Appendix A.

Lab protocols for samples with partially degraded and low concentrations of DNA

A.1. Background

Due to the opportunistic nature of sample collection for marine fishery species, DNA quality and tissue size may vary widely between samples. In our study, after the first round of DNA extractions using the Qiagen DNeasy 96-well Blood & Tissue Kits (Qiagen Inc., Valencia, CA), a number of subsampled tissues had DNA concentrations which were too low for use in our RAD sequencing protocol. In addition, several of our samples contained partially degraded DNA (**Figure A.1.**). We were able to retain these samples for sequencing by implementing minor changes to the standard RAD library preparation protocol (Baird et al., 2008; Etter et al., 2011).

A.2. Low-concentration DNA extraction

In the standard Qiagen DNeasy 96-well Protocol for animal tissue, 200uL of Buffer AE is added to each sample to elute DNA from each well of the 96-well plate (Qiagen, 2006). The plate is then incubated for one minute at room temperature, and centrifuged for four minutes. The protocol notes that to obtain higher concentrations of DNA, it is possible to perform a second elution step using the 200uL Buffer AE flow-through from the previous elution (Qiagen, 2006).

We modified this base protocol for samples with low concentrations of DNA (0.2 – 3.1ng/uL) to include two to three elutions (per recommendation from I. Jimenez, University of

Washington). All samples were first eluted in 120uL of Buffer AE. After samples incubated and were centrifuged, the 120uL of Buffer AE flow-through was used to perform a second elution. Lastly, for samples with DNA concentrations of less than 1 ng/uL, an additional 60uL of Buffer AE was added to each well for a third elution.

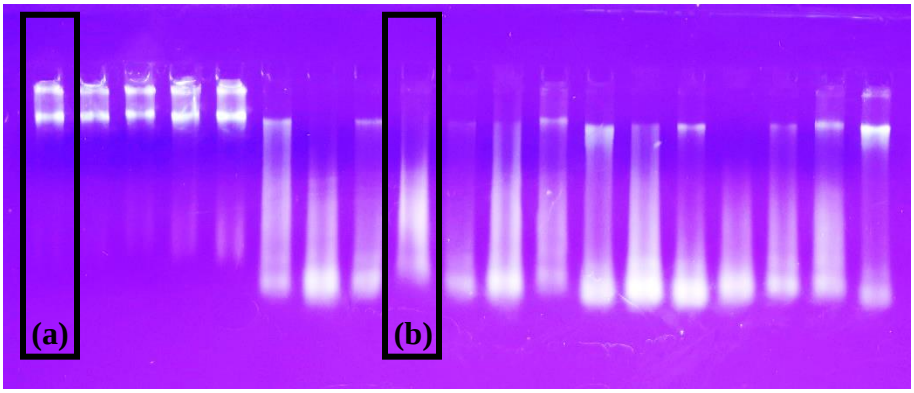
Two elution steps increased DNA concentrations by 121.8% ± 2.5%, on average. Three elution steps increased DNA concentrations by an average of 274.5% ± 368.9%, with only one sample showing no increase in DNA concentration. Although some of this may be a result of human error during the first extraction, or a difference in subsampled tissue size between extractions, the altered protocol did consistently lead to a higher DNA concentration per sample.

A.3. RAD-SEQUENCING WITH PARTIALLY DEGRADED DNA

Alterations to the standard RAD library preparation protocol were restricted to the first four steps, which involve restriction enzyme digestion, P1 adapter ligation, pooling, and shearing (Baird et al., 2008; Etter et al., 2011). First, samples were incubated at 37°C overnight during SbfI-HF digestion and P1 adapter ligation (per recommendation from I. Jimenez, University of Washington), rather than the suggested one-hour incubation period. Samples with partially degraded DNA were then pooled into separate sublibraries, with 12 samples per sublibrary. Having separate sublibraries with only degraded DNA allowed us to shear samples with lower quality DNA for approximately half the number of cycles as those with high quality DNA (per recommendation from I. Jimenez, University of Washington). Of the twelve partially degraded samples that were prepared for sequencing using the following modifications, all but one were retained through the missing data filter and included in the final data set.

A.4. FIGURES

Figure A.1. Picture of extracted DNA on a 1% agarose gel after gel electrophoresis, for the sequenced samples with (a) high quality DNA, and (b) partially degraded DNA.



A.5. SOURCES

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., . . . Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE, 3*(10), e3376. doi:10.1371/journal.pone.0003376

Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011). Local De Novo Assembly of RAD Paired-End Contigs Using Short Sequencing Reads. *PLoS ONE, 6*(4), e18561. doi:10.1371/journal.pone.0018561

Qiagen, Inc. (2006) Protocol: Purification of Total DNA from Animal Tissues (DNeasy 96 Protocol). In *DNeasy Blood & Tissue Handbook*. pp. 35-40. Retrieved from www.qiagen.com/us/resources.

