Assessing stock structure of Pacific cod (*Gadus macrocephalus*) by combining ecological and evolutionary perspectives

Authors:

Fisher, M.C.[1]

Affiliations:

1. School of Aquatic and Fishery Sciences, University of Washington, Seattle WA 98105, USA

Abstract

This study provides ecological context for genetic population structure of Korean Pacific cod stocks by analyzing otolith microchemistry data and comparing it with existing population genetic analysis. A total of 167 samples from five spawning aggregates around the Korean peninsula were analyzed for otolith edge and core microchemistry, and sequenced for SNP genetic markers. I first used PERMANOVAs and an NMDS to establish that profiles composed of eight elements differed significantly between sampling sites, with particular differentiation at Jinhae Bay, Pohang, and Yellow Sea sites. This proves that otolith microchemistry can be useful in a marine setting, and identified specific elements that vary in concentration along the Korean coastline. A Mantel test showed that edge microchemistry was moderately correlated with core microchemistry, suggesting that most individuals were caught on spawning grounds with similar water chemistry as their natal spawning ground. Hierarchical agglomerative cluster analysis somewhat verified this result; although the majority of individuals at each site clustered together using both edge and core microchemistry data, there was greater spread within sites in the core than in the edge data. This is likely due to the presence of multiple age cohorts and slight variation in microchemistry across years. When compared to genetic analysis, otolith microchemistry was able to differentiate between spawning groups on a finer scale and between temporal replicates both within and between years. Overall, the combination of both data sets provided a more comprehensive description of population structure around the Korean peninsula, with particular insight into Jinhae Bay spawning patterns.

Introduction

Fisheries management plans that include population genetic structure reduce the probability of overexploitation of local populations and allow for a more sustainable yield[1]. Additionally, evolving threats to fisheries species from climate change have stressed the importance of preserving standing genetic variation and locally adapted populations, both of which can promote species resilience. As next-generation sequencing methods allow us to determine genetic structure on much finer scales in species characterized by high gene flow and large population size, the application of population genetic research to the management of purely marine species becomes more feasible.

Yet we still have much to learn about the ecological context in which population genetic patterns evolve. This is especially true of marine species, which encounter fewer barriers to dispersal and tend to consist of large, geographically expansive populations. By combining population genetic approaches with ecological studies of individual life history, we can gain additional insight into the drivers of evolutionary patterns on ecological timescales in marine populations.

This may also provide marine fisheries managers with a more complete toolkit to identify and manage population structure. For example, genetic analyses only detect migration on evolutionary scales, and only when migrants are reproductively successful. Recent studies have suggested that migration without reproduction is more common in fish species than previously believed[2]. Skipped spawning, in which a mature fish with an annual reproductive cycle does not spawn during a given year, is now thought to characterize many species and populations of fish. While this may not have an effect on population structure on evolutionary timescales, it has a significant impact on the yearly composition of local populations and therefore resource

management. Fisheries are primarily managed on ecological timescales, and so the connectivity of populations through individual movement and behavior is a key consideration when discussing stock structure.

Clearly, there is a need to couple studies of individual life history with population genetics in marine fish species. In the marine environment, traditional mark-recapture studies used to characterize individual behavior are limited in size and scope. Otolith microchemistry provides an alternative to such tagging methods. Otoliths accumulate layers of calcium carbonate on a protein matrix throughout a fish's life, beginning in early larval stages. The layers incorporate some elemental chemistry from the body of water surrounding the fish. Otolith microchemistry can therefore be used as a type of "flight path" when matched to chemical profiles of water bodies. Previous studies have successfully combined genetic and otolith microchemistry data sets to distinguish between populations of Patagonian toothfish caught off of South America and Antarctica[3], to prove limited mixing between black rockfish populations[4], and to provide complementary information on the natal origin and genetic structure of a migratory cyprinid species[5].

However, such studies used 20 or fewer microsatellites or allozymes for genetic analyses. Next-generation sequencing methods, which generate thousands of loci for finer-scale genetic population structure analysis, may be more conducive to ecological comparisons. Additionally, otolith microchemistry is often applied to estuarine or anadromous species, or across large geographic distances, to ensure spatial differences in elemental profiles. Otolith microchemistry applications to proximal marine populations is comparatively limited, primarily because of the relative homogeneity in element concentrations[6].

Pacific cod is an excellent case study to address key management questions on population structure, while testing applications of new technologies. A marine finfish species which supports major fisheries across its North Pacific range, Pacific cod form coastal spawning aggregates during the winter months to breed. While such aggregates can be in close geographic proximity, previous genetic analyses suggest that Pacific cod can display distinct genetic breaks between aggregates[7–10]. This population structure is critical to sustainable fisheries management of the species; a simulation study of Pacific cod stocks in Alaska found that establishing management units based on genetic research maintained stock sizes at target levels and even increased catch, while not managing for genetically distinct populations could reduce stock sizes below target levels[1].

The maintenance of genetically distinct spawning aggregates is likely a result of site fidelity and short larval dispersal distances, estimated from studies conducted along the western US and Canadian coastlines[8,11]. However, direct evidence of site fidelity in the form of individual migration is primarily derived from a single mark-recapture study conducted in the northwestern United States[11]. Field observations of non-spawning mature adults in Pacific cod spawning aggregates also raise the possibility that Pacific cod may display skipped spawning, a common behavior observed in their Atlantic relatives[2]. More direct evidence of site fidelity, and the linkage of individual behavior to genetic population structure, would provide needed insight into the biology of this species and greatly benefit Pacific cod fisheries management.

The purpose of this study was to explore Pacific cod population structure and site fidelity around the Korean peninsula according to otolith microchemistry, and to then compare these results with next-generation sequencing data. In doing so, the study addresses three major questions:

1. How different are otolith microchemistry concentrations across Pacific cod spawning grounds around the Korean peninsula?
2. Does otolith microchemistry support existing evidence that Pacific cod display site fidelity to their natal spawning site?
3. What additional insights and management tools can be gained by comparing and/or integrating next-generation genetic data with otolith microchemistry data?

Methods

Sample Collection

Tissue samples and otoliths were collected from 322 individuals across seven Pacific cod spawning aggregates in coastal waters around the Korean peninsula (**Table 1; Fig. 1**). Three aggregates were sampled during the 2007 – 2008 winter spawning season, three aggregates during the 2014-2015 season, and one during the 2015 – 2016 season. Two additional spawning groups were sampled as temporal replicates, one at Geoje as a between-year replicate during the 2013-2014 season, and one at Jinhae Bay as a within-year replicate during the 2007-2008 season. Tissue samples were preserved in 95-100% non-denatured ethanol, stored at 4°C.

Otolith Data Collection

A total of 261 otoliths were scanned for thirteen trace elements using Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) at the University of Oregon (**Fig. 2**). Element concentrations were then standardized against a calcium standard. Of the thirteen original elements, eight had relative concentrations that were large enough to produce a signal

that could be distinguished from the calcium baseline, and retained as "informative" for further analysis: boron (B11), barium (Ba138), Lithium (Li7), magnesium (Mg24), manganese (Mn55), lead (Pb208), strontium (Sr88) and zinc (Zn66). Otoliths were then aged (**Fig. 2**) and each year of growth partitioned into 4 additional sections, A through D. Since multiple LA-ICP-MS measurements were taken within each section, concentration ratios were averaged together to reach a single concentration ratio for that section. This study utilizes the mean concentration ratios from measurements at the core of the otolith (Year 0; Section A), and from measurements at the edge of the otolith (Year varies by individual; Section D). Otolith core concentrations were assumed to be indicative of the natal spawning ground, and otolith edge concentrations indicative of the spawning ground on which the individual was captured for sampling.

Genetic Data Collection

DNA was extracted from 322 fin clip and tissue samples using DNeasy Blood & Tissue Kits (Qiagen, Inc.). DNA was then quantified using Quant-iT PicoGreen dsDNA Reagent (Invitrogen, Carlsbad, CA) and visualized on an agarose gel. A total of 283 samples were prepared for RAD sequencing according to Baird et al. (2008)[13] and Etter et al. (2011)[14], with modifications to incorporate Agencourt AMPure XP beads (Beckman Coulter, Inc., Pasadena, CA) for purification[15]. Sequencing of 150bp single end (n = 157) and paired end (n = 126) reads was completed on an Illumina HiSeq 4000 (Illumina, Inc.) at the University of Oregon's Genomics and Cell Characterization Core Facility.

Quality filtering and demultiplexing of raw RAD sequencing data, as well as de novo construction of a reference database of RAD loci, SNP discovery, and genotyping was completed using a combination of the Stacks 1.44[16,17], Bowtie[18], and NCBI's Basic Local Alignment Search

Tool, BLAST[19], according to the procedures outlined in Gruenthal et al. (in press)[15] and Brieuc et al. (2013)[20]. A genotype file containing putative polymorphic SNPs present in ≥ 80% of fish per spawning aggregate was randomly filtered to include one SNP per RAD tag. Final filtering was then used to remove loci with minor allele frequencies (MAFs) with < 0.05 in all spawning aggregates, loci that did not conform to Hardy – Weinberg equilibrium, and loci missing more than 20% of genotypes in any of the spawning aggregates. Individuals were removed from the data set if they were missing genotypes at more than 30% of loci, or were found to be potentially contaminated using a combination of heterozygosity distribution and ML-Relate[21].

