1) and created a female linkage map for each haploid family. Merging maps from the 2 families produced a 3353 cM consensus map consisting of 7035 markers and 26 linkage groups with an average spacing of 2 cM between the 1658 unique positions (Supplementary Table S1). This number of linkage groups is within the expected range of chromosome pairs in odd-year pink salmon in Washington populations which varies from 26–27 because of a Robertsonian translocation (Phillips and Kapuscinski 1987). We are not able to establish the true chromosome number in each female; however, the 26 linkage groups obtained in our consensus map fit the dominant karyotype (Phillips and Kapuscinski 1987) and should comprise all chromosome arms in the pink salmon genome. We successfully annotated 23% of all markers on the consensus linkage map (Supplementary Table S1). Of the 8036 markers reported in Seeb et al. (2014), 2881 matched a locus on our map. Population analyses were based on the full set of 8036 markers from Seeb et al. (2014), while the map was used to highlight regions of interest based on those 2881 markers. That only 2881 of 8036 loci were mapped is likely a reflection that the map originated from markers polymorphic in 2 females from the Puget Sound region only.

Population Genomics

We detected a total of 216 unique outlier loci (2.7%) combined over 6 distinct genome scans (Table 1; Figure 2). The tree based on allele frequencies from the remaining 7820 neutrally-behaving markers showed a deep historical split between lineages (Figure 1b) and more recent population divergence within lineages (as seen in Zhivotovsky et al. 1994; Churikov and Gharrett 2002; Seeb et al. 2014).

Latitudinal genome scans revealed 91 candidates for selection within lineages (tests b, c) and 38 outlier loci between lineages within sites (tests d, e, f) with 14 ambiguous loci present in both categories (Figure 2). The genome scan based on all 6 populations (test a) revealed 101 unique outliers (Table 1). In total, ambiguous outliers numbered 115 and were not considered further.

Twenty-seven of the 77 candidates for selection within lineages mapped to 12 linkage groups (Supplementary Table S2 and Supplementary Figure S2). Five were detected in both lineages (Figure 2) and were among those previously considered for strong signatures of parallel selection (Seeb et al. 2014). After plotting outliers on the linkage map, we

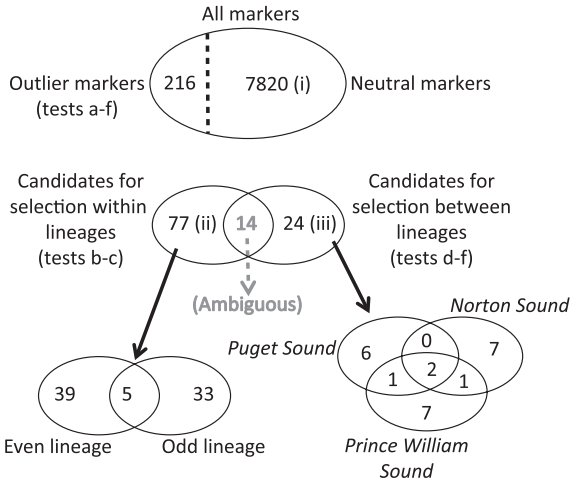

Figure 2. Overview of outlier loci detected from 6 genome scan tests and defined marker sets. Numbers followed by a parenthesis represent one of three different marker sets (i–iii) as defined in the text. The top oval shows numbers of all unique outlier loci detected over all 6 tests and number of putatively neutral markers. The middle shows overlap between markers detected as candidates for selection within and/or between lineages. The bottom shows numbers of outliers detected in one or more genome scans including candidates for selection within lineages (left), and candidates for divergent selection between lineages within sampling locations (right).

further identified a genomic location on LG 6 at 62.51 cM showing a signature of parallel selection evidenced by the joint mapping of 3 outliers within the even-year lineage and 2 within the odd-year lineage (Figure 3). Interestingly, all of these 5 markers were only outliers in one lineage. In another example, 1 marker (*OgoRAD*_57650) significant in both lineages was colocated on LG 3 at 67.11 cM with another marker (*OgoRAD*_7571) that was only significant in the odd-year lineage (Supplementary Figure S3). Nine of the 77 candidates for selection within lineages were successfully annotated (Supplementary Table S2). None of these represented known candidate genes for environmental adaptation, and the low number of annotations prevented functional comparisons among marker sets.