Appendix B.

Genotyping discrepancy between samples prepared for sequencing using 300ng v. 500ng of DNA

B.1. Background

Standard RAD library preparation protocols call for 500ng of DNA per sample (Baird et al., 2008; Etter et al., 2011). Although we successfully increased DNA concentrations using our altered extraction protocol (Appendix A), some samples still did not possess 500ng of DNA. We adjusted our normalization methods and the first two steps of the RAD library preparation protocol to allow for the sequencing of samples with 300ng of DNA, and the subsequent pooling of 300ng-based sublibraries with 500ng-based sublibraries into a single lane for sequencing.

B.2. Methods

We first normalized our 300ng samples in a total volume of 12uL, as opposed to 20uL, in order to maintain consistent DNA concentrations between 300ng and 500ng samples before pooling. Reagent ratios were then altered for SbfI-HF digestion and P1 RAD adapter ligation to account for this change in total volume (**Table B.1.**; per recommendation from M.S. Brieuc, Centre for Ecological and Evolutionary Synthesis).

In order to test the effect of protocol changes on genotyping discrepancy rates, a set of twelve samples were sequenced using both the 300ng and the 500ng protocols, and a set of 20 samples were sequenced with true replicates, for which the same protocol was used during the

preparation of each. Samples then underwent the same quality filtering steps as described in the main text.

We calculated the proportion of loci with different genotypes across only loci that had been genotyped in both the original sample and the replicate. To better understand the impact of genotyping error rate on conclusions drawn at the population level, a principal component analysis (PCA) was completed on all samples from the southern population, with the 300ng replicates included in the data set.

B.3. RESULTS

Genotyping discrepancy was negatively correlated with the number of loci genotyped in both samples, such that samples with a low proportion of missing data tended to share more genotypes with their replicate (**Figure B.1.**). This was true of samples prepared with different protocols, and samples with true replicates.

Of the samples with more than 5000 loci genotyped in both the original and the 300ng / true replicate, the average genotyping discrepancy between samples prepared with the 300ng and 500ng protocols was 6.31%, whereas the average discrepancy between true replicates was 4.37% (**Figure B.2.**). This suggests that genotyping discrepancy due to protocol differences alone was only ~2% on average. According to the PCA, the described error rate did not appear to have a significant effect on the population level; 300ng replicates clustered closely with the original samples prepared using the 500ng protocol, across the first and second principal components (**Figure B.3.**).

B.4. TABLES & FIGURES

Table B.1. Comparison of RAD library preparation protocols used for samples with a total of 500ng and 300ng of DNA. Protocol changes were restricted to the first three steps.

DNA Normalization

		500ng	**300ng**
Concentration (ng/uL)		**Total Volume (uL)**	**Total Volume (uL)**
25ng/uL		20	12

SbfI-HF Digestion

		500ng	**300ng**
Solution	**Stock []**	**Amt for 1 sample (uL)**	**Amt for 1 sample (uL)**
Cutsmart buffer	10x	2.5	1.5
H2O (milliQ)	-	2	1.2
SbfI-HF	20000U/mL	0.5	0.3
RNAse A	100mg/mL	0.1	0.06
TOTAL	-	5	3

P1 RAD Adapter Ligation

		500ng	**300ng**
Solution	**Stock []**	**Amt for 1 sample (uL)**	**Amt for 1 sample (uL)**
H2O	-	1.95	1.17
NEB2	10x	1	0.6
rATP	100mM	0.3	0.18
T4 DNA ligase	2000000 U/ml	0.25	0.15
TOTAL	-	3.5	2.1