Data Subsetting

After filtering out individuals which were not present in both genetic and otolith microchemical datasets, data were subset to only include individuals two years and older, according to otolith aging methods. Fish less than two years old are considered juveniles, and therefore not spawning fish, which are the subject of this study. Additionally, it is not known whether there is an ontogenetic shift in spawning migration patterns between adult and juvenile fish; therefore, the inclusion of juvenile fish in these analyses could lead to the mischaracterization of mature Pacific cod spawning behavior.

The size of the data set was further reduced for computational simplicity and to achieve more even sample sizes. Individuals were randomly removed from both Jinhae Bay 2007-2008 sample sets, so that the number of samples in each group reached n = 30. This left a total of 167 samples collected across five spawning aggregates, with temporal replicates taken at two sites (**Fig. 1, Table 1**).

Analysis of Otolith Microchemistry Data

Multivariate normality of otolith microchemistry data was assessed using Mardia's multivariate normality test and QQPlot visualization with the R package MVN[22]. The data did not conform to multivariate normality. Data transformations, including natural log, $\log_{10}$, cube root, and square root transformations, did not normalize the data set. Thus the multivariate analytical methods (below) for this data do not include distributional assumptions.

Visualization of otolith microchemistry data by element at the core and edge of the otolith showed markedly different scales (**Fig. 3**), so each element's concentrations were relativized by maxima using the R package vegan[23]. Calculation of the coefficient of variance of column totals produced a value of 264.47, which indicates that the degree of variability in the columns would have had a large effect on results[24]. A Euclidean distance matrix was then calculated separately for the elemental ratios measured at the core of the otolith, and the elemental ratios measured at the edge of the otolith.

To determine whether elemental fingerprints (composed of all eight element concentration ratios) differed significantly between the sampled Pacific cod spawning grounds, I performed a permutational analysis of variance (PERMANOVA) of the microchemistry at the edge of the otolith against spawning site, with 100,000 permutations. To explore potential temporal differences in otolith microchemistry, I also performed a PERMANOVA of elemental fingerprints at the edge of the otolith for Jinhae Bay samples against spawning month, and Geoje samples against spawning year. I then conducted one-way ANOVAs for those elements which could be normalized to determine element-specific differences between sampling site, spawning

year, and spawning month. This consisted of four elements: barium (natural log-transformed), magnesium (natural log-transformed), strontium, and zinc (natural log-transformed). I then ran a non-metric Multi-dimensional Scaling Analysis (NMDS) on otolith edge microchemistry, and fit element vectors to the ordination using vegan[23]. This allowed identification of elements which were correlated with the separation of specific sampling sites across the ordination space. Although this step is commonly completed for otolith microchemistry analysis with a Linear Discriminant Analysis[25–27], NMDS has been used when data do not conform to multivariate normality[3]. The NMDS was run with the wrapper function metaMDS in vegan[23], with a maximum of 400 iterations and minimum of 40 runs[24]. In order to determine the optimum number of dimensions, "k", I ran the NMDS with 1 through 7 dimensions (seven being the total number of sampling units), then evaluated a plot of dimensions against associated stress values. The minimum stress value was achieved with seven dimensions. However, a two-dimensional NMDS cannot adequately visualize variation across seven dimensions. I therefore used three dimensions for the final NMDS, as this was the lowest number of dimensions from which I had no real risk of drawing false inferences[28] (stress < 0.10).

To approach the question of site fidelity, a Mantel test was used to determine whether there was a correlation between the elemental fingerprint at the otolith core and the otolith edge for each fish, which would indicate similar elemental signatures of the individual's natal and final spawning ground. I ran the Mantel test using Pearson's correlation coefficient and 100,000 permutations.

While the Mantel test may suggest correlation between the elemental fingerprint at the core and edge of the otolith, it is not possible to explore specific differences in groupings of individuals, or individual assignment to groups. To do so, hierarchical polythetic agglomerative

cluster analysis was run using the function hclust in the R stats package[29]. I first conducted the analysis with three separate clustering methods: the Ward method[30] (ward.d2), complete linkage (default), and the average or UPGMA method. Clustering produced using the Ward method most closely matched expectations and groups visualized in the NMDS ordination, and so was used for final analyses. With NbClust[31], a gap index was calculated based on the Ward clustering method for both edge and core data sets. The best number of clusters and the optimal partition of individuals was determined based on the gap index, and a dendrogram generated for both core and edge microchemistry using the R package dendextend[32]. To compare clustering of individuals between the otolith edge and core datasets, I used a tanglegram, a common visualization in phylogenetics in which two dendrograms are drawn opposite each other, and then leaves connected with auxiliary lines[33]. The tanglegram function was called from the R package dendextend[32].

Analysis of Genetic Population Structure and Individual Assignment

The number of genetically distinct populations in the data was determined using STRUCTURE[34]. STRUCTURE places samples into groups whose members share similar patterns of variation using a Bayesian clustering algorithm with applied MCMC estimation. This allows for a progressive convergence toward reliable allele frequency estimates in each population, and membership probabilities of sampled individuals[34]. STRUCTURE was run using a burnin of 50,000 and 10,000 MCMC reps, with K = 1 through 9 assumed populations, and 3 iterations per value of K. The likelihood of the data was calculated for each value of K, and then the mean log-likelihoods plotted to determine the optimum value of K (the smallest "stable" K value that maximizes the global likelihood of the data). The population assignment for each

individual under the optimal value of K was then exported for comparison to otolith microchemistry.

Comparison of Genetic Assignment and Otolith Microchemistry

The first step in this comparison was to qualitatively assess the differences in spatial and temporal stock structure that are distinguishable from genetic analyses versus otolith microchemistry. I then visually assessed the groupings of individuals according to genetic assignment and otolith core microchemistry. To do so, I conducted an NMDS with the wrapper function metaMDS in vegan[23], using the same parameters as with the edge microchemistry NMDS - a maximum of 400 iterations and minimum of 40 runs[24]. To determine the optimum number of dimensions, "k", I ran the NMDS with 1 through 7 dimensions (the number of sampling units), then evaluated a plot of dimensions against associated stress values. The minimum stress value was achieved with seven dimensions, as in the NMDS conducted on edge microchemistry data. I again used three dimensions for the final NMDS, as this was the lowest number of dimensions from which I had no real risk of drawing false inferences[28] (stress < 0.10). I then coded point shape according to whether the individual's genetic material was primarily assigned to the West or the South/Southeast population.

Results

Elemental Fingerprints of Spawning Grounds

A PERMANOVA of elemental fingerprints (combination of all eight element concentration ratios) at the edge of the otolith by sampling site, spawning year at Geoje, and spawning month at Jinhae Bay all showed significant variation in elemental fingerprints ($p <$

0.01; **Tables 2a-c**). However, the proportion of variance in elemental fingerprints that could be accounted for by each of these factors, rather than by residuals, differed. Whereas sampling site could explain 45% of the variance in elemental fingerprints, spawning year could only explain 13% and spawning month 14% of total variation (**Fig. 4**).

I was also able to conduct a one-way analysis of variance, or ANOVA, to explore element-specific differences between sampling sites, spawning year, and spawning month for four of the eight elements that could be normalized (**Tables 3-5**). All four elements - barium, magnesium, strontium, and zinc - showed significant variation among sampling sites. Variation in magnesium concentrations across spawning years at Geoje was also significant, whereas there was a significant variation in barium concentrations across spawning months at Jinhae Bay.

A non-metric Multi-dimensional Scaling (NMDS) ordination provided further insight into how element concentrations differed between sampling sites. The NMDS, run with three dimensions, produced a stress value of 0.0934, which is associated with good representation of the data set[28], and a linear fit of R^2=0.963 (**Table 6; Fig. 5** inset). The ordination showed distinct separation between the Jinhae Bay, Pohang, and the Yellow Sea (YSBlock) samples (**Fig. 5**). Vectors indicating the fit of elements along the ordination axes suggest that boron (B11), barium (Ba138), and strontium (Sr88) concentrations lead to the differentiation of samples from Jinhae Bay. Samples from Pohang also appear to differentiate across the ordination space, correlated with magnesium (Mg24) and zinc (Zn66) concentrations, and somewhat correlated with lithium (Li7) and lead (Pb208). Yellow Sea (YSBlock) samples were correlated with higher manganese concentrations. The NMDS also displayed separation across ordination space in Jinhae Bay temporal replicates, which was found to be significant in the PERMANOVA. As suggested by the ANOVAs run for this site, this separation was correlated with concentration of the element

barium. The NMDS also suggests that concentration of boron differed between spawning months. Although Geoje temporal samples (2014 and 2015), which were found to have significantly different elemental fingerprints in the PERMANOVA, did not differentiate from each other across the ordination space, this could be a result of other samples being much more highly differentiated (i.e. Pohang, Yellow Sea, and Jinhae Bay), or of variation of only two out of three dimensions being represented in the NMDS plot.

Site Fidelity

A Mantel test between Euclidean distance matrices calculated from otolith core concentrations and otolith edge concentrations produced a Mantel R statistic of 0.3441, and a *p* value of 1×10^{-5}. This suggests a moderate positive correlation between full elemental fingerprints of each individual's natal spawning ground and the spawning ground at which they were caught.

