**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**

Inclusion of population structure in marine fisheries management protects local populations from overexploitation, and allows for a more sustainable yield. Population structure is often assessed from a genetic perspective; however, the actual distribution and migration of individuals which contribute to these evolutionary patterns is important for fishery managers who work primarily on ecological timescales. We place genetic population structure in the context of individual life histories for Korean Pacific cod (*Gadus microcephalus*) stocks by comparing otolith microchemistry data with existing population genetic analysis. We first established that profiles composed of eight elements differed significantly between sampling sites by using otolith edge microchemistry. Specific elements could be used to differentiate between sampling sites in Jinhae Bay, Pohang, and the Yellow Sea in a non-metric Multidimensional Scaling analysis. This proves that otolith microchemistry can be useful in an entirely marine setting, where it is commonly assumed that bodies of water will not significantly differentiate by element profiles. A Mantel test showed that edge microchemistry was significantly correlated with core microchemistry, suggesting that most individuals were caught on spawning grounds with similar water chemistry as their natal spawning ground. Hierarchical cluster analysis…

We conclude that…

**Introduction**

1. Use of population structure in marine fisheries management
   1. 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 (Spies & Punt 2015).
   2. Providing an ecological context for the observed genomic patterns is also important for fisheries management, as it allows additional insight into the drivers of these patterns on ecological timescales.
   3. Multidisciplinary approaches are particularly important for the management of marine species, which are typically associated with high gene flow and have complex demographic interactions (Selkoe et al. 2008).
2. Pacific cod is an excellent case study for combining these two types of data sets, as its biology directly impacts population structure, but much is unknown about the specifics of individual fishes’ life histories.
   1. Tagging studies and prior genetic analysis, primarily on US Pacific cod populations, suggest that Pacific cod display site fidelity to spawning aggregates. Much of the tagging evidence for this behavior comes from a mark-recapture study conducted by Rand et al. (2014)7 in the northwestern United States. Rand et al. (2014) found that during the summer and fall, Pacific cod were more widely distributed at lower densities, traveling up to 600km from the spawning aggregation where they were released. Pacific cod were significantly more likely to be found within 100km of their original spawning aggregation during the spring spawning season, than during the fall.
   2. Prior studies using microsatellite and RAD loci suggest limited mixing on evolutionary scales between Pacific cod spawning aggregates in the western half of their north Pacific range8–10.
3. 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, to prove limited mixing between black rockfish populations, and to provide complementary information on the natal origin and genetic structure of a migratory cyprinid species. This study was the first known to the authors to use next-generation sequencing methods, which generate thousands of loci for finer-scale genetic population structure analysis, compared to the 20 or fewer microsatellite or allozyme loci used in previous studies.

The purpose of this study was to explore Pacific cod spawning population structure and site fidelity around the Korean peninsula according to otolith microchemistry, and to then compare these results from otolith microchemistry with genetic population structure. In doing so, we address three key questions:

1. How different are otolith microchemical concentrations across Pacific cod spawning grounds around the Korean peninsula?
2. Is there evidence suggesting that Pacific cod return to their natal spawning ground during the spawning season in which they were caught?
3. Does the genetic population structure that developed over evolutionary timescales align with spatial spawning patterns and site fidelity determined from individual life histories?

**Methods**

*Sample Collection*

A total of 322 fin clips and tissue samples were collected from ten Pacific cod spawning aggregates in the coastal waters around the Korean peninsula (Table, Figure). Three aggregates were sampled during the 2007 – 2008 winter spawning season, three aggregates during the 2014-2015 season, and one each during the 2015 – 2016 season. These do not include sample groups taken as temporal replicates, one between-year replicate during the 2013-2014 season, and one within-year replicate during the 2007-2008 season. Fin clips and tissues were preserved in 95-100% non-denatured ethanol, stored at 4C.

*Otolith Analysis*

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. Element concentrations were then standardized against a calcium standard. Eight elements were found to be informative and retained for further analysis: boron (B11), barium (Ba138), Lithium (Li7), magnesium (Mg24), manganese (Mn55), lead (Pb208), strontium (Sr88) and zinc (Zn66). Otoliths were then aged and each year of growth partitioned into 4 additional sections. 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).

*RAD sequencing and SNP genotyping*

DNA was extracted from 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) and Etter et al. (2011), with modifications to incorporate Agencourt AMPure XP beads (Beckman Coulter, Inc., Pasadena, CA) for purification (Gruenthal et al. in prep). 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 pipeline (Catchen et al. 2011; Catchen et al. 2013), Bowtie (Langmead et al. 2009), and NCBI’s Basic Local Alignment Search Tool, BLAST (Altschul et al. 1990), according to the procedures outlined in Gruenthal et al. (in prep) and Brieuc et al. (2013). A genotype file containing putative polymorphic SNPs present in ≥ 80% of fish per spawning aggregate was 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 (Kalinowski et al. 2006).