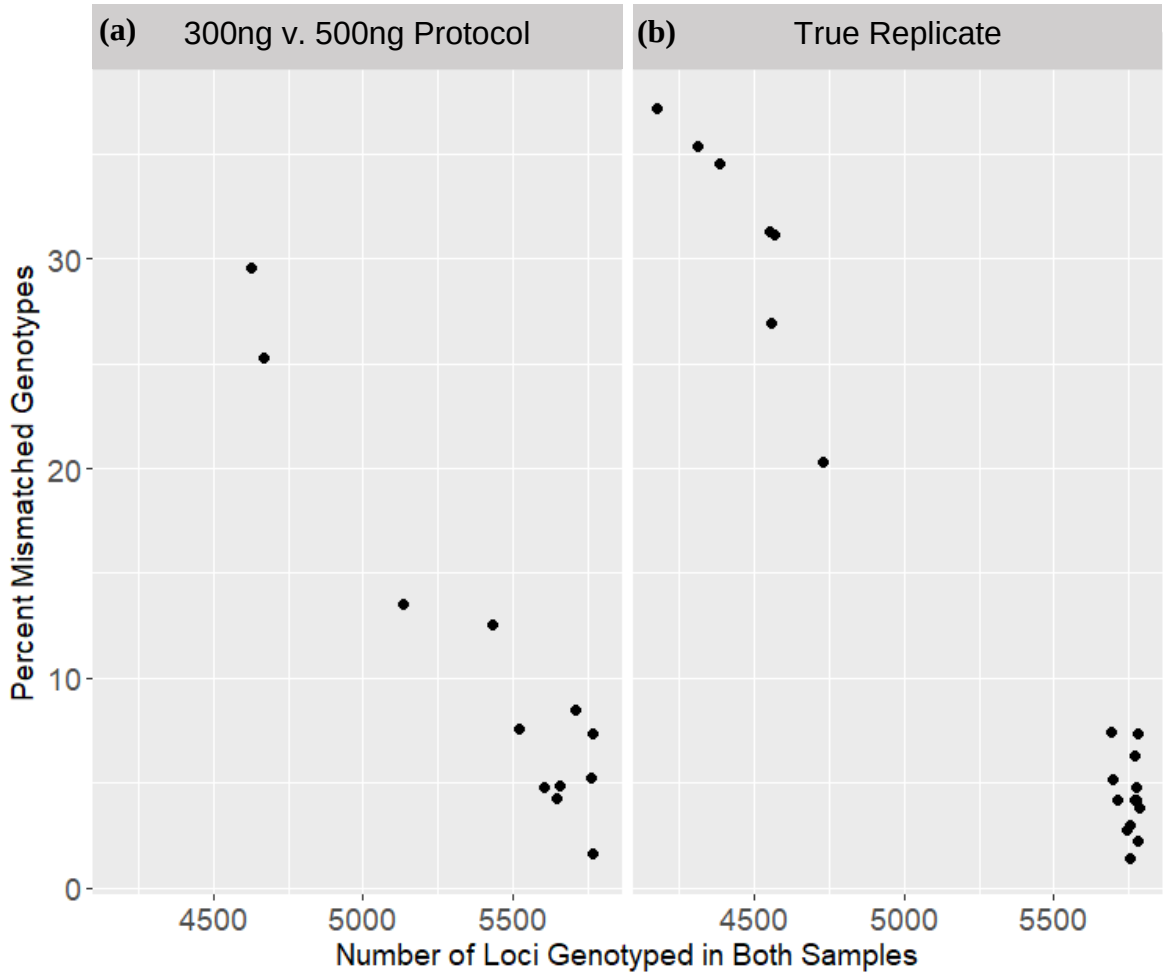


Figure B.1. Genotyping discrepancy rate plotted against the number of loci that were genotyped in the sample and its (a) 300ng replicate, or (b) true replicate.

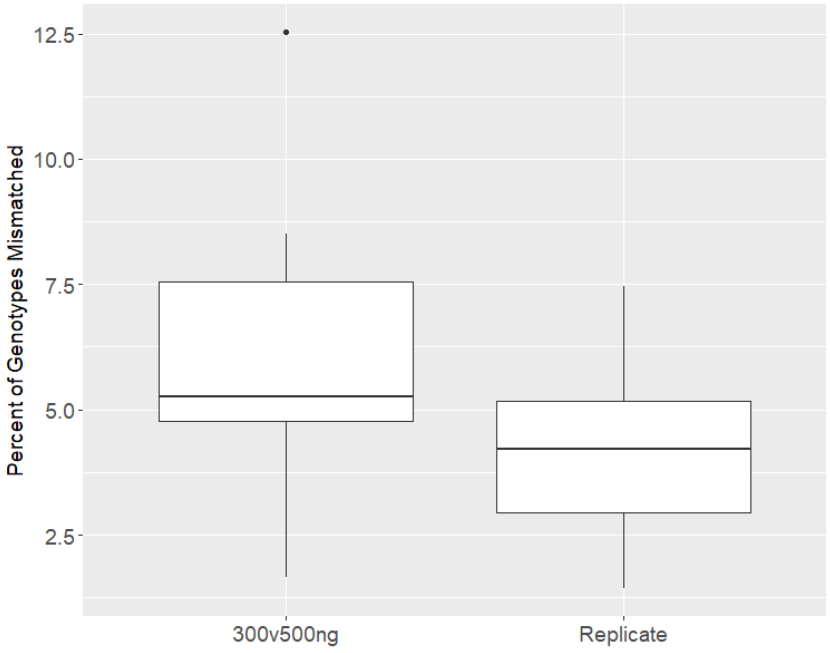


Figure B.2. Boxplot (25/50/75th quartiles, maximum, minimum, and outliers) of genotyping discrepancy for samples with more than than 5,000 loci genotyped.