Site fidelity was also visually explored through a cluster analysis, to determine if individuals that were caught on the same spawning ground would cluster together in an analysis based on core microchemistry. Hierarchical agglomerative cluster analysis analyzed with the Tibshirani et al. (2001)[35] gap statistic suggested that the data set was represented best by two clusters in both the otolith edge and otolith core datasets. In the otolith edge data set, one cluster was composed of approximately half of the Jinhae Bay samples, with the remaining samples across all sites in the second cluster; in the core data set, one cluster consisted of all Jinhae Bay samples as well as some samples from the Geoje and Yellow Sea (YSBlock) sites, while the second cluster consisted of a mixture of samples from all sites except Jinhae Bay (**Fig. 6a&b**).

Overall, individuals from the same sampling site tended to cluster together on lower subtrees, although upper subtrees consisted of a mix of different sampling sites.

The tanglegram comparison of clustering between edge and core otolith microchemistry showed that, on average, there was less clustering within sampling sites in the core microchemistry data. The extent to which this occurred varied by sampling site, ranging from the tight clustering of individual samples at Jinhae Bay to the greater spread of individual samples collected from the Yellow Sea (**Fig. 7a-d**). Between-year temporal samples at Geoje clustered together in both the edge and core microchemistry data; the within-year temporal sample at Jinhae Bay showed consistently separation of early and late spawners in both the edge and core microchemistry data. Overall greater intermixing of individuals between sampling sites according to core microchemistry was also evident when groupings suggested by cluster analysis were overlaid on the NMDS produced with edge microchemistry; hulls had far more overlap across sampling sites for core microchemistry than edge microchemistry (**Fig. 8a,b**).

Genetic population structure

STRUCTURE analyses suggested two genetically distinct populations, delineated between the western and south/southeastern coast. Assignment tests through the program primarily assigned individuals to the coast on which they were sampled. However, there were six individuals sampled from sites on the south/southeastern coast which had the majority of their genetic material assigned to the west coast: two fish from Pohang, two from Geoje (one from the 2013-2014 season, one from the 2014-2015 season) and three from Namhae sampling sites.

Genetic and otolith microchemistry comparisons

The NMDS of three dimensions conducted on otolith core microchemistry had a stress value of 0.0933 (**Table** 7) and a linear fit of 0.966 (**Fig. 9** inset). The individuals from the south/southeastern population that were genetically assigned to the western population did not show otolith core microchemistry which would confidently suggest that their true natal spawning aggregate was that of the Yellow Sea. When genetic population structure is overlaid on an NMDS of otolith core microchemistry (**Fig. 9**), there is not enough separation across the ordination space to identify outlier individuals in specific populations, except the Jinhae Bay spawning aggregates. However, one individual from the Pohang spawning aggregate which was genetically assigned to the west coast population is an apparent outlier on the NMDS. While this may be a result of individual variation in the uptake of water chemistry into the otolith, it is also possible that this individual derives from an unsampled spawning aggregate. The genetic assignment procedure will assign every individual to a given baseline, even if that baseline does not include the true population of origin.

In contrast to genetic data, which showed potential migrants from the west population to the southern coast, cluster analysis of otolith microchemistry showed several Yellow Sea samples clustering with the southern coast. However, fine-scale patterns in the cluster analysis should be interpreted carefully, given the low overall distinction between non-Jinhae Bay sites.

Discussion

Otolith Microchemistry Variation & Site Fidelity

Otolith microchemistry has been primarily used to discern fish migratory patterns in estuary systems or in anadromous species. This study proves that microchemistry can be useful in discerning between marine bodies of water, particularly along coastlines. Not only did elemental fingerprint vary by sampling site location, but there was also significant evidence of temporal variation in element concentrations in PERMANOVAs. It should be noted that the spawning site explained a far greater proportion of total variation than either spawning year or month. However, follow-up ANOVAs suggest that this may be a result of the number of elements which varied significantly in their concentrations. Four out of the four elements tested in ANOVAs showed significant variation among sampling sites. However, only one out of the four elements tested showed significant variation between spawning years and months.

When differences in sampling site elemental fingerprints were visually assessed through an NMDS, four sites became particularly distinguished across the ordination space. Samples from Pohang loaded higher than other sites on zinc and lead, two heavy metals. This is likely due to runoff from the port city of Pohang, which is a highly industrialized region. Jinhae Bay samples loaded high on strontium, barium and boron, all of which have been associated with salinity[36]. Jinhae Bay was a nearshore spawning aggregate along the southern coast of Korea, which consists of many fjord-like bays and inlets that may create a more heterogeneous salinity profile. However, the difference in element concentrations associated with salinity may also be a function of time period; the Jinhae Bay samples were taken during the 2007-2008 spawning season, whereas the Geoje, Namhae and Pohang samples were taken during either the 2013-2014 or 2014-2015 spawning season. Finally, Yellow Sea samples had higher manganese concentrations. This could be a result of this site being the furthest offshore, or because this was

the only site along the western, rather than the southern, coast. Additional samples from the western coast of the Korean peninsula could provide further insight into this pattern.

Temporal differences in element concentrations at Jinhae Bay, between months of the same spawning year, were visible in the NMDS and found to be significant in a PERMANOVA; temporal differences were also significant at Geoje, between spawning years, although not as evident in the NMDS ordination. This significant variation in elemental fingerprints between spawning years and months in the temporal replicates indicates that element concentrations are measurably variable over time.

This temporal variation could explain why core microchemistry did not consistently produce as distinct cluster by sampling site, or even region, as did the edge microchemistry. It may also have lowered the Mantel R statistic when correlating edge and core microchemistry. Since individuals caught at the same sampling site encompassed a variety of cohorts (i.e. ages), the element concentrations at natal spawning grounds may have varied enough by year for a spawning site to appear more similar to other sites within the same year than to itself across years.

Taking this into account, the dataset as a whole does suggest that Pacific cod display site fidelity to their natal spawning aggregates. This supports the previous tagging study completed in the northeastern Pacific, and conclusions drawn from prior microsatellite analyses[8,11]. Further subdivision of the data set by age cohorts may show stronger signals of site fidelity by removing temporal variation in elemental fingerprints at sampling sites.

Otolith microchemistry v. Genetic assignment testing

Genetic and otolith microchemistry analysis allowed for differentiation of spawning aggregates on different scales, with otolith microchemistry allowing for finer-scale spatial and temporal differentiation of individuals. Genetic data identified two highly differentiated populations. Of the two prior studies of population structure in Korean Pacific cod that were conducted with microsatellite loci[9,10], this study agrees with the most recently published manuscript by Gwak & Nakayama (2011)[9] which identified significant divergence between south and west coast spawning aggregates. The sharp genetic break between these two populations is in contrast to recent next-generation sequencing results on northeastern Pacific cod populations along the Alaskan and Canadian coast, which show more of an isolation-by-distance pattern[15].

Otolith microchemistry also suggested differences in elemental fingerprints between the western spawning aggregate, Yellow Sea, with the remaining samples. However, microchemistry was able to further distinguish between spawning aggregates within the south/southeastern population, particularly the Jinhae Bay and Pohang sites, and suggested site fidelity to natal spawning aggregates. Site fidelity to subgroups within a single genetic population at first appears contradictory; however, this is likely because even very low levels of effective migration among subgroups is enough to drown out significant differentiation (generally estimated at one to ten migrants per generation[37]). Such low levels of migration in a very large population might not be detected with the number of samples collected for this study.

This was the first study to assess genetic differentiation between temporal replicates in Pacific cod spawning aggregates. Structure analyses suggested no significant divergence between spawning years or months at two of the sampling sites. Otolith microchemistry, in

contrast, showed significant variation between temporal replicates, both between spawning years at Geoje and between spawning months at Jinhae Bay.

One of the most interesting conclusions that could be drawn from genetic and otolith comparisons regards the Jinhae Bay spawning aggregates. A temporal sample of different months (December and February) within the same spawning season was taken at this location because field observations suggested that two separate aggregates formed at different times during the spawning season[38]. Genetic data show that on evolutionary timescales, these spawning groups are not significantly differentiated. However, in cluster analysis of otolith microchemistry, these two spawning groups have quite distinct elemental fingerprints – and this pattern is replicated across both edge and core microchemistry (**Fig. 7c**). This suggests that ecologically, Pacific cod which spawn in Jinhae Bay are composed of two distinct groups: early and late spawning fish. However, there appear to be enough migrants between these groups that any distinction is drowned out on evolutionary timescales.

Conclusions

The ability to determine the natal spawning ground of an individual fish makes possible a wide range of studies on fish movement and behavior; this study establishes otolith microchemistry as a tool for such future studies in Pacific cod. A follow-up to this study should include a comparison of assignment tests based on otolith versus genetic data, using discriminant function analysis with jackknifing to assign individuals based on otolith microchemistry. This would allow for more conclusive verification of site fidelity, and allow exploration of potential migrants identified through genetic analysis. Another critical next step is to explore temporal

variation in otolith microchemistry at spawning sites, as variation over time has the potential to drown out signals of spatial variation across spawning years.