Nine of the 24 unique candidates for divergent selection between lineages (marker set iii) mapped to 8 linkage groups (Supplementary Table S2 and Supplementary Figure S2). Two of these candidates (*OgoRAD_5628* and *OgoRAD_27545*) were detected at all 3 sampling locations of which 1 (*OgoRAD_27545*) mapped to LG 8 (Supplementary Figure S4), and 2 outliers (*OgoRAD_36809* and *OgoRAD_64756*) were detected independently in 2 separate locations (Figure 2). These results support the existence of multiple genomic regions of potential adaptive variation that may be unique within each of the pink salmon lineages (Supplementary Figure S2, Supplementary Figure S5). Only 3 of the 24 outliers in marker set (iii) were successfully annotated to known proteins; however, these did not include known candidate genes (see Supplementary Material, Table S2).

Most loci exhibiting signatures of selection within or between lineages (marker sets ii and iii) were only detected in 1 test (Figure 2); however, outlier categories were largely supported by the 3 AMOVA tests. Candidates for selection within lineages showed an expected pattern of increased differentiation among populations within lineages (F_{SC}) whereas candidates for divergent selection between lineages mainly exhibited increased differentiation between lineages (F_{CT} in Figure 4) .

No differences in diversity for neutral loci (marker set i) or candidates for selection within lineages (marker set ii) were observed (Figure 5a,b). However, a consistent pattern of significantly reduced diversity (H_O) in even-year lineage populations, and increased H_O in the odd-year lineage, was observed among outliers for divergent selection between lineages (marker set iii; Figure 5c).

Discussion

We demonstrated the usefulness of combining haploid mapping with RAD sequencing to efficiently generate a high-density linkage map with 7035 markers spanning 1658 unique genomic locations. We then combined this linkage map with genome scans to identify outlier loci and genomic regions with signatures of selection within or between temporally isolated lineages of pink salmon.

Haploid Crosses Facilitate Mapping of a Complex Genome

The use of haploids allowed efficient screening for confounding paralogous sequence variants detected as any heterozygote genotypes (Spruell et al. 1999; Everett and Seeb 2014). This greatly facilitated the creation of a linkage map with loci known to segregate disomically, a critical condition assumed by most linkage mapping algorithms (Broman et al. 2003). Identifying paralogous loci is more challenging in diploids as true heterozygotes cannot readily be distinguished from the joint scoring of paralogous loci (Seeb et al. 2011).

One potential caveat with RAD sequencing is that loci are often detected *de novo* in each study compromising comparisons among studies. However, here we demonstrate the ability to merge 2 independent RAD-based linkage maps to construct a single large consensus map; an important finding in light of the increasing popularity of this approach for constructing de novo linkage maps in non-model species.

Linkage Map Reveals Novel Signatures of Parallel Selection

The combined results of this study and those of Seeb et al. (2014) are the first to report genome-wide screens to detect and identify multiple genomic regions with signatures of selection in pink salmon. These 2 studies describe a genomic background dominated by neutral variation interspersed by smaller regions reflecting patterns of selection within as well as between lineages (cf., Aspinwall 1974) further adding to previous findings of adaptive divergence among populations within lineages (Kovach et al. 2012; Gharrett et al. 2013).