*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 Pacific cod spawning behavior.

The size of the data set was further reduced for computation simplicity and to achieve more even sample sizes. Individuals were randomly removed from both Jinhae Bay 2007-2008 early and late spawning groups, so that the number of samples in each group reached n = 30.

*Statistical analysis of otolith microchemistry data*

Microchemistry data normality using Mardia’s multivariate normality test and QQPlot visualization with the R package MVN (cite). The data did not conform to multivariate normality. Data transformations, including natural log, log10, 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 and sampling site showed markedly different scales (Figure 1), so each element’s concentration ratios were relativized by maxima using the R package vegan3. 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 differed significantly between the sampled Pacific cod spawning grounds, I ran a PERMANOVA on the elemental ratios at the edge of the otolith against a combination variable of sampling site and spawning year, with 100,000 permutations. To further explore differences between sites and years for each individual element, an ANOVA was conducted for elements that conformed, or could be transformed, to the assumed normal distribution. This consisted of four elements: Barium138 (natural log-transformed), Magnesium24 (natural log-transformed), Strontium88, Zinc66 (natural log-transformed). I then ran a non-metric Multi-dimensional Scaling Analysis (NMDS) on elemental ratios from the edge of the otolith, and fit element vectors to the ordination using vegan3, to identify which elements were correlated with the separation of specific sampling sites across the ordination space. Although this step is commonly completed by conducting a Linear Discriminant Analysis4–6, NMDS has been used when data do not conform to multivariate normality7. The NMDS was run with the wrapper function metaMDS in vegan3, with a maximum of 400 iterations and minimum of 40 runs (McCune & Grace 2002). In order 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, which was used for the final NMDS.

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. Hierarchical cluster analysis was run to explore groups of individuals based on otolith microchemistry at the core and the edge of the otolith. With NbClust8, a gap index was calculated based on the Ward clustering method (ward.d2 in R) 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 each analysis using the R package dendextend (Galili 2015). 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 (Scornavacca et al. 2011). The tanglegram function was called from the R package dendextend (Galili 2015).

*Analysis of genetic population structure and individual assignment*

The number of genetically distinct populations in the data was determined using STRUCTURE9. 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 individuals9. STRUCTURE was run using default parameters (burnin and MCMC reps = 10,000) 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 visually assess the groupings of individuals according to genetic assignment and otolith core microchemistry. The core of the otolith was used to generate a dendrogram because it represents the natal spawning ground of that individual. By color-coding dendrogram leaves by sampling site, we can therefore indicate how that individual’s location may have changed, or remained the same, between the natal and final spawning seasons. This suggests individual migration or site fidelity according to microchemistry data, which can then be matched to the genetic assignment of that individual. Genetic assignment of individuals was plotted in a heatmap associated with the otolith core dendrogram, using the package gplots (Warnes et al. 2016).

**Results**

*Microchemistry of Spawning Grounds*

A permutational analysis of variance (PERMANOVA) of microchemical data by sampling site, sampling year (temporal sample at Geoje), and sampling month (temporal sample at Jinhae Bay) showed significant differences in elemental fingerprints between sampling sites, and spawning year and month (*p* < 0.01; **Table**). However, the amount of variability in elemental fingerprints that could be accounted for by each of these factors, rather than by the residual, differed across factors (**Figure**)... 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. (barium, magnesium, strontium, and zinc). All four ANOVAs were highly significant (*p* < 0.01; **Table**); however, the amount of variability in the element concentration that could be contributed to sampling site, spawning year, and spawning month was quite low…

A non-metric Multi-dimensional Scaling (NMDS) ordination provided insight into how element concentrations differed between sampling units. The NMDS, run with seven dimensions, produced a stress value of 0.01215, which is associated with excellent representation of the data set (Clarke 1993), and a linear fit of R2=0.999. The ordination showed distinct separation between the Jinhae Bay sampling site (both early and late in the 2007-2008 spawning season) and the remainder of the data set (**Fig. 3**). 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 this site. Samples from Pohang (2014-2015 spawning season) also appear to differentiate across the ordination space; this differentiation is strongly correlated with magnesium (Mg24) and zinc (Zn66) concentrations, and somewhat correlated with lithium (Li7) and lead (Pb208). Although Geoje temporal samples (2014 and 2015), which were found to have significantly different elemental fingerprints in the PERMANOVA, did not differentiate across the ordination space, this could be a result of other samples being much more highly differentiated (i.e. Pohang, Yellow Sea Block, and Jinhae Bay), or of variation of only two dimensions being represented in the NMDS visualization.