The results of this study also support the use of multidisciplinary approaches to study population structure in marine species with complex demographic interactions and dispersal.[39] Otolith and genetic data provided a more comprehensive view of population structure in Pacific cod, which suggested that ecological structure from spawning migrations – such as the early and late spawning groups identified at Jinhae Bay – are not always represented in genetic data. Truly sustainable fisheries management must include both ecological and genetic analysis when accounting for population structure.

Sources

1. Spies, I. & Punt, A. E. The utility of genetics in marine fisheries management: a simulation study based on Pacific cod off Alaska. *Can. J. Fish. Aquat. Sci.* **72,** 1415–1432 (2015).
2. Rideout, R. M. & Tomkiewicz, J. Skipped Spawning in Fishes: More Common than You Might Think. *Mar. Coast. Fish.* **3,** 176–189 (2011).
3. Ashford, J. R., Arkhipkin, A. I. & Jones, C. M. Can the chemistry of otolith nuclei determine population structure of Patagonian toothfish Dissostichus eleginoides? *J. Fish Biol.* **69,** 708–721 (2006).
4. Miller, J. A., Banks, M. A., Gomez-Uchida, D. & Shanks, A. L. A comparison of population structure in black rockfish (*Sebastes melanops*) as determined with otolith microchemistry and microsatellite DNA. *Can. J. Fish. Aquat. Sci.* **62,** 2189–2198 (2005).
5. Collins, S. M. *et al.* Population Structure of a Neotropical Migratory Fish: Contrasting Perspectives from Genetics and Otolith Microchemistry. *Trans. Am. Fish. Soc.* **142,** 1192–1201 (2013).
6. Sturrock, A. M., Trueman, C. N., Darnaude, A. M. & Hunter, E. Can otolith elemental chemistry retrospectively track migrations in fully marine fishes? *J. Fish Biol.* **81,** 766–795 (2012).
7. Canino, M. F., Spies, I. B., Cunningham, K. M., Hauser, L. & Grant, W. S. Multiple ice-age refugia in Pacific cod, Gadus macrocephalus: MULTIPLE ICE-AGE REFUGIA IN PACIFIC COD. *Mol. Ecol.* **19,** 4339–4351 (2010).
8. Cunningham, K. M., Canino, M. F., Spies, I. B. & Hauser, L. Genetic isolation by distance and localized fjord population structure in Pacific cod (Gadus macrocephalus): limited effective dispersal in the northeastern Pacific Ocean. *Can. J. Fish. Aquat. Sci.* **66,** 153–166 (2009).
9. Gwak, W. S. & Nakayama, K. Genetic variation and population structure of the Pacific cod Gadus macrocephalus in Korean waters revealed by mtDNA and msDNA markers. *Fish. Sci.* **77,** 945–952 (2011).
10. Kim, M.-J., An, H. S. & Choi, K.-H. Genetic characteristics of Pacific cod populations in Korea based on microsatellite markers. *Fish. Sci.* **76,** 595–603 (2010).
11. Rand, K. M., Munro, P., Neidetcher, S. K. & Nichol, D. G. Observations of Seasonal Movement from a Single Tag Release Group of Pacific Cod in the Eastern Bering Sea. *Mar. Coast. Fish.* **6,** 287–296 (2014).
12. Helser, T. personal communication. (2016).
13. Baird, N. A. *et al.* Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* **3,** e3376 (2008).
14. Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A. & Johnson, E. A. Local De Novo Assembly of RAD Paired-End Contigs Using Short Sequencing Reads. *PLoS ONE* **6,** e18561 (2011).
15. Drinan, D. P. *et al.* Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (Gadus macrocephalus). *Evol. Appl.* (in press).

16. Catchen, J. C., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. Stacks: building and genotyping loci de novo for short-read sequences. *G3 Genes Genomes Genet.* **1,** 171–182 (2011).
17. Catchen, J. C., Hohenlohe, P., Bassham, S., Amores, A. & Cresko, W. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22,** 3124–2140 (2013).
18. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).
20. Brieuc, M. S. O., Waters, C. D., Seeb, J. E. & Naish, K. A. A Dense Linkage Map for Chinook salmon (*Oncorhynchus tshawytscha*) Reveals Variable Chromosomal Divergence After an Ancestral Whole Genome Duplication Event. *G3amp58 GenesGenomesGenetics* **4,** 447–460 (2014).
21. Kalinowski, S. T., Wagner, A. P. & Taper, M. L. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* **6,** 576–579 (2006).
22. Korkmaz, S., Goksuluk, D. & Zararsiz, G. MVN: An R package for assessing multivariate normality. *R J.* **6,** 151–162 (2014).
23. Oksanen, J. *et al. vegan: Community Ecology Packages.* (2018).
24. McCune, B. & Grace, J. B. Chapter 9: Data Transformations. in *Analysis of Ecological Communities* 67–79 (MjM Software, 2002).
25. Barnett-Johnson, R., Teel, D. J. & Casillas, E. Genetic and otolith isotopic markers identify salmon populations in the Columbia River at broad and fine geographic scales. *Environ. Biol. Fishes* **89,** 533–546 (2010).
26. Standish, J., Sheehy, M. & Warner, R. Use of otolith natal elemental signatures as natural tags to evaluate connectivity among open-coast fish populations. *Mar. Ecol. Prog. Ser.* **356,** 259–268 (2008).
27. Bradbury, I. R., Campana, S. E. & Bentzen, P. Estimating contemporary early life-history dispersal in an estuarine fish: integrating molecular and otolith elemental approaches. *Mol. Ecol.* **17,** 1438–1450 (2008).
28. Clarke, K. R. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* **18,** 117–143 (1993).
29. R Core Team. R: A language and environment for statistical computing. *R Found. Stat. Comput.* (2013).
30. Ward, J. H. & Hook, M. E. Application of an Hierarchical Grouping Procedure to a Problem of Grouping Profiles. *Educ. Psychol. Meas.* **23,** 69–81 (1963).
31. Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. NbClust: An R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* **61,** 1–36 (2014).
32. Galili, T. dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btv428

33. Scornavacca, C., Zickmann, F. & Huson, D. H. Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics* **27,** i248–i256 (2011).
34. Porras-Hurtado, L. *et al.* An overview of STRUCTURE: applications, parameter settings, and supporting software. *Front. Genet.* **4,** (2013).
35. Tibshirani, R., Walther, G. & Trevor, H. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. B* **63,** 411–423 (2001).
36. Matta, M. E., Short, J., Helser, T., Ormseth, O. & Miller, J. Spatial and temporal variation in otolith elemental chemistry of young-of-year Pacific cod in the Gulf of Alaska.
37. Mills, L. S. & Allendorf, F. W. The One-Migrant-per-Generation Rule in Conservation and Management. *Conserv. Biol.* **10,** 1509–1518 (1996).
38. Gwak, W. Re: Cod samples. (2015).
39. Selkoe, K. A., Henzler, C. M. & Gaines, S. D. Seascape genetics and the spatial ecology of marine populations. *Fish Fish.* **9,** 363–377 (2008).

Figures

Figure 1. Map of sampling locations around the Korean peninsula that were used in this study.

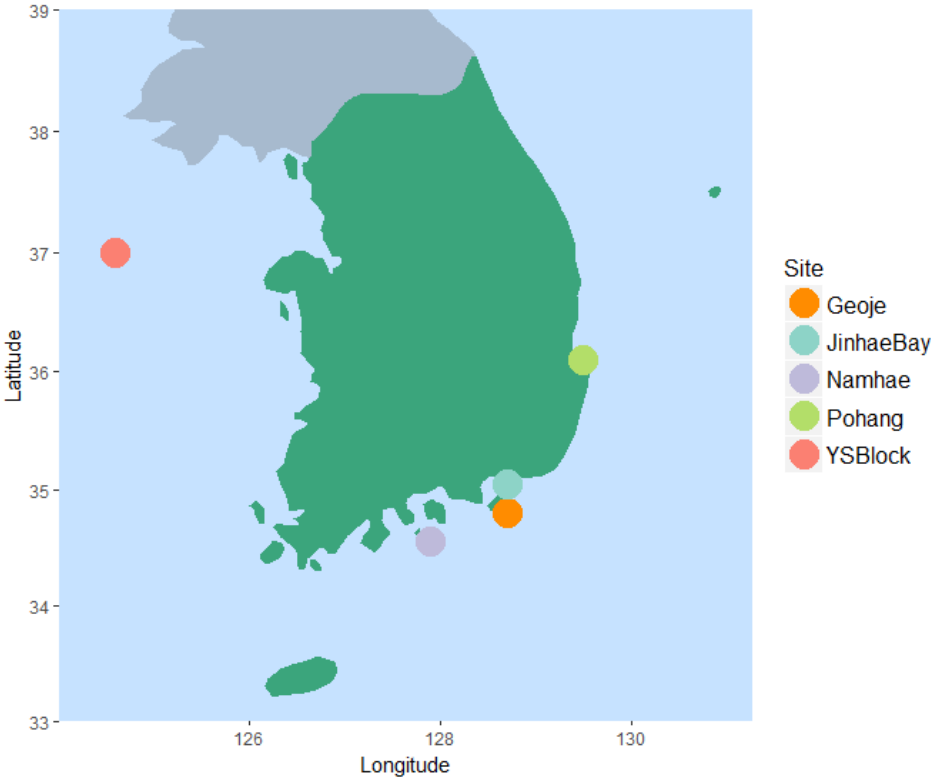


Figure 2. Illustration of LA-ICPMS scan line on Pacific cod otolith. Tick marks delineate winter months during each year of life, used to age the individual. Point marks core of otolith.