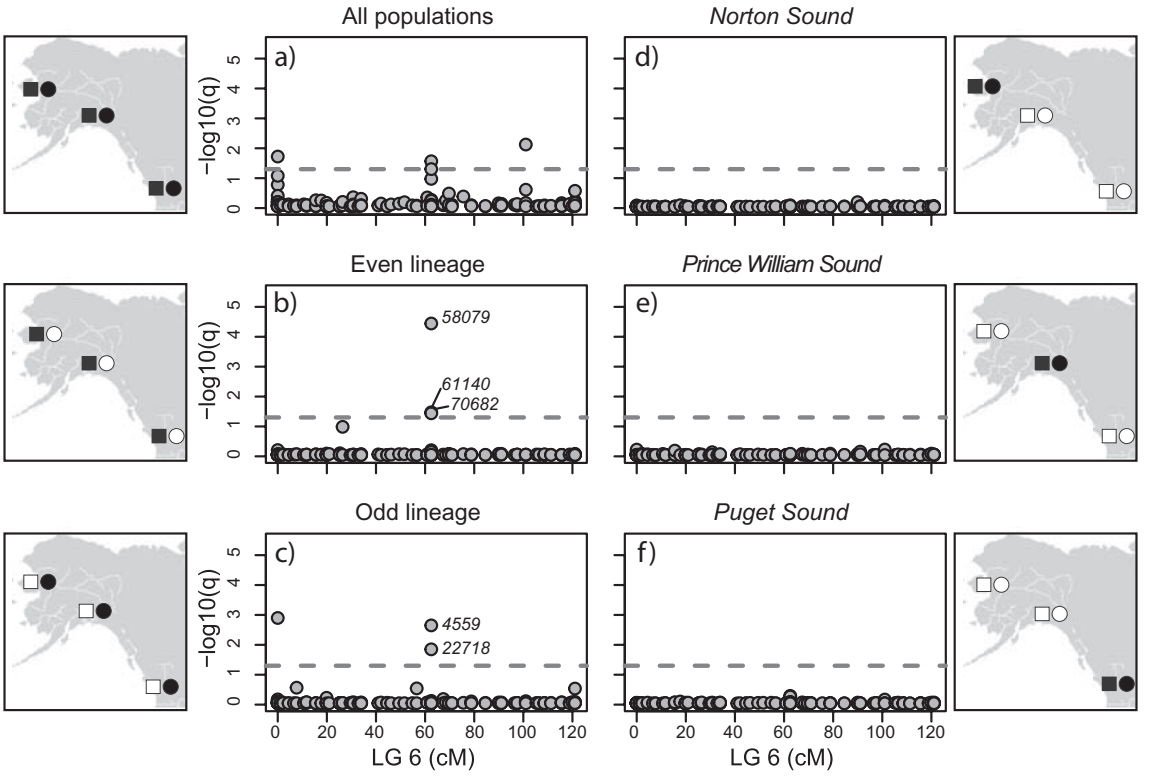

Figure 3. BayeScan results for linkage group 6: Three markers (*OgoRAD_58079*, *OgoRAD_61140* and *OgoRAD_70682*) at position 62.51 cM reveal a signature of divergent selection among populations within the even-year lineage (**a**, **b**). A similar pattern at the same position is seen for two different markers (*OgoRAD_4559* and *OgoRAD_22718*) within the odd-year lineage (**a**, **c**). None of these markers were also suggested to be affected by divergent selection between lineages (**d**–**f**). Letters correspond to the different genome scans described in methods and Table 1. Population samples included in each genome scan are shown as black symbols on adjacent maps, while white symbols represent populations not included in the respective genome scan. Even-year populations are shown as squares and odd-year populations as circles. Marker prefixes (*OgoRAD_*) have been omitted for the purpose of visualization. The broken horizontal gray line represents the significance threshold ($q = 0.05$).

Seeb et al. (2014) present evidence for parallel patterns of selection at 15 loci between lineages; here, using a linkage map, we strengthen those findings by showing shared genomic locations of outliers only detected in 1 of the 2 lineages. This leads us to conclude that some fraction of outliers, only detected in a single lineage, nevertheless represent genomic regions that may have exhibited parallel responses to selection in both lineages (Figure 2; see also Seeb et al. 2014).

Mapping outliers onto a linkage map to identify genomic regions of adaptive importance is not a novelty of this study (e.g., Bradbury et al. 2013; Gagnaire et al. 2013b; Larson et al. 2014). However, we provide a concrete example illustrating the complimentary power of performing map-assisted genome scans to distinguish signatures of selection shared between, or unique to, reproductively isolated lineages. These results emphasize the importance of generating linkage maps for non-model organisms currently lacking such resources.

Lineage-Specific Genetic Variation: Past and Future Evolutionary Responses

Considering the assumed existence of different pools of standing variation between lineages, we would also expect some fraction of lineage-specific outliers to reflect true differences in adaptive responses. Here, we found a striking difference in observed heterozygosity for outliers detected between lineages (marker set iii), clearly suggesting a scenario of unique pools of standing (adaptive) genetic variation. The consistent pattern of reduced diversity in the even-year lineage may originate from historical selective sweeps. Alternatively, increased diversity in the odd-year lineage may have increased its potential to adapt to post-glacial conditions through, for example, balancing selection maintaining high diversity. These explanations are not mutually exclusive, but adaptive divergence between lineages is further supported by the numerous outliers and genomic regions only showing signatures of selection within a single lineage (Figure 2, Supplementary Figure S2). These findings represent the first genomic insights of adaptive signatures between pink salmon lineages. Genomic signatures may help to explain ecological evidence for adaptive variation between lineages such as failed attempts to transplant even-year populations to rivers only populated by the odd-year lineage (Heard 1991) and outbreeding depression in lineage hybrids (Gharrett et al. 1999).