*Site Fidelity*

A Mantel test between distance matrices calculated from otolith core concentrations and otolith edge concentrations suggested slight positive correlation (p = 1 x 10-5, r = 0.3372) between full elemental fingerprints of each individual’s natal spawning ground and the spawning ground at which they were caught.

Hierarchical cluster analysis analyzed with the Tibshirani et al. (2001) gap statistic suggested only two clusters in both the otolith edge and otolith core datasets (**Fig. 5a&b**). The first cluster is composed entirely of Jinhae Bay samples, although Jinhae Bay samples could also be found in the second cluster. A tanglegram to visually compare the two suggests that…

* Jinhae Bay individuals…
* The dendrogram of the otolith core data seems to be less segregated by sampling site in the second, larger cluster than the otolith edge data.

(to be honest, I had a difficult time getting much useful information out of the dendrograms / cluster analysis as a whole.)

*Genetic population structure*

STRUCTURE analyses suggested two genetically distinct populations, delineated between the western and southern 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 southern coasts which had the majority of their genetic material assigned to the western coast baseline: two fish from Pohang, two from Geoje (one from the 2013-2014 season, one from the 2014-2015 season) and two from Namhae.

*Genetic and otolith microchemistry comparisons*

* Dendrogram with heat map showed some similarity between otolith core clustering and genetic population assignment
  + Most yellow Sea samples, from the genetically distinct western coast, were clustered together in a subtree on the dendrogram
  + Perhaps most interestingly, five out of the six individuals from the southern sites which were identified as most genetically similar to western samples did *not* cluster in the subtree with the majority of western samples. However, given that no hierarchical structure within the second major cluster of this dendrogram was found to be significant using the gap statistic, this clustering pattern should be approached with caution.
* Otolith core microchemistry offered more differentiation between sites in the same region, particularly Jinhae Bay, than the genetic data.

**Discussion**

1. Otolith microchemistry has been primarily used to discern fish migratory patterns in estuary systems or in anadromous species. In this study, we prove that microchemistry can be useful in discerning between marine bodies of water, particularly along the coast.
   1. Pohang samples loaded high on heavy metals such as zinc and lead, likely due to runoff from this highly industrial port.
   2. Jinhae Bay samples loaded high on strontium, barium and boron. The southern coast of Korea has fjord-like bays and inlets; this site was further into the bay than the other southern sampling sites, Geoje and Namhae. This would explain why elements associated with salinity profiles differentiate Jinhae Bay samples. The difference in salinity profile 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.
   3. Yellow Sea samples loaded high on Manganese. 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.
2. This study adds to the literature suggesting that Pacific cod display site fidelity to natal spawning aggregates. However, this study was limited by a lack of multivariate normality in the data set; future research efforts 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.
   1. The hierarchical cluster analysis of edge versus core microchemistry displayed unexpected results. The NMDS ordination suggested that it would be possible to discern between four groups of sampling sites; however, cluster analysis only showed two significant clusters based on the gap statistic. While the first cluster consisted solely of Jinhae Bay samples for both edge and core otolith data, which might be expected based on their clear separation in NMDS ordination space, some Jinhae Bay samples were also assigned to the second cluster.
3. Genetic population structure v. otolith microchemistry
   1. The ability to discern between sampling sites varied between genetic and otolith microchemistry data. Genetic analysis through STRUCTURE was able to discern between the eastern, western, and southern coasts. Otolith microchemistry analysis through hierarchical clustering was only able to discern between two groups: the majority of Jinhae Bay samples, and all other sampling sites. However, NMDS ordination suggested that otolith microchemistry may be used to differentiate between the southern and western sampling sites, while also being able to distinguish between sites on the same coast. Perhaps hierarchical cluster analysis was not the most correct way to approach clustering of otolith microchemistry data?
   2. Neither genetic nor otolith microchemistry data showed major differences between years, although Jinhae Bay samples taken during the latter half of the 2007-2008 spawning season were more separated from other sites in the NMDS ordination than those taken during the beginning half of the season. This suggests that certain elemental profiles in Jinhae Bay may fluctuate throughout the spawning season.
   3. Selkoe et al. (2008) suggest that when genetic and otolith microchemistry data show little correlation, comparisons can still prove valuable by targeting ambiguous results in genetic data through microchemistry analysis. We see this in practice in this study, where individuals flagged as potential migrants in the genetic data were also scattered among southern sites in the otolith core cluster analysis. Again, this must be interpreted carefully.