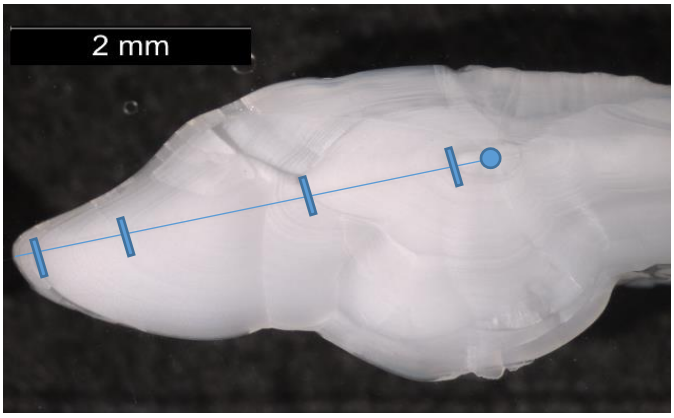


Figure 3. Histograms of element concentration distributions, before relativization. Element concentrations were split by location on otolith (core, ".c" and edge, ".e"). Note non-normal distributions and differing range of x axis values across elements.

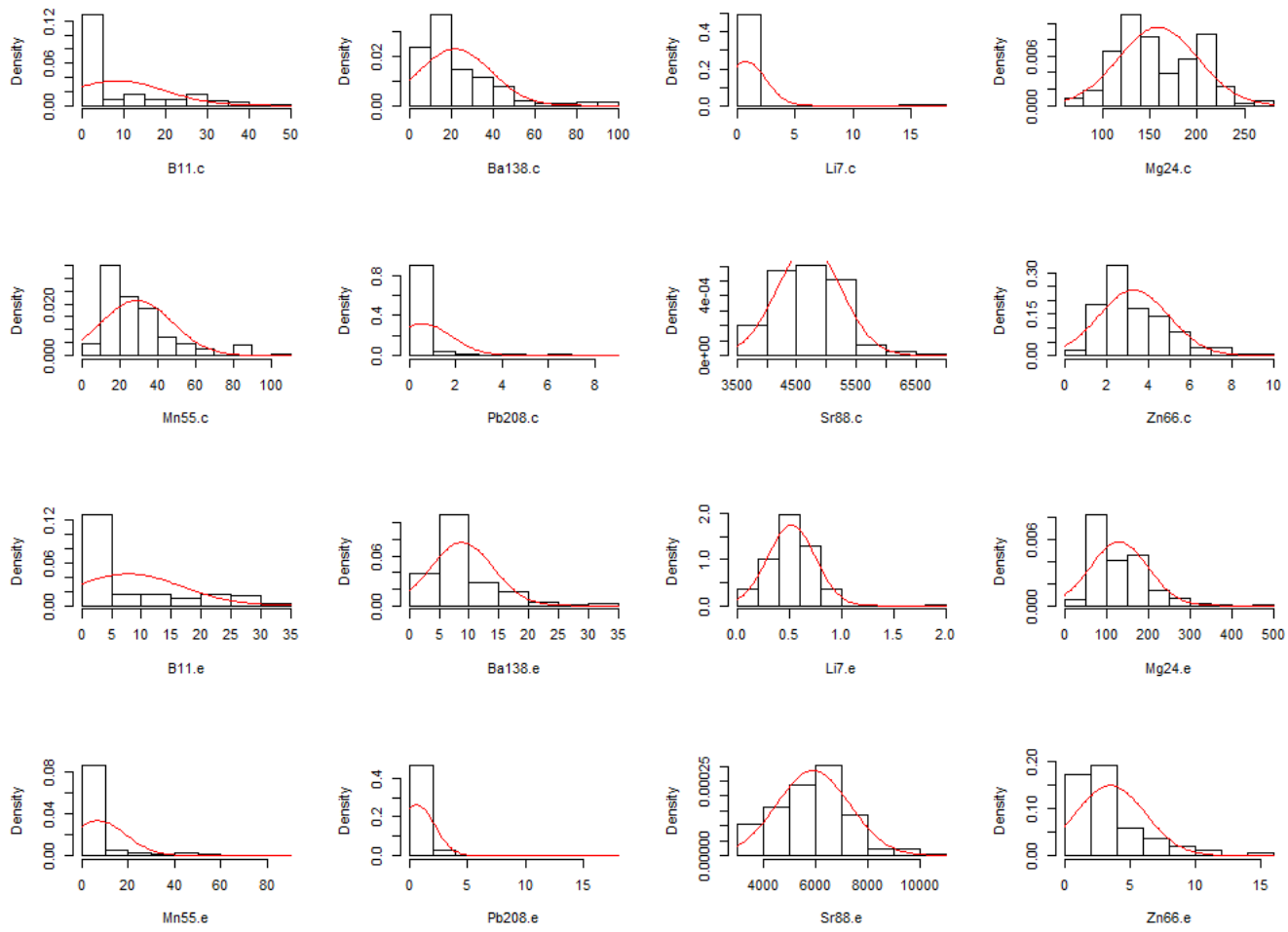


Figure 4. Partition of PERMANOVA Sum of Squares between factor and residuals.

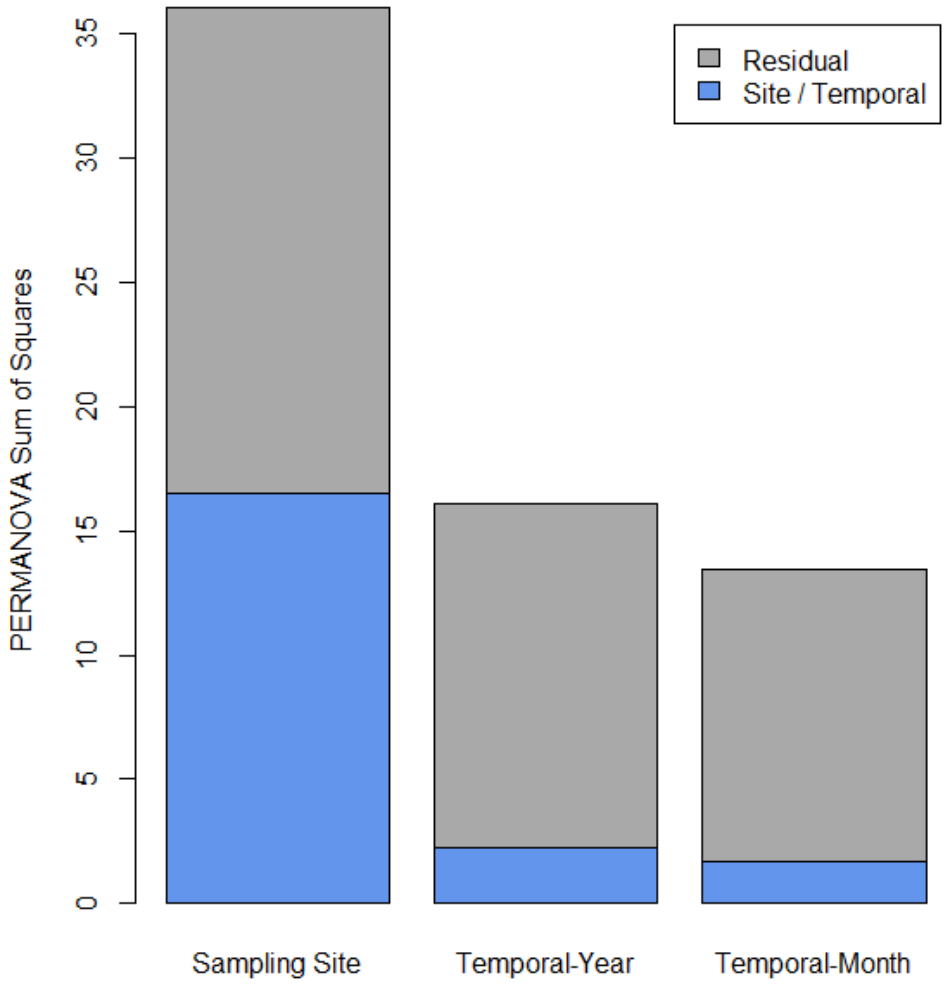


Figure 5. Non-metric Multidimensional Scaling ordination of all samples using edge microchemistry. Ellipses show standard deviation around the centroid of each cluster. Vectors display correlation between specific element loadings and ordination axes. Elements are: boron (B11), barium (Ba138), strontium (Sr88), magnesium (Mg24), manganese (Mn55), zinc (Zn66), lithium (Li7) and lead (Pb208). Inset displays Shepard Plot to evaluate fit of NMDS to data (observed dissimilarity v. observed distance).