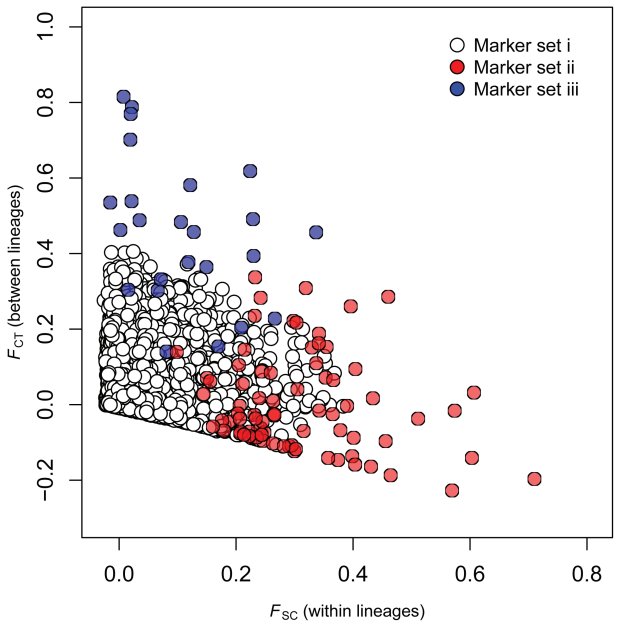

Figure 4. Distribution of genetic variation as explained by the following hierarchical levels; among populations within lineages (F_{SC}), and between lineages (F_{CT}). For each locus; F_{SC} and F_{CT} values are plotted for the three different marker sets; (i) neutral markers, (ii) candidates for selection within lineages, and (iii) candidates for divergent selection between lineages.

We cannot rule out that some outliers between lineages reflect a demographic bottleneck in 1 or both lineages rather than unique responses to natural selection (Akey et al. 2004; Jensen et al. 2005). However, while a bottleneck event may lead to individual false outlier loci (Akey et al. 2004; Jensen et al. 2005), bottleneck events are also expected to leave a genome-wide effect. The consensus pattern of observed heterozygosity for the 7820 neutral loci does not suggest any population- or lineage-specific effects (Figure 5a). Further, estimates of genetic diversity were very consistent within and between lineages (Seeb et al. 2014), suggesting no severe effects of demographic bottlenecks. Compared with a bottleneck signal, genes affected by natural selection are more likely to maintain such signals through purifying selection. Thus, coupled with previous observations and the conservative statistical approach applied here, we believe that our results reflect at least some genomic regions affected by natural selection, supporting the existence of different adaptive potentials between lineages.

A similar scenario of adaptive differences between 2 diverged lineages was proposed by Prunier et al. (2012). They reported climate associated outliers in a Western and Eastern lineage of the black boreal spruce (*Picea mariana*) and found 16 out of 23 adaptive SNPs to be specific to 1 lineage (8 in each). The authors then discussed 2 main explanations warranting comparison with our results for pink salmon. First, independent drift within each lineage could have led to elimination of future adaptive alleles in 1 lineage, impeding its ability to respond to subsequent environmental selection.