**Sources**

1. Helser, T. (2016).

2. 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).

3. Oksanen, J. et al. *vegan: Community Ecology Packages.* (2018).

4. 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).

5. 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).

6. 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).

7. 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).

8. 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).

9. Porras-Hurtado, L. *et al.* An overview of STRUCTURE: applications, parameter settings, and supporting software. *Front. Genet.* **4,** (2013).

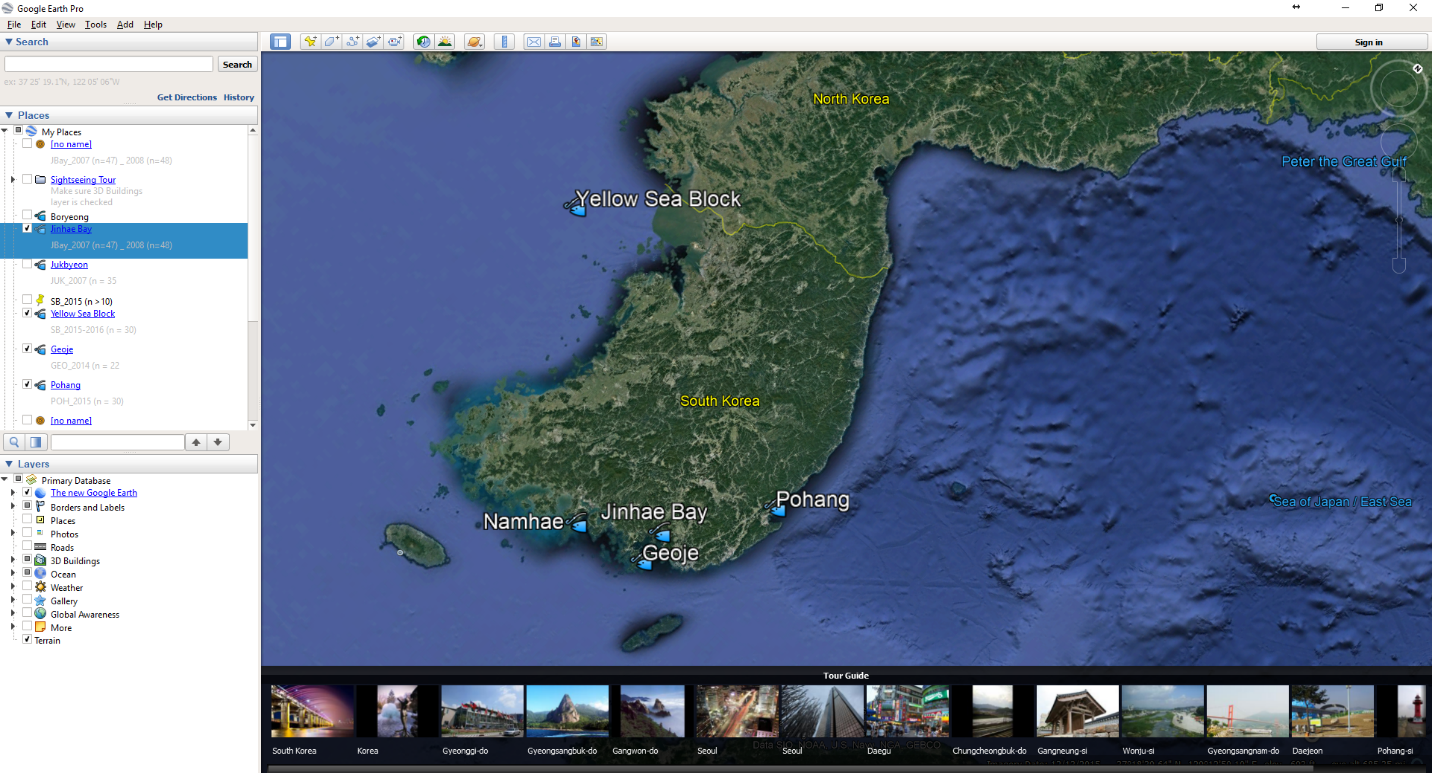
**Tables**

**Table 1.** Number of samples for each sampling site and spawning season \*that were used in this study.\* Site names correspond to those in **Fig. 1**.