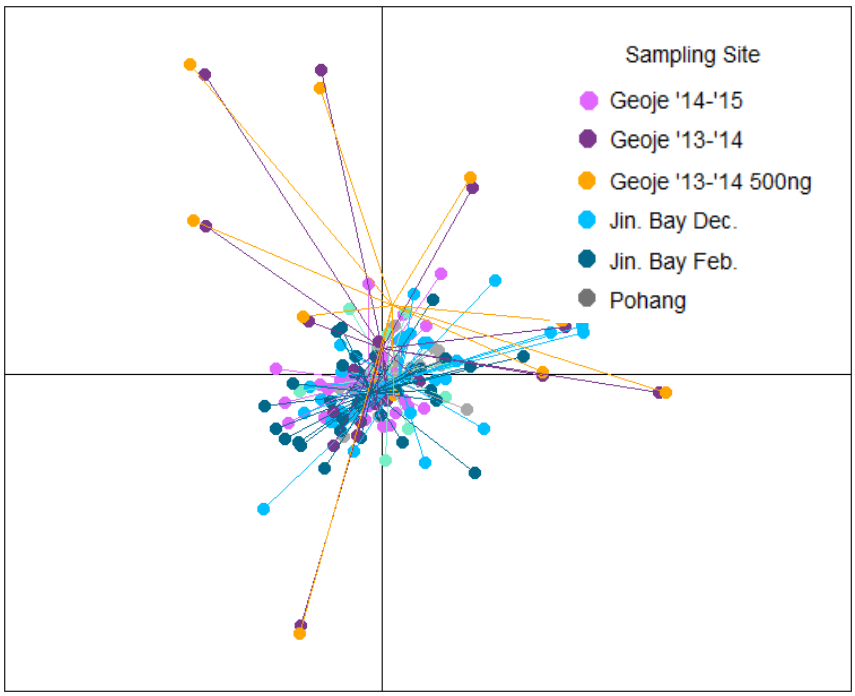


Figure B.3. Principal component analysis on all southern samples, including samples prepared with both 500ng (orange) and 300ng (dark purple) protocols.

B.5. SOURCES

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., . . . Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE, 3*(10), e3376. doi:10.1371/journal.pone.0003376

Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011). Local De Novo Assembly of RAD Paired-End Contigs Using Short Sequencing Reads. *PLoS ONE, 6*(4), e18561. doi:10.1371/journal.pone.0018561

Appendix C.

Simulations to quantify genotyping error using low stack depth

C.1. Background

Although Mastretta-Yanes et al. (2015) report that a stack depth (flag: -m) of three is acceptable when calling genotypes with the Stacks pipeline (Catchen et al., 2011; Catchen et al., 2013), as a rule of thumb we use a more conservative stack depth of ten, and then re-call genotypes per Brieuc et al. (2014) to account for the undercalling of heterozygotes by Stacks (Brieuc et al., 2014).

However, we observed a high proportion of missing data in the samples from the eastern population when a stack depth of ten was used for this study. Although the eastern samples had nearly equivalent coverage as the western samples, with similar distributions of loci aligned to the reference database **(Figure C.1.**), the eastern samples had approximately half the number of raw reads when compared to the samples from the western population (**Figure C.2.**), which led to far lower read depths per locus.

C.2. Methods

To understand the implications of using a read depth of three, we conducted a short simulation study (**Figure C.3.**). We removed a random subset of reads from fifty western population samples after alignment to the reference database, so that each subset sample had 50% fewer reads than the original sample. The size of these new subset files were approximately

equivalent to those of the eastern population. We then completed the remainder of the Stacks v1.44 pipeline (Catchen et al., 2011; Catchen et al., 2013), employing the same parameters as used in the original data set (pstacks m = 5, M = 3, N = 4, n = 3, max_locus_stacks = 3). Genotypes were called by Stacks using a minimum stack depth of three. To match the loci genotyped in the original and the subset samples, we wrote out all SNPs at all loci for the subset samples, and then filtered for SNPs that were present in the original samples. From this genotypes file for the subset samples, we created a list of SNPs genotyped in at least one aggregate, and re-filtered the original sample genotypes files to include only those SNPs. We then re-called genotypes in the subset samples according to Brieuc et al. (2014).