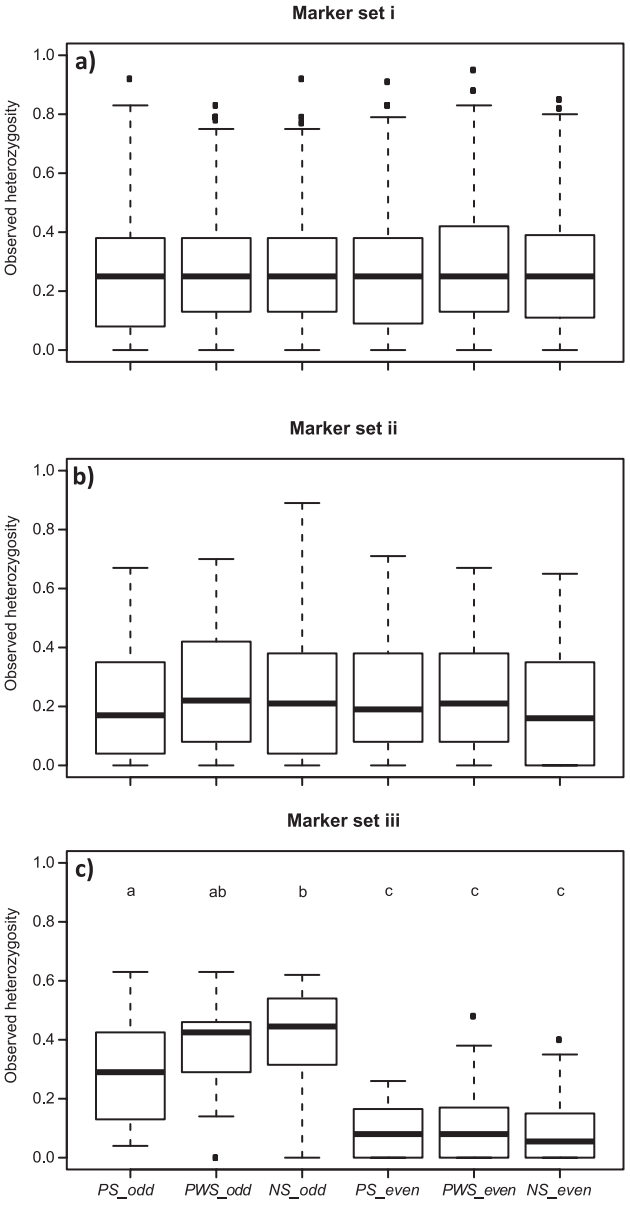

Figure 5. Box and whisker plots of observed heterozygosity (*Ho*) within each population for the neutral loci (marker set i), candidates for selection within lineages (marker set ii), and candidates for divergent selection between lineages (marker set iii). (**a**) ANOVA tests were not significant in more than 2 out of 1000 repetitions among all pairwise comparisons with neutral markers. (**b**) No significant differences were observed for the candidates for selection within lineages. (**c**) The ANOVA based on the 24 candidates for divergent selection between lineages revealed a consistent pattern of reduced diversity (*Ho*) in the even-year lineage populations compared with odd-year populations. Populations not sharing letters are significantly different ($P < 0.05$).

A similar scenario seems likely for explaining lineage-specific responses to selection in pink salmon. Indeed, Seeb et al. (2014) found no correlation in locus specific F_{ST} estimates

between lineages for the vast majority of loci, in agreement with independent evolutionary trajectories (see Figure 4 in Seeb et al. 2014). Second, Prunier et al. (2012) proposed that other untested environmental variables varying between the lineages' habitats could explain the emergence of region-specific outlier loci. However, contemporary environmental heterogeneity is not expected to explain lineage specific outliers in pink salmon, as lineage pairs are only separated temporally while inhabiting identical habitats. Hence, we note that differences in standing genetic variation between lineages of pink salmon are likely to originate from either drift throughout the period of reproductive isolation and/or from historical adaptations to distinct refugial environments.

While we cannot distinguish the above 2 scenarios, our findings have further relevance for understanding broader scale ecosystem impacts. Independent population dynamics cause interannual shifts in high and low pink salmon abundances, with odd-year populations tending to dominate to the south while even-year populations dominate at more northerly locations. Krkosek et al. (2011) discussed potential explanations of these interannual asynchronies in abundance and hypothesized that density dependence and stochasticity are the main drivers. We argue that not only ecological drivers, but also differences in genomic backgrounds, play a role in explaining varying interannual population dynamics across the distribution of pink salmon. Lastly, these large asynchronies in local abundance of the 2 lineages have resulted in major interannual shifts between 2 alternate states of a complex marine ecosystem (Springer and van Vliet 2014). Such large-scale ecosystem impacts have also been observed from periodical cicadas (Koenig and Liebhold 2005). For pink salmon, we predict that lineages will exert different responses to future climate change, affecting the degree of such interannual shifts in ecosystem states with broad consequences for the productivity of numerous other species.