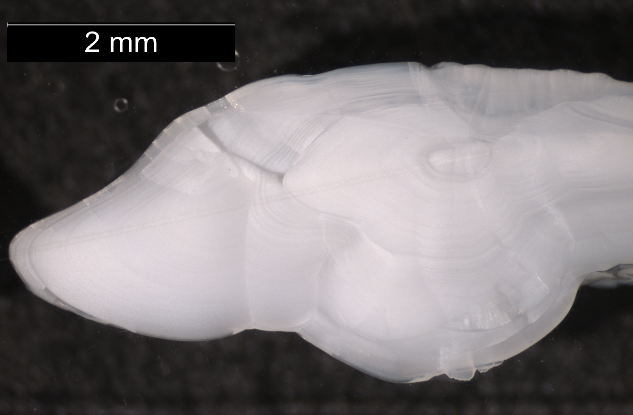
|  |  |  |
| --- | --- | --- |
| **Sampling Site** | **Spawning Season** | **N. Samples** |
| Yellow Sea | 2015-2016 | 18 |
| Namhae | 2014-2015 | 13 |
| Geoje | 2013-2014 | 16 |
| Geoje | 2014-2015 | 30 |
| Jinhae Bay | 2007-2008 (early) | 30 |
| Jinhae Bay | 2007-2008 (late) | 30 |
| Pohang | 2014-2015 | 30 |