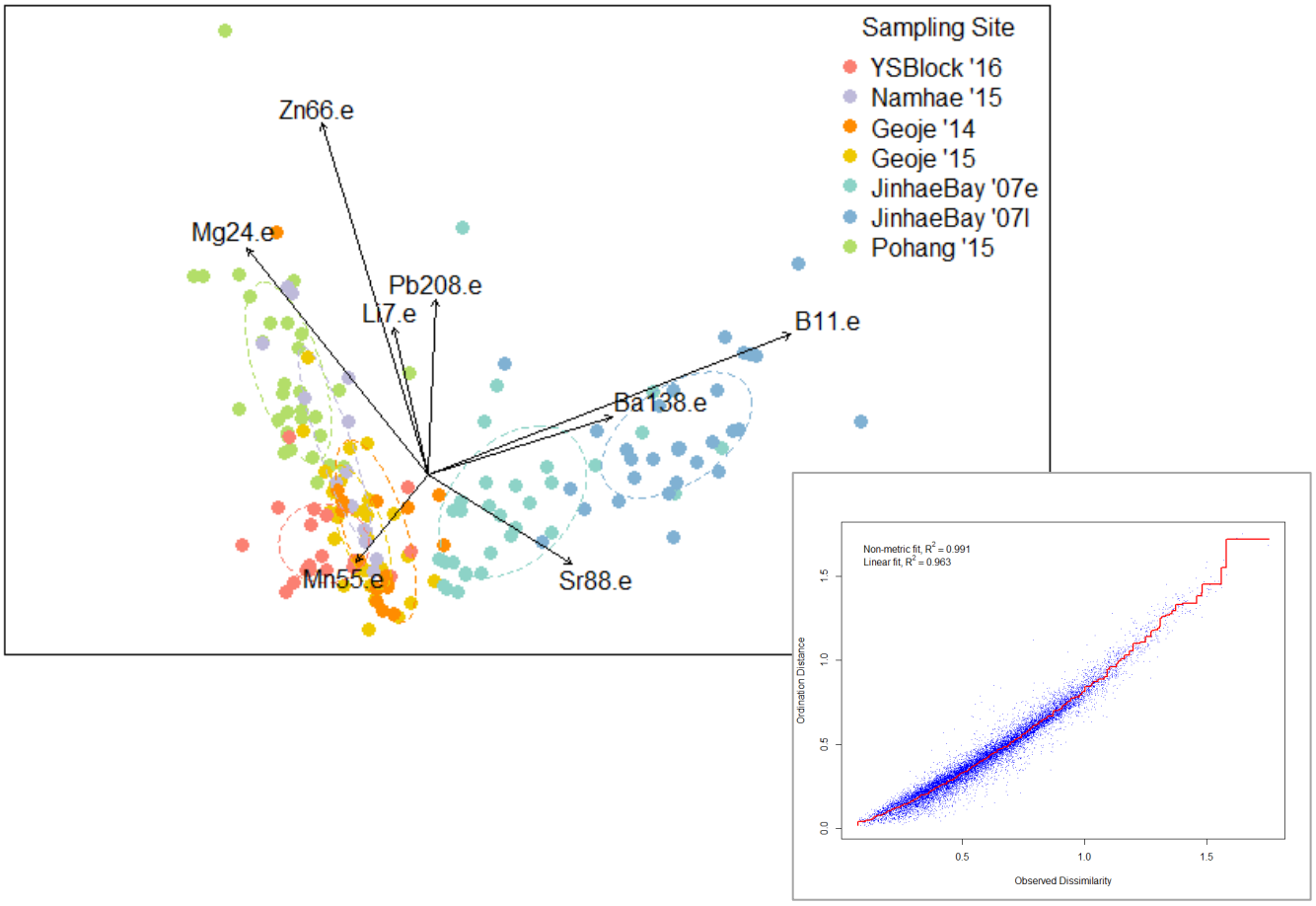


Figure 6. Dendrograms of individuals by sampling site, using the Ward method for cluster analysis. Cluster analysis was completed separately using **(a)** otolith core microchemistry, and **(b)** otolith edge microchemistry. Vertical line at top of dendrograms indicates split identified by gap statistic.

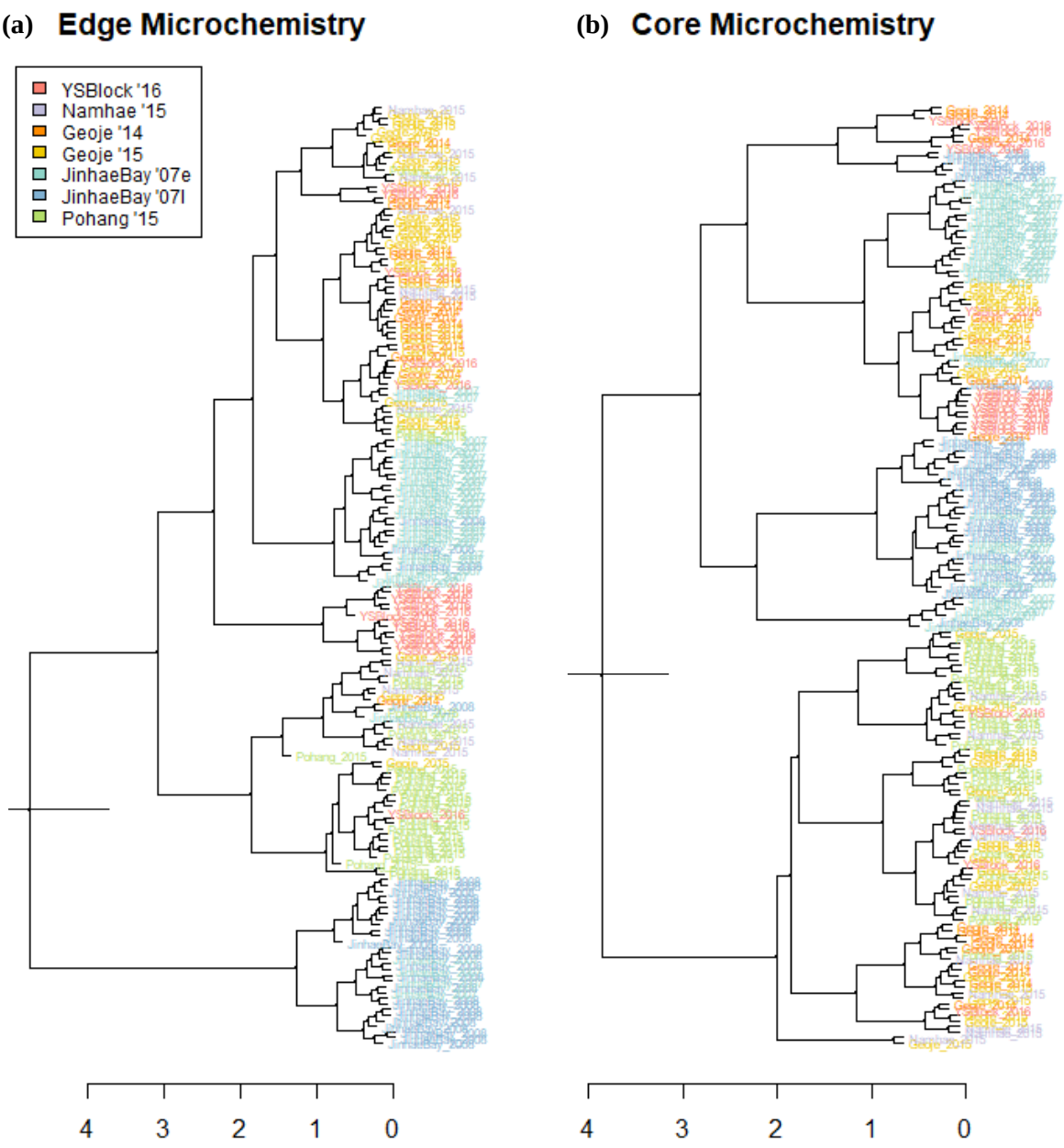


Figure 7. Tanglegram comparing cluster analyses of samples using otolith edge and core microchemistry. Highlighted internal lines show (a) Yellow Sea samples, (b) Pohang samples, (c) Jinhae Bay samples, and (d) Geoje samples.