Future Directions

We note that the current resolution of our linkage map is not expected to capture all relevant signals of divergence across the pink salmon genome (see relevant discussions in Santure et al. 2013; Johnston et al. 2014). However, we show that new maps can be added to existing resources, and we are currently expanding our map with even- and odd-year populations from Asia to further increase genomic coverage of species-wide variation. Yet, loci showing signatures of divergent selection here represent a foundation of candidate genomic regions for future studies seeking more precise identification of the genes underlying adaptive differences between lineages. One particularly promising approach to increase annotation of linkage maps includes alignment of anonymous RAD loci against fully annotated reference genomes of closely related species such as the rainbow trout (Berthelot et al. 2014). This approach is expected to expand as genomic resources become increasingly available.

Candidate genes of particular interest include metabolic pathway genes; it is known that odd-year populations use more energy for growth and less for lipid storage compared with even-year populations (Beamish 2012). This observation has been used to hypothesize a metabolic explanation for differences in growth patterns and climate responses between lineages (Beamish 2012). Particularly interesting genes include the pituitary growth hormone (GH) and insulin-like growth factor-I (IGF-I) genes (Beamish and Mahnken 2001). In conclusion, the unique replication of allochronic lineages of pink salmon will serve as a prime model for understanding replicated responses to climate change between 2 independent gene pools of this keystone species.

Supplementary Material

Supplementary material can be found at "http://www.jhered.oxfordjournals.org/".

Funding

Gordon and Betty Moore Foundation (1453); The Danish Council for Independent Research's career program Sapere Aude (12-126687) to M.T.L.

Acknowledgments

We thank Chris Habicht, the Alaska Department of Fish and Game, and Sewall Young, Washington Department of Fish and Wildlife, for providing samples of pink salmon. We thank Carita E. Pascal for invaluable contributions in the laboratory. Wesley A. Larson, Fred W. Allendorf and 3 anonymous reviewers are thanked for constructive comments that all improved the manuscript. Washington Department of Fish and Wildlife staff kindly provided access to activities at the Hoodsport Hatchery.

Conflict of Interest

The authors declare no conflict of interest.

References

Abbot P, Withgott JH. 2004. Phylogenetic and molecular evidence for allochronic speciation in gall-forming aphids (*Pemphigus*). Evolution. 58:539–553.

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol. 2:1591–1599.

Alexander RD, Bigelow RS. 1960. Allochronic Speciation in Field Crickets, and a New Species, *Acheta veletis*. Evolution. 14:334–346.

Aspinwall N. 1974. Genetic analysis of North American populations of the pink salmon, *Oncorhynchus gorbuscha*, possible evidence for the neutral mutation-random drift hypothesis. Evolution. 28:295–305.

Baker CS. 2013. Journal of heredity adopts joint data archiving policy. J Hered. 104:1.

Bams RA. 1976. Survival and propensity for homing as affected by presence or absence of locally adapted paternal genes in 2 transplanted populations of pink salmon (*Oncorhynchus gorbuscha*). J Fish Res Board Can. 33:2716–2725.

Barrett RD, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. Nat Rev Genet. 12:767–780.

Barrett RD, Schluter D. 2008. Adaptation from standing genetic variation. Trends Ecol Evol. 23:38–44.

Beacham TD, Murray CB. 1988. Variation in developmental biology of pink salmon (*Oncorhynchus gorbuscha*) in British Columbia. Can J Zool. 66:2634–2648.

Beamish RJ. 2012. Observations and speculations on the reasons for recent increased in pink salmon production. International Workshop on Explanations for the High Abundance of Pink and Chum Salmon and Future Trends. North Pacific Anadromous Fish Commission Technical Report No. 8.1-8.

Beamish RJ, Mahnken C. 2001. A critical size and period hypothesis to explain natural regulation of salmon abundance and the linkage to climate and climate change. Prog Oceanogr. 49:423–437.

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, Bento P, Da Silva C, Labadie K, Alberti A, et al. 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat Commun. 5:3657.

Bradbury IR, Hubert S, Higgins B, Bowman S, Borza T, Paterson IG, Snelgrove PV, Morris CJ, Gregory RS, Hardie D, et al. 2013. Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. Evol Appl. 6:450–461.

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 19:889–890.