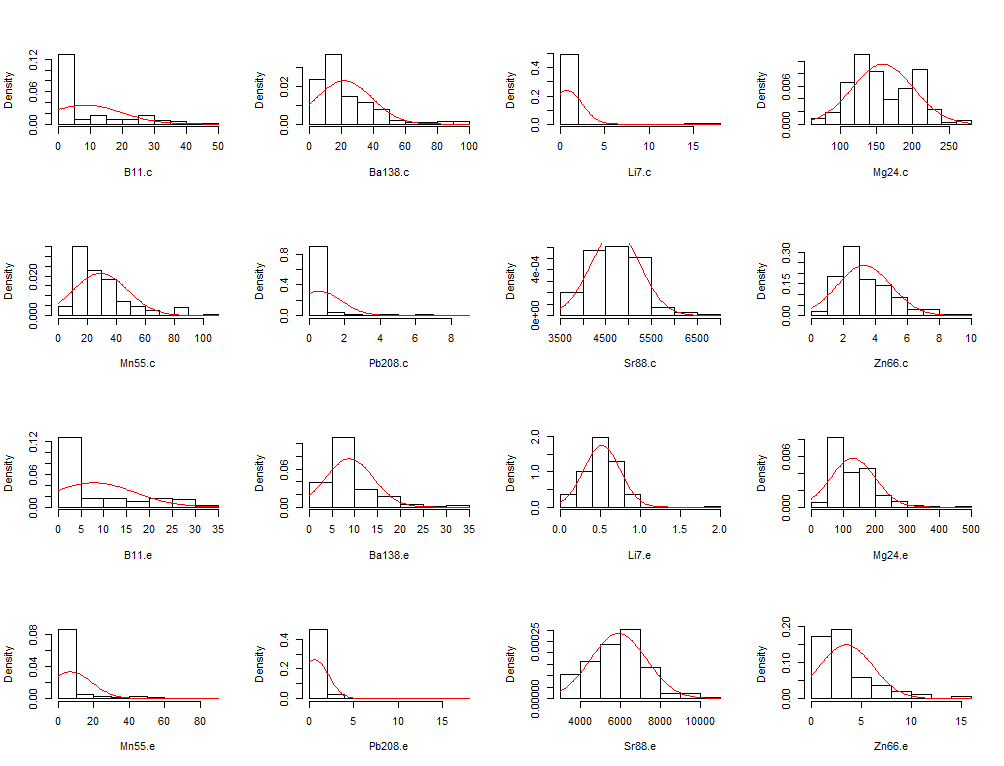
**Figures**

**Figure 1.** Map of sampling locations around the Korean peninsula.

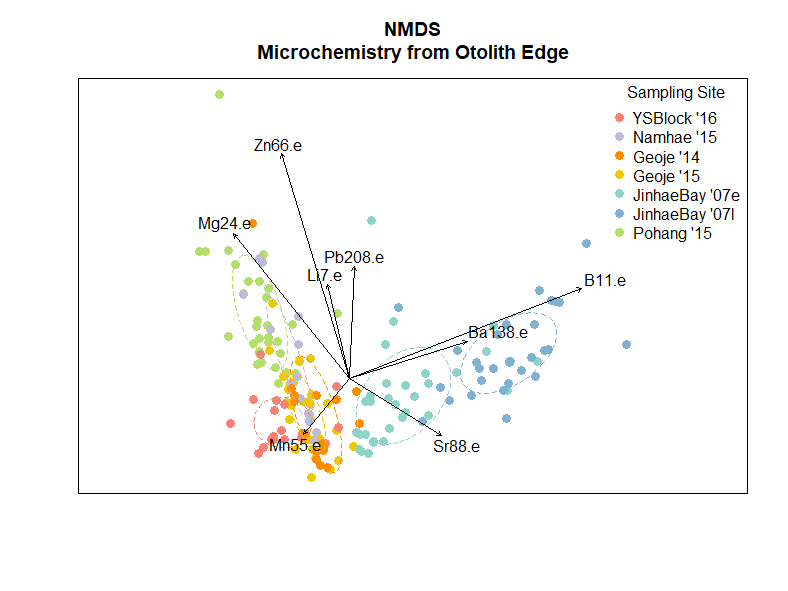
****

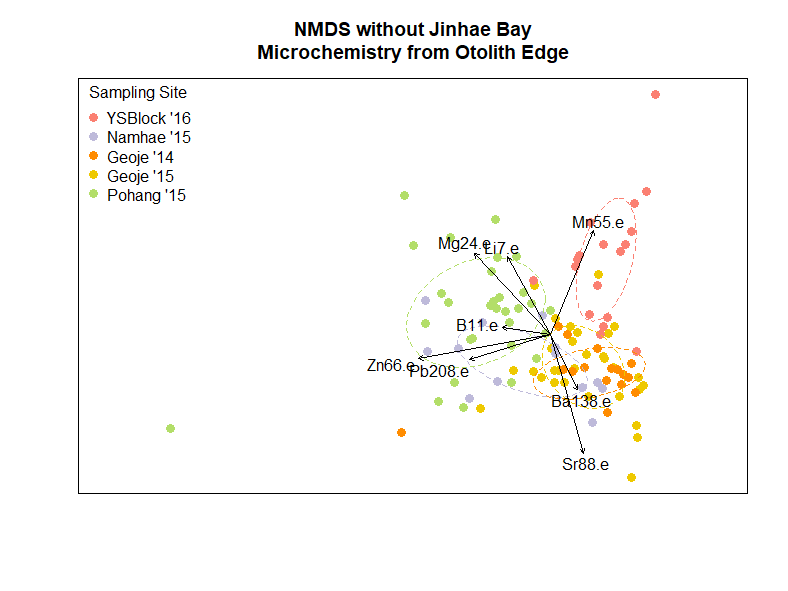
**Figure 2.** Illustration of LA-ICPMS scan line on Pacific cod otolith. Tick marks delineate winter months during each year of life, used to age the individual.