This sequence provided us with four final data sets that contained the same SNPs: two for the original samples, one with genotypes called by Stacks and one with recalled genotypes; and two for the subset samples, one with genotypes called by Stacks and one with recalled genotypes (**Figure C.3.**). Note that we did not go through any additional filtering steps for loci or individuals with high proportions of missing data, for loci with minor allele frequencies, or for loci out of Hardy-Weinberg equilibrium expectations.

We then compared genotypes between the original and the subset samples, for both the genotypes called by Stacks and the genotypes that had been recalled according to Brieuc et al. (2014). The genotyping error rate was calculated by dividing the number of genotypes called differently in the subset sample by the total number of genotypes called in both the original and the subset sample.

Since the primary motive for re-calling genotypes is to avoid under-calling heterozygotes (Brieuc et al., 2014), we also calculated F_{is} per sampling site. Note that because we chose which individuals to subset at random, sample sizes were small within sampling sites; as a result, our

objective was primarily to compare relative change in $F_{is,}$ and we did not attempt to interpret absolute values.

C.3. RESULTS

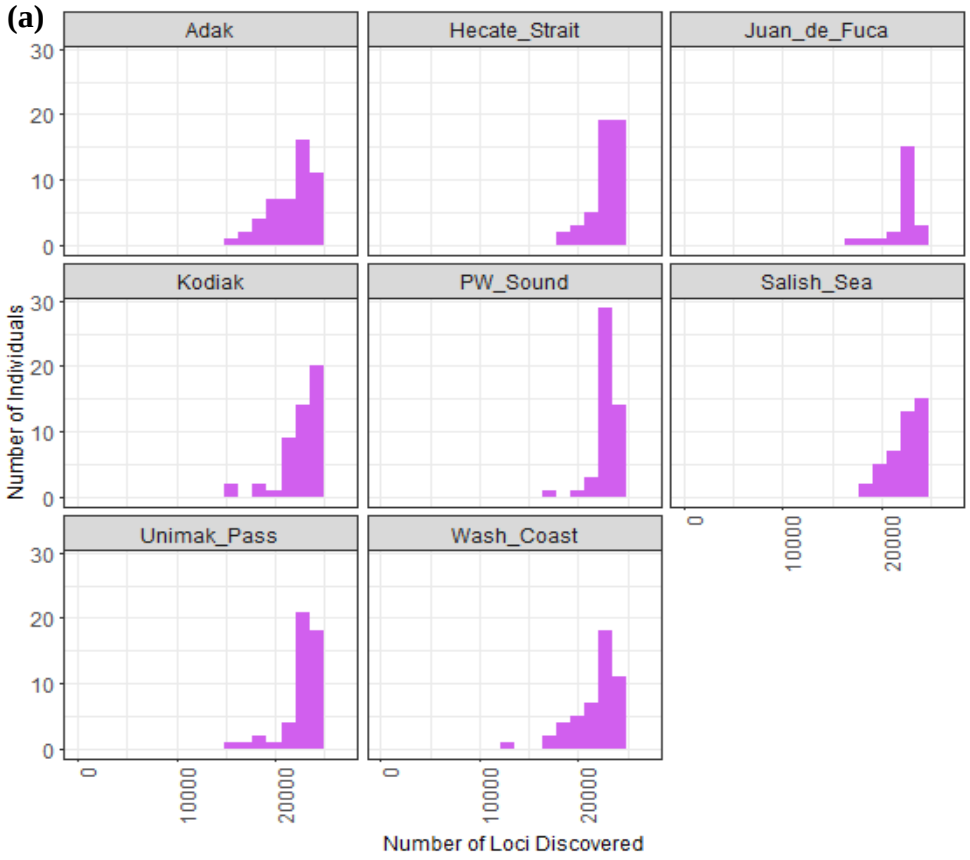
When the genotypes were recalled according to Brieuc et al., (2014), 7.7% of the genotyped loci were called differently in the subset sample than in the original sample (**Figure C.4.a**). Most of this error rate could be attributed to genotypes that were homozygous in the original sample, but were called heterozygous in the subset sample. In contrast, only 0.3% of genotypes were called differently in the subset samples when genotypes were called by Stacks (**Figure C.4.b**).

We did not observe exceptionally high F_{is} values per sampling site when using the Stacks-called genotypes, although the F_{is} values calculated from the Stacks-called genotypes were higher than those calculated from the re-called genotypes, for every sampling site (**Figure C.5.**). F_{is} varied slightly between the original and subset samples.