Brykov VA, Polyakova N, Skurikhina LA, Kukhlevsky AD. 1996. Geographical and Temporal Mitochondrial DNA Variablity in Popualtions of Pink Salmon. J Fish Biol. 48:899–909.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. G3 (Bethesda). 1:171–182.

Churikov D, Gharrett AJ. 2002. Comparative phylogeography of the two pink salmon broodlines: an analysis based on a mitochondrial DNA genealogy. Mol Ecol. 11:1077–1101.

De Mita S, Thuillet AC, Gay L, Ahmadi N, Manel S, Ronfort J, Vigouroux Y. 2013. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. Mol Ecol. 22:1383–1399.

Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA. 2011. Local de novo assembly of RAD paired-end contigs using short sequencing reads. PLoS One. 6:e18561.

Everett MV, Miller MR, Seeb JE. 2012. Meiotic maps of sockeye salmon derived from massively parallel DNA sequencing. BMC Genomics. 13:521.

Everett MV, Seeb JE. 2014. Detection and mapping of QTL for temperature tolerance and body size in Chinook salmon (*Oncorhynchus tshawytscha*) using genotyping by sequencing. Evol Appl. 7:480–492.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 10:564–567.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 131:479–491.

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

Gagnaire PA, Normandeau E, Pavey SA, Bernatchez L. 2013a. Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (*Coregonus clupeaformis*). Mol Ecol. 22:3036–3048.

Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L. 2013b. The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. Evolution. 67:2483–2497.

Gharrett AJ, Joyce J, Smoker WW. 2013. Fine-scale temporal adaptation within a salmonid population: mechanism and consequences. Mol Ecol. 22:4457–4469.

Gharrett AJ, Smoker WW, Reisenbichler RR, Taylor SG. 1999. Outbreeding depression in hybrids between odd- and even-broodyear pink salmon. Aquaculture. 173:117–129.

Gilbert CH. 1912. Age at maturity of the Pacific coast salmon of the genus *Oncorhynchus*. Bull US Bur Fish. 32:1–22.

Heard WR. 1991. Life history of pink salmon. In: Pacific salmon. (Groot C, Margolis L, editors). University of British Columbia Press: Vancouver (BC) pp. 119–230.

Heliövaara K, Väisänen R, Simon C. 1994. Evolutionary ecology of periodical insects. Trends Ecol Evol. 9:475–480.

Hemmer-Hansen J, Nielsen EE, Therkildsen NO, Taylor MI, Ogden R, Geffen AJ, Bekkevold D, Helyar S, Pampoulie C, Johansen T, et al.; FishPopTrace Consortium. 2013. A genomic island linked to ecotype divergence in Atlantic cod. Mol Ecol. 22:2653–2667.

Huang X, Madan A. 1999. CAP3: A DNA sequence assembly program. Genome Res. 9:868–877.

Jensen JD, Kim Y, DuMont VB, Aquadro CF, Bustamante CD. 2005. Distinguishing between selective sweeps and demography using DNA polymorphism data. Genetics. 170:1401–1410.

Johnston SE, Orell P, Pritchard VL, Kent MP, Lien S, Niemelä E, Erkinaro J, Primmer CR. 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (*Salmo salar*). Mol Ecol. 23:3452–3468.

Koenig WD, Liebhold AM. 2005. Effects of periodical cicada emergences on abundance and synchrony of avian populations. Ecology. 86:1873–1882.

Kosambi DD. 1943. The estimation of map distances from recombination values. Ann Eugenic. 12:172–175.

Kovach RP, Gharrett AJ, Tallmon DA. 2012. Genetic change for earlier migration timing in a pink salmon population. Proc Biol Sci. 279:3870–3878.

Krkosek M, Hilborn R, Peterman RM, Quinn TP. 2011. Cycles, stochasticity and density dependence in pink salmon population dynamics. Proc Biol Sci. 278:2060–2068.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9:357–359.

Larson WA, Seeb LW, Everett MV, Waples RK, Templin WD, Seeb JE. 2014. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). Evol Appl. 7:355–369.

Lewontin RC, Krakauer J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics. 74:175–195.

Lindner KR, Seeb JE, Habicht C, Knudsen KL, Kretschmer E, Reedy DJ, Spruell P, Allendorf FW. 2000. Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. Genome. 43:538–549.