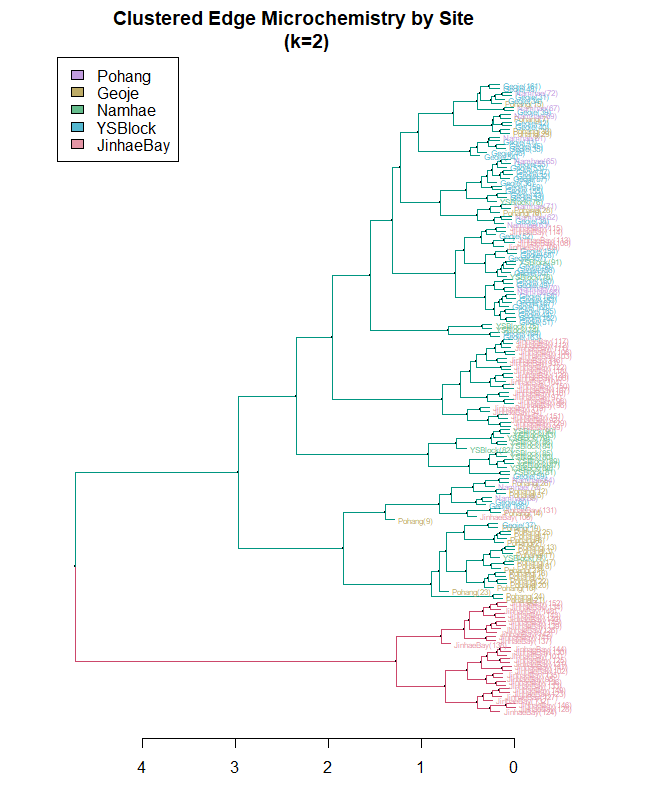
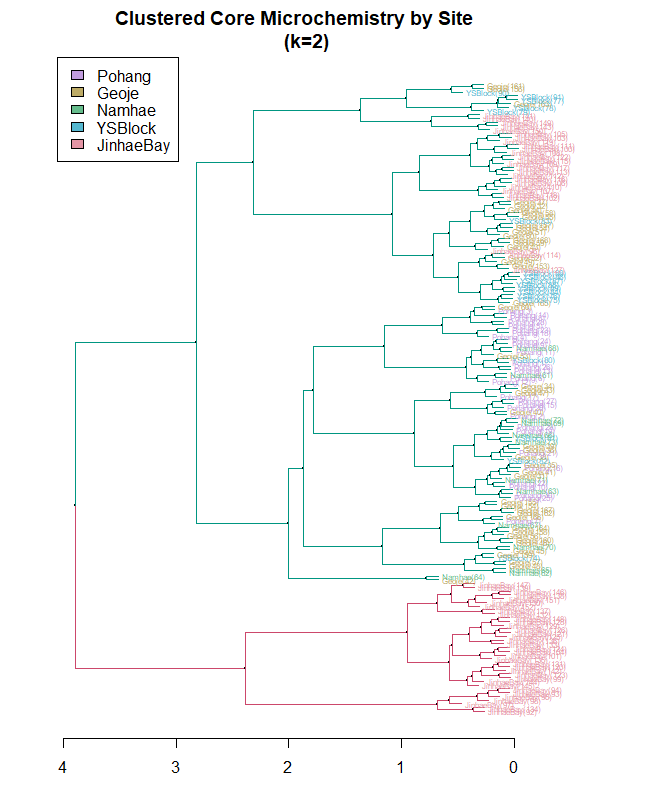
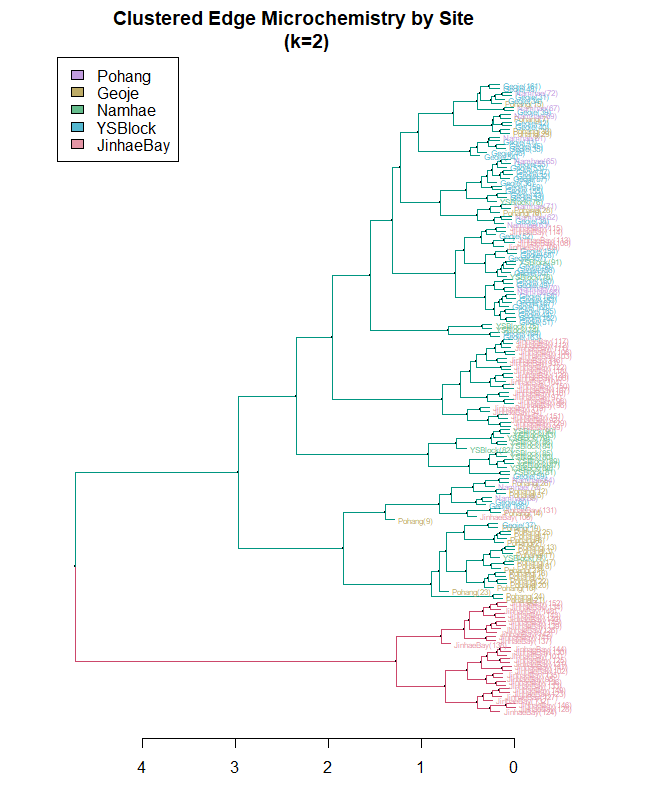
**Figure 3.** Histograms of element concentration distributions, before relativization. Element concentrations were split by location on otolith (core, “.c” and edge, “.e”). Note non-normal distributions and differing range of x axis values across elements.

**Figure 4.** Non-metric Multidimensional Scaling ordination of all samples. 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).



**Figure 5.** Non-metric Multidimensional Scaling ordination excluding Jinhae Bay samples. 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).

**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. Two sets of branches are highlighted in accordance with the “best cluster” determination from the gap statistic. Leaves are color-coded according to sampling site. *Scale will be the same for both graphs in final paper.*