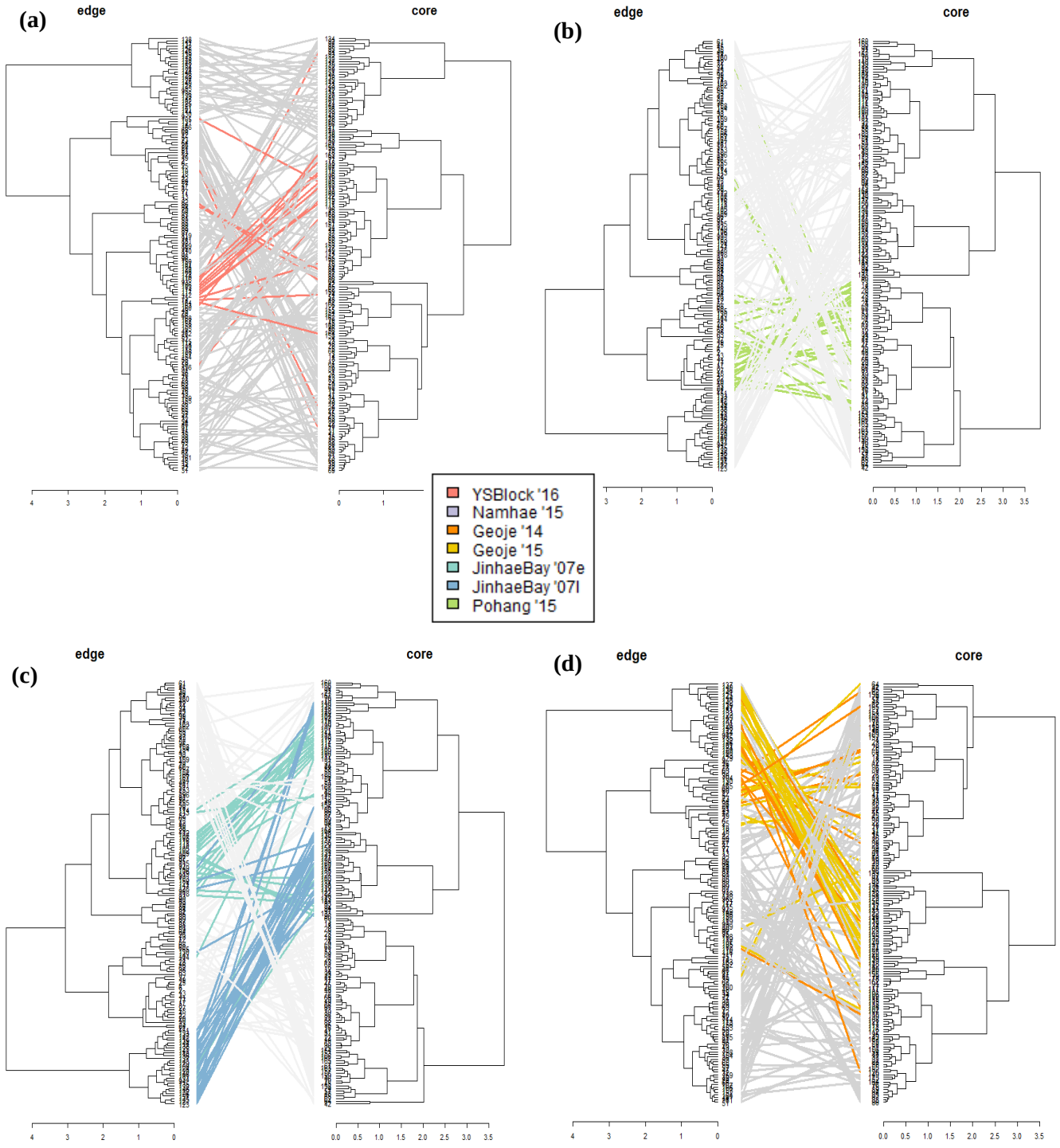


Figure 8. Non-metric Multidimensional Scaling ordination of all samples using otolith edge microchemistry, overlaid with groupings from hierarchical cluster analysis on **(a)** edge microchemistry, and **(b)** core microchemistry.

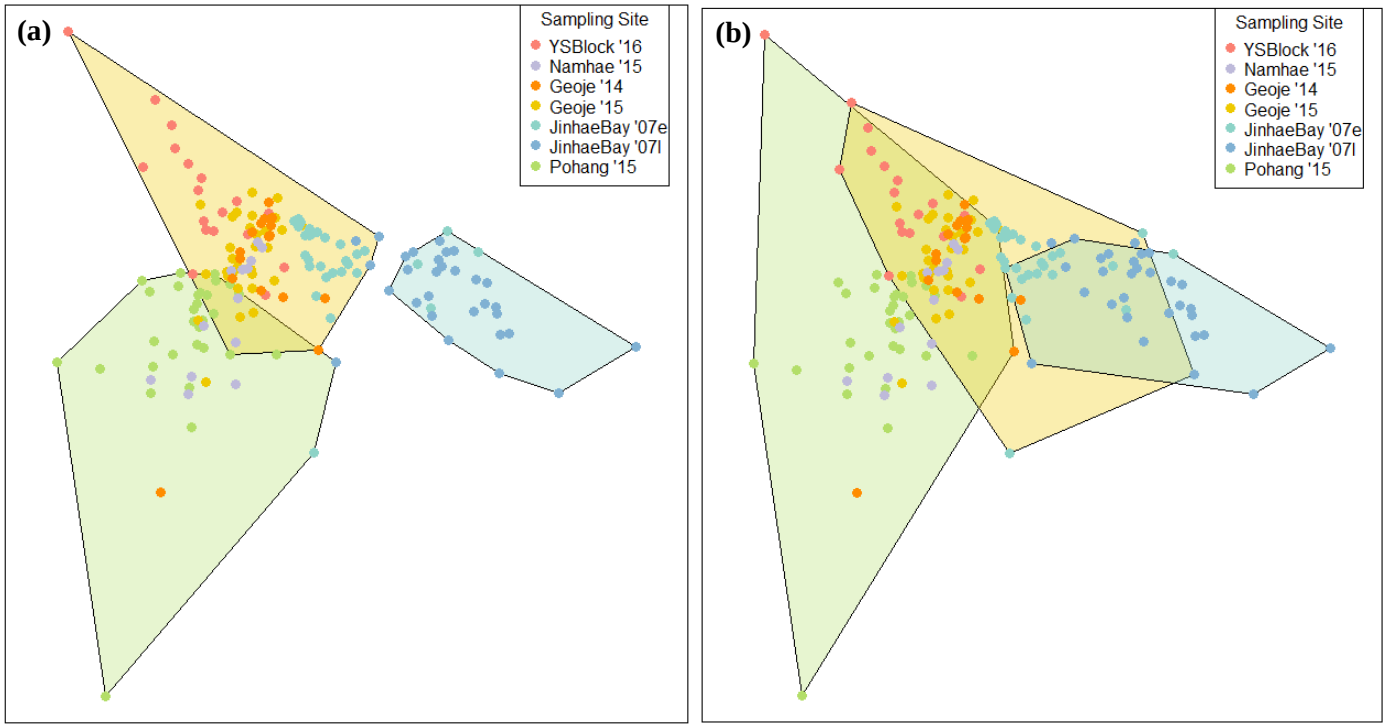
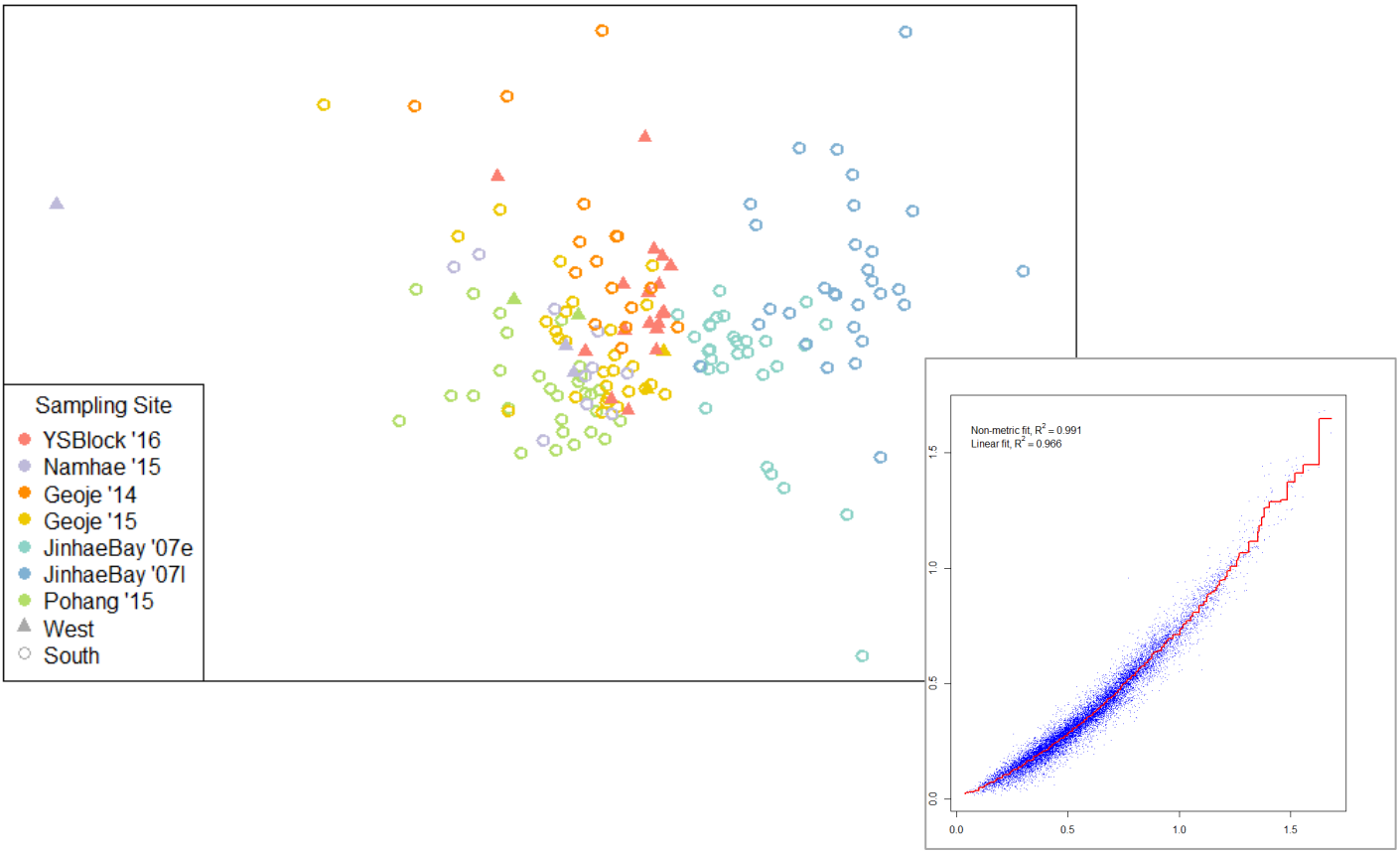


Figure 9. Non-metric Multidimensional Scaling ordination of all samples using otolith core microchemistry, overlaid with groupings from genetic assignment tests. Point color indicates sampling site; point shape indicates whether sample was assigned to the genetically distinct west or south/southeast populations. Inset displays Shepard Plot to evaluate fit of NMDS to data (observed dissimilarity v. observed distance).