Matsuoka MP, Gharrett AJ, Wilmot RL, Smoker WW. 2004. Genetic linkage mapping of allozyme loci in even- and odd-year pink salmon (*Oncorhynchus gorbuscha*). J Hered. 95:421–429.

Miller MR, Brunelli JP, Wheeler PA, Liu S, Rexroad CE 3rd, Palti Y, Doe CQ, Thorgaard GH. 2012. A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. Mol Ecol. 21:237–249.

Narum SR, Hess JE. 2011. Comparison of F(ST) outlier tests for SNP loci under selection. Mol Ecol Resour. 11(Suppl 1):184–194.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J Mol Evol. 19:153–170.

Nielsen R. 2005. Molecular signatures of natural selection. Annu Rev Genet. 39:197–218.

Phillips RB, Kapuscinski AR. 1987. A Robertsonian polymorphism in pink salmon (*Oncorhynchus gorbuscha*) involving the NOR region. Cytogenet Cell Genet. 44:148–152.

Prunier J, Gérardi S, Laroche J, Beaulieu J, Bousquet J. 2012. Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. Mol Ecol. 21:4270–4286.

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

Rice WR. 1989. Analyzing tables of statistical tests. Evolution. 43:223–225.

Rousset F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour. 8:103–106.

Santos H, Rousselet J, Magnoux E, Paiva MR, Branco M, Kerdelhué C. 2007. Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. Proc Biol Sci. 274:935–941.

Santure AW, De Cauwer I, Robinson MR, Poissant J, Sheldon BC, Slate J. 2013. Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. Mol Ecol. 22:3949–3962.

Seeb JE, Pascal CE, Grau ED, Seeb LW, Templin WD, Harkins T, Roberts SB. 2011. Transcriptome sequencing and high-resolution melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. Mol Ecol Resour. 11:335–348.

Seeb JE, Pascal CE, Ramakrishnan R, Seeb LW. 2009. SNP genotyping by the 5 '-nuclease reaction: advances in high-throughput genotyping with nonmodel organisms In: Methods in molecular biology, single nucleotide polymorphisms, vol. 578: (Komar AA, ed), pp. 277–292. Humana Press, New York.

Seeb JE, Seeb LW. 1986. Gene mapping of isozyme loci in chum salmon. J Hered. 77:399–402.

Seeb LW, Waples RK, Limborg MT, Warheit KI, Pascal CE, Seeb JE. 2014. Parallel signatures of selection in temporally isolated lineages of pink salmon. Mol Ecol. 23:2473–2485.

Springer AM, van Vliet GB. 2014. Climate change, pink salmon, and the nexus between bottom-up and top-down forcing in the subarctic Pacific Ocean and Bering Sea. Proc Natl Acad Sci U S A. 111:E1880–E1888.

Spruell P, Pilgrim KL, Greene BA, Habicht C, Knudsen KL, Lindner KR, Olsen JB, Sage GK, Seeb JE, Allendorf FW. 1999. Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon. J Hered. 90:289–296.

Takezaki N, Nei M, Tamura K. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Mol Biol Evol. 27:747–752.

Thornton KR, Jensen JD. 2007. Controlling the false-positive rate in multilocus genome scans for selection. Genetics. 175:737–750.

Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S, Ihara-Ujino T. 2012. Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. Heredity (Edinb). 109:349–360.

Wu Y, Bhat PR, Close TJ, Lonardi S. 2008. Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. PLoS Genet. 4:e1000212.

Wu Y, Close TJ, Lonardi S. 2011. Accurate construction of consensus genetic maps via integer linear programming. IEEE/ACM Trans Comput Biol Bioinform. 8:381–394.

Zhivotovsky LA, Gharrett AJ, Mcgregor AJ, Glubokovsky MK, Feldman MW. 1994. Gene Differentiation in Pacific Salmon (*Oncorhynchus Sp*) - Facts and Models with Reference to Pink Salmon (*Oncorhynchus gorbuscha*). Can J Fish Aquat Sci. 51:223–232.

Zueva KJ, Lumme J, Veselov AE, Kent MP, Lien S, Primmer CR. 2014. Footprints of directional selection in wild Atlantic salmon populations: evidence for parasite-driven evolution? PLoS One. 9:e91672.

Received June 17, 2014; First decision August 15, 2014; Accepted August 27, 2014

Corresponding Editor: Fred Allendorf