C.4. CONCLUSIONS

The extremely low error rate when calling Stacks genotypes with a minimum depth of three, without substantially inflating F_{is}, validates our decision to use a low stack depth in our analyses. The results of this simulation verify conclusions drawn by Mastretta-Yanes et al. (2015), in that a stack depth of three can substantially increase the number of loci while minimizing the SNP and allele error rates. However, this assumes that the genotypes called by Stacks are correct. Although we deemed our Stacks-called genotypes as appropriate for our analyses, we did observe an increase in F_{is} from the re-called genotypes.

C.5. FIGURES


(a)
Adak
Hecate_Strait
Juan_de_Fuca
Kodiak
PW_Sound
Salish_Sea
Unimak_Pass
Wash_Coast
Number of Individuals
0
10
20
30
0
10000
20000
Number of Loci Discovered

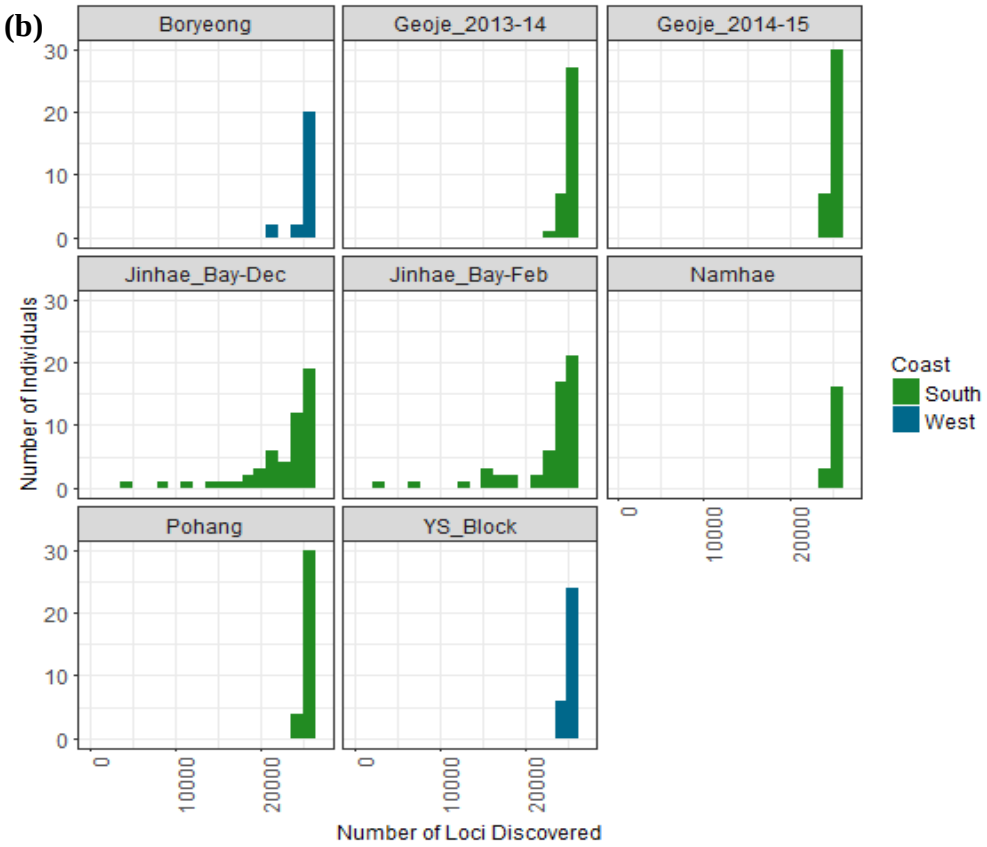
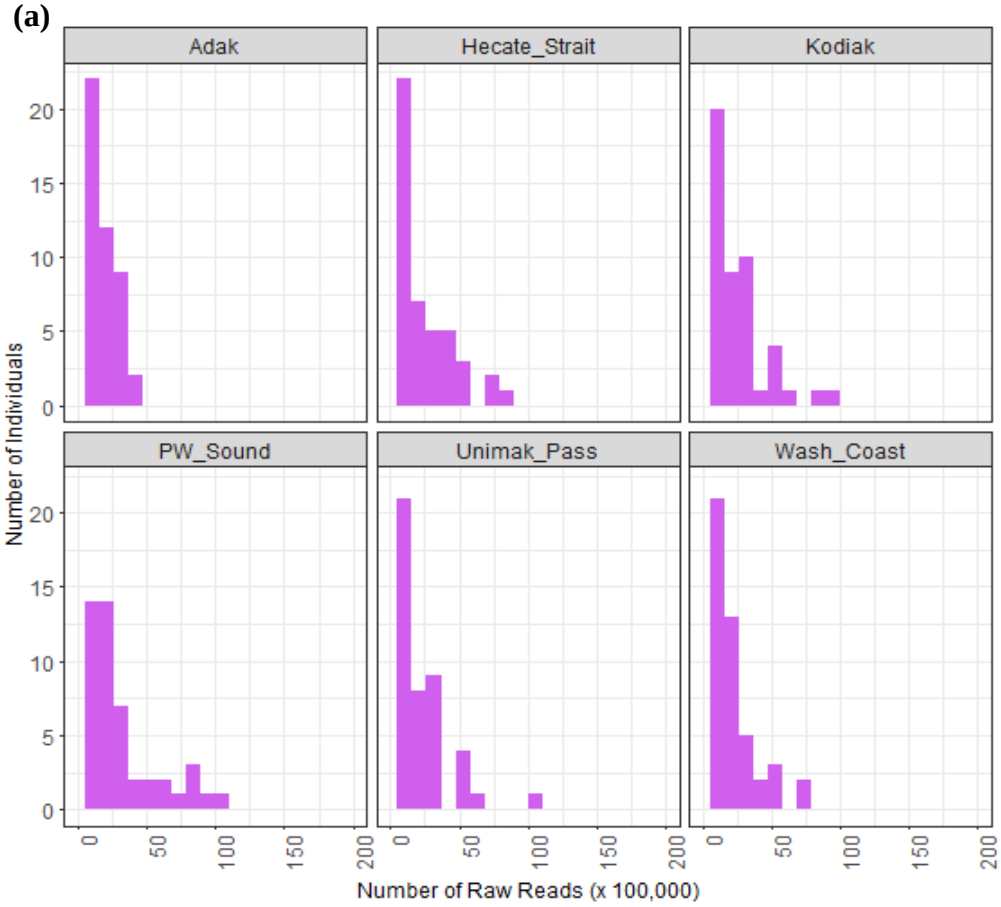