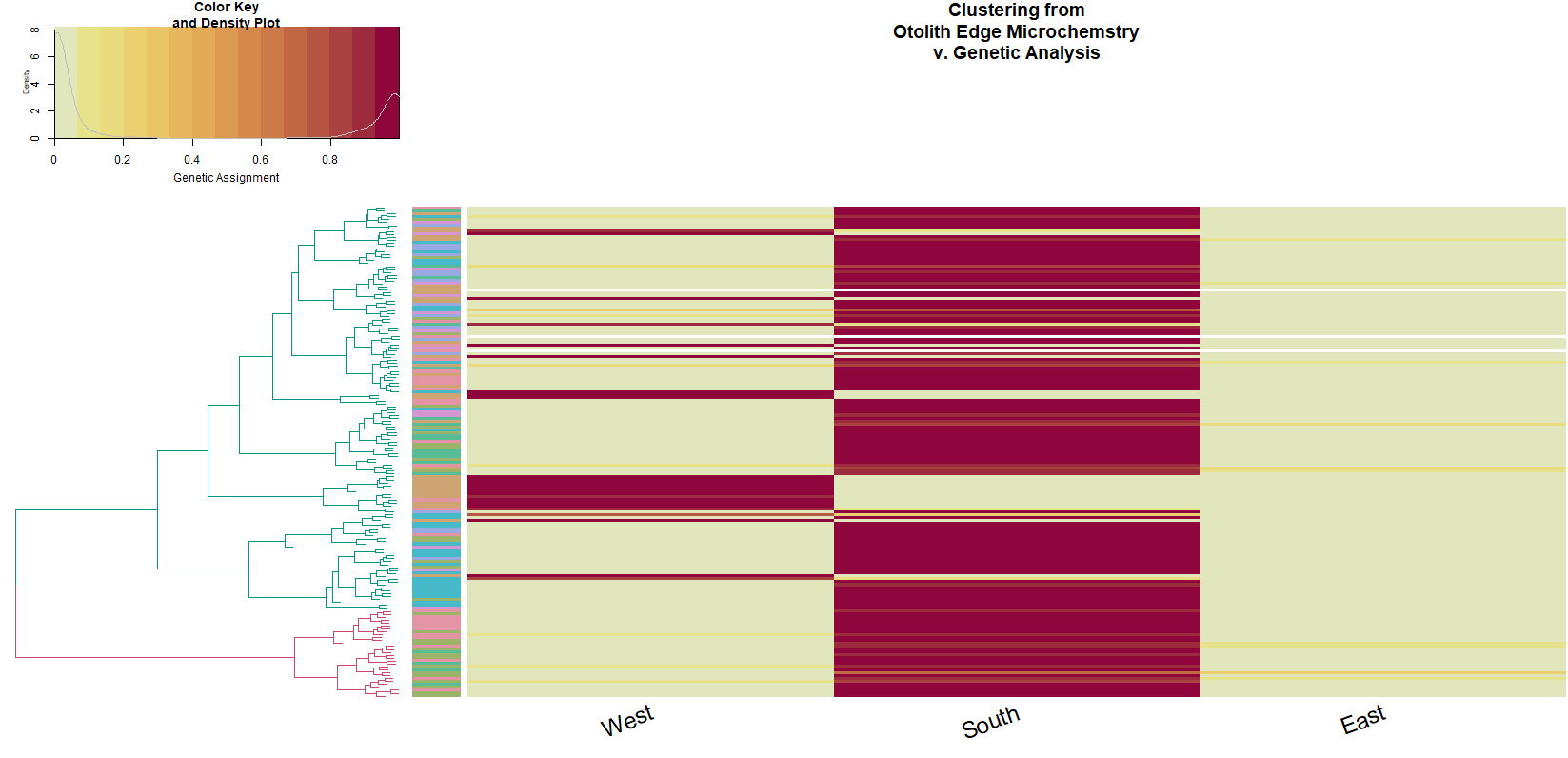


**(a)**

**(b)**

**Figure 6.** Tanglegram comparing cluster analyses of samples using otolith edge and core microchemistry. *Interior lines will be color-coded by sampling site and dendrogram branches will not be colored in final paper.*

**Figure 7.** Dendrogram paired with heat map to compare clustering of individuals between genetic and otolith data sets. Dendrogram displays cluster analysis using otolith microchemistry. *In final paper, color bars associated with dendrogram will be according to sample site.* Heat map displays proportion of genome of each individual assigned to the western or southern coast population. Intensity of color represents proportion of genome assigned.



**Appendix**

Metadata

Scripts

All R scripts used for this study can be found in the “scripts” folder in the [Otolith Analyses directory](https://github.com/mfisher5/PCod-Korea-repo/tree/master/otolith_analyses) of Mary Fisher’s Github page. *For the final paper, I will have a README in this folder to describe the purpose of each script.*

Additionally, downloading the [Otolith Analyses directory](https://github.com/mfisher5/PCod-Korea-repo/tree/master/otolith_analyses) or branching [the PCod-Korea-repo](https://github.com/mfisher5/PCod-Korea-repo) will provide the file structure and data needed to replicate this study’s analyses in R.