Tables

Table 1. Number of samples for each sampling site and spawning season that were collected from around the Korean peninsula, and the number of samples used for this study. Site names correspond to those in Figure 1. *1. Provided by Dr. Sukyung Kang, National Institute of Fisheries Science. 2. Provided by Dr. Wooseok Gwak, Gyeongsang National University.*

Sampling Site	Spawning Season	Tissue Samples Collected	Otolith Samples Collected	Samples Retained
Yellow Sea (YS) Block[1]	2015-2016	31	31	18
Boryeong[2]	*2007-2008*	*22*	*0*	*0*
Namhae[1]	2014-2015	16	16	13
Geoje[1]	2013-2014	22	22	16
Geoje[1]	2014-2015	34	34	30
Jinhae Bay[2]	2007-2008 (Dec.)	46	48	30
Jinhae Bay[2]	2007-2008 (Feb.)	44	48	30
Pohang[1]	2014-2015	31	31	30
Jukbyeon[2]	*2007-2008*	*35*	*0*	*0*

Table 2. PERMANOVA table for test of variation in elemental fingerprints at the edge of the otolith among (a) sampling sites, (b) spawning years at Geoje, 2014 v. 2015, and (c) spawning months at Jinhae Bay, December and February.

(a)

	Degrees of Freedom	Sum of Squares	F	P value
Sampling Site	4	16.389	34.096	1×10^{-5}
Residual	162	19.467		

(b)

	Degrees of Freedom	Sum of Squares	F	P value
Year	1	1.7266	6.4847	2×10^{-5}
Residual	44	11.7150		

(c)

	Degrees of Freedom	Sum of Squares	F	P value
Month	1	2.3093	9.7433	1×10^{-5}
Residual	58	13.7467		

Table 3. One-way ANOVA tables for test of relationships between sampling site and elemental fingerprints at the edge of the otolith.

	Degrees of freedom	Sum of Squares	Mean Sum of Squares	F	P value
Barium					
Sampling site	4	5.10	1.28	5.43	0.000394
Residuals	163	38.27	0.23		
Magnesium					
Sampling site	4	22.73	5.68	48.13	$<2x10^{-16}$
Residuals	163	19.24	0.12		
Strontium					
Sampling site	4	0.86	0.22	17.96	$3.08x10^{-12}$
Residuals	163	1.96	0.01		
Zinc					
Sampling site	4	28.48	7.12	20.59	$9.55x10^{-14}$
Residuals	163	56.37	0.35		

Table 4. One-way ANOVA tables for test of relationships between spawning year at Geoje and elemental fingerprints at the edge of the otolith.

	Degrees of freedom	Sum of Squares	Mean Sum of Squares	F	P value
Barium					
Spawning Year	1	0.125	0.125	0.579	0.451
Residuals	44	9.52	0.216		
Magnesium					
Spawning Year	1	3.05	3.05	24.83	$1.02x10^{-5}$
Residuals	44	5.40	0.123		
Strontium					
Spawning Year	1	0.034	0.034	2.592	0.115
Residuals	44	0.574	0.013		
Zinc					
Spawning Year	1	0.002	0.002	0.005	0.944
Residuals	44	17.93	0.408		

Table 5. One-way ANOVA tables for test of relationships between spawning months at Jinhae Bay and elemental fingerprints at the edge of the otolith.

	Degrees of freedom	Sum of Squares	Mean Sum of Squares	F	P value
Barium					
Spawning Month	1	4.68	4.68	27.33	$2x10^{-6}$
Residuals	59	10.10	0.171		
Magnesium					
Spawning Month	1	0.109	0.109	1.15	0.288

Residuals	59	5.59	0.095		
Strontium					
Spawning Month	1	0.019	0.019	1.53	0.22
Residuals	59	0.737	0.012		
Zinc					
Spawning Month	1	0.337	0.337	0.798	0.375
Residuals	59	24.90	0.422		

Table 6. Stress values from NMDS of edge data.

Dimensions (K)	**Stress Value**
1	0.2559
2	0.1422
3	0.0934
4	0.0609
5	0.0390
6	0.0255
7	0.0122

Table 7. Stress values from NMDS of core data.

Dimensions (K)	**Stress Value**
1	0.2478
2	0.1372
3	0.0933
4	0.0656
5	0.0401
6	0.0240
7	0.0105

Appendix A - Metadata

This data set consists of genetic and otolith microchemistry data for Pacific cod, to be used to determine stock structure of Pacific cod around the Korean peninsula. Mature Pacific cod were sampled from winter spawning aggregates in coastal South Korean waters over the course of several spawning seasons. For each individual, a fin clip or tissue sample was taken and preserved in 100% denatured ethanol, and either one or both otoliths (ear bones) collected and stored dried. Samples were collected by Dr. Sukyung Kang of the National Institute of Fisheries Science (spawning seasons 2014 – 2016), and Dr. WooSeok Gwak of Gyeongsang University (2007-2008 spawning season). Tissue samples from the 2007-2008 spawning season are subsamples from Gyeongsang National University.

The spawning aggregates span three regions of South Korean coastal waters: the western coast (2 sites), southern coast (4 sites), and eastern coast (1 site). The data also includes temporal replicates; one site was sampled during both the 2013-2014 and 2014-2015 spawning seasons to test between-year differences, and one site sampled at the beginning and end of the 2007-2008 spawning season to test within-year differences.

Raw genetic data is currently stored on a fireproof hard drive in the Hauser Lab, at the University of Washington, Seattle WA. Raw otolith data is saved on desktop computers in the Age & Growth Lab of the Alaska Fisheries Science Center, Seattle WA. Processed genetic and otolith data are both available on Github: link to genetic data, link to otolith data. Original tissue and otolith samples are stored in the Molecular Ecology Research Lab, at the University of Washington, Seattle WA.

Original metadata collected with the samples (listed below in *explanatory variables* table) is summarized in a google spreadsheet, available by request from Mary Fisher (mfisher5@uw.edu).

Explanatory Variables

Site	Sampling Site
Season	Sampling Spawning Season
Region	Sampling Region (West / South / East)
TL	Total length of fish (cm)
W	Weight of fish (g)
GW	Gonad weight of fish (g)
Sex	Sex of fish (M/F)*
Maturity	Sexual maturity of fish at time of capture*
Age	Age of fish, as determined from otolith (years)
OW	Otolith weight (g)

* Low confidence that these variables were correctly measured in the field

Response Variables

1. Genetic markers
    a. 5,405 single nucleotide polymorphisms (SNPs), recorded as "Locus_"
2. Otolith chemical markers
    a. All chemicals are expressed as a *relative* concentration to a Calcium standard

Abbreviation	Element
B11	Boron
Ba138	Barium 138
Li7	Lithium 7
Mg25	Magnesium 25
Mn55	Manganese 55
Pb208	Lead 208
Si86	Silicon 86
Zn66	Zinc 66

Individual Codes

Each individual sample is coded according to sampling site, sampling date. Following the code is a specific numerical identifier for that sample, and any appropriate suffixes (see next section).

Abbreviation	Sampling Site Location	Sampling Date	Analysis Group
PO010715	Pohang	1/7/2015	Pohang 2015
PO020515	Pohang	2/5/2015	Pohang 2015
GE011215	Geoje	1/12/15	Geoje 2015
GE012315	Geoje	1/23/2015	Geoje 2015
GEO020414	Geoje	2/4/2014	Geoje 2014
YS121316	Yellow Sea Block 161	12/13/2016	Yellow Sea 2016
JB121807	Jinhae Bay	12/18/2007	Jinhae Bay 2007, Early Spawn
JB021108	Jinhae Bay	2/11/2008	Jinhae Bay 2007, Late Spawn
JUK07	Jukbyeon	12/10/2007	Jukbyeon 2007
BOR07	Boryeong	1/24/2007	Boryeong 2007

Appendix B - Scripts

All R scripts used for this study can be found in the "scripts" folder in the Otolith Analyses directory of Mary Fisher's Github page. Additionally, downloading the Otolith Analyses directory or branching the PCod-Korea-repo will provide the file structure and data needed to replicate this study's analyses in R.