Figure C.1. Distribution of loci discovered per individual for (a) eastern and (b) western Pacific cod populations, after alignment to the reference database. Loci counts were taken for all individuals sampled (not all individuals were retained in final analysis). Within each population, individuals are broken out by sampling site, including temporal replicates. There were 25,669 loci in the reference database.

(a)


Adak
Hecate_Strait
Kodiak
PW_Sound
Unimak_Pass
Wash_Coast
Number of Individuals
0
5
10
15
20
0
50
100
150
200
Number of Raw Reads (x 100,000)

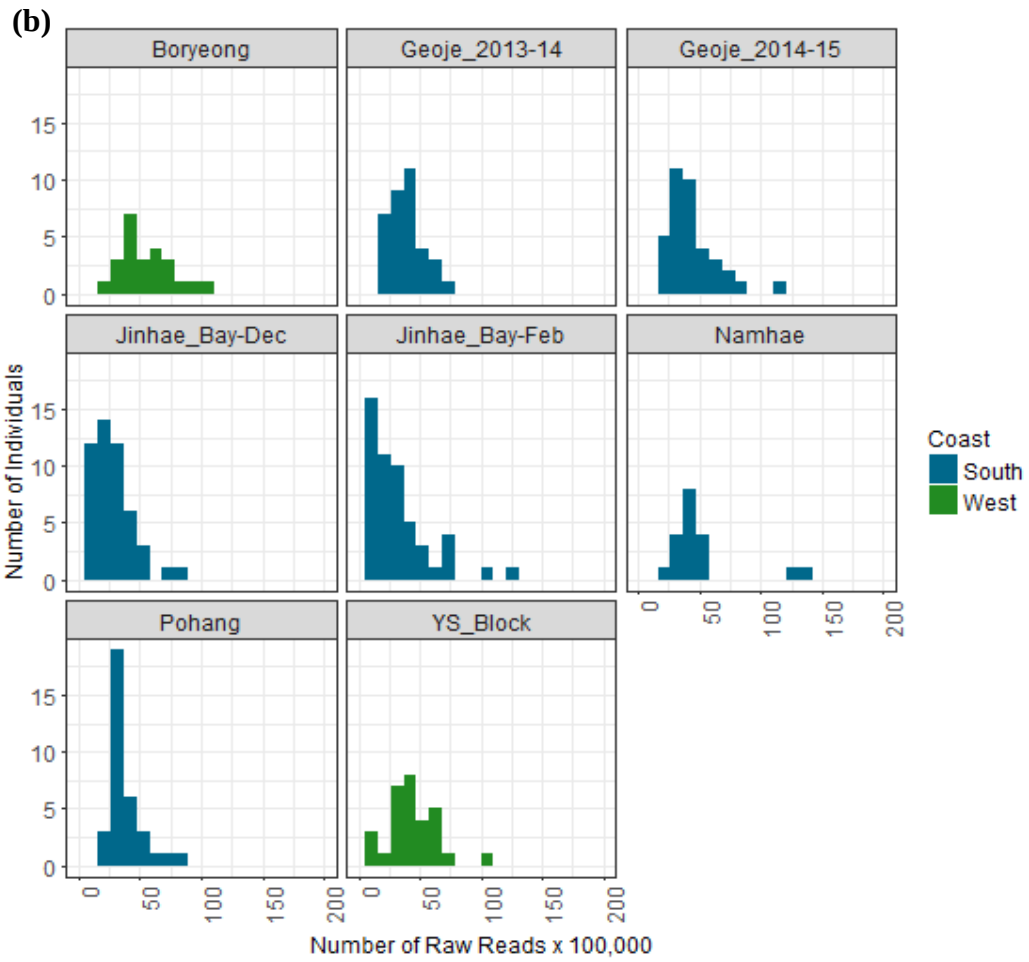


Figure C.2. Distribution of raw reads in demultiplexed FastQ files for **(a)** eastern and **(b)** western Pacific cod populations. Total read depth was calculated for all individuals sampled (not all individuals were retained in final analysis). Within each population, individuals are broken out by sampling site, including temporal replicates.

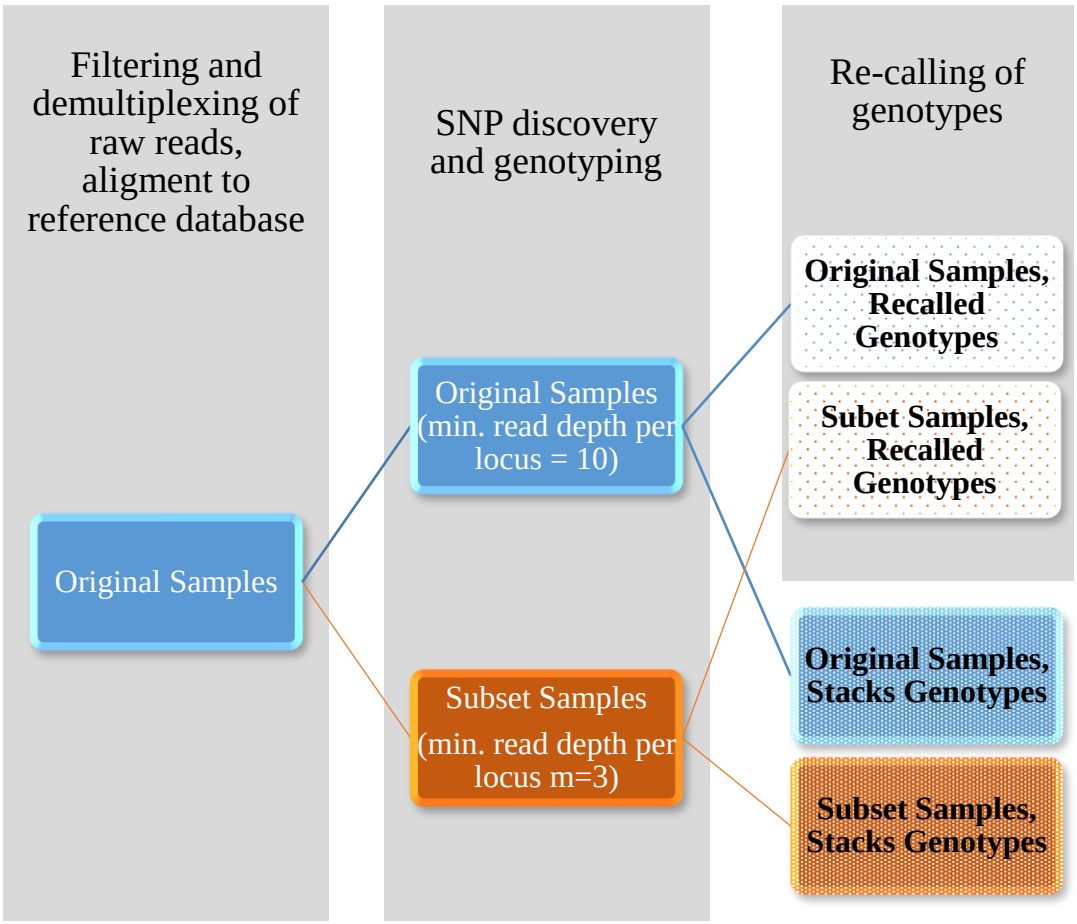


Figure C.3. Diagram of data processing steps used for this simulation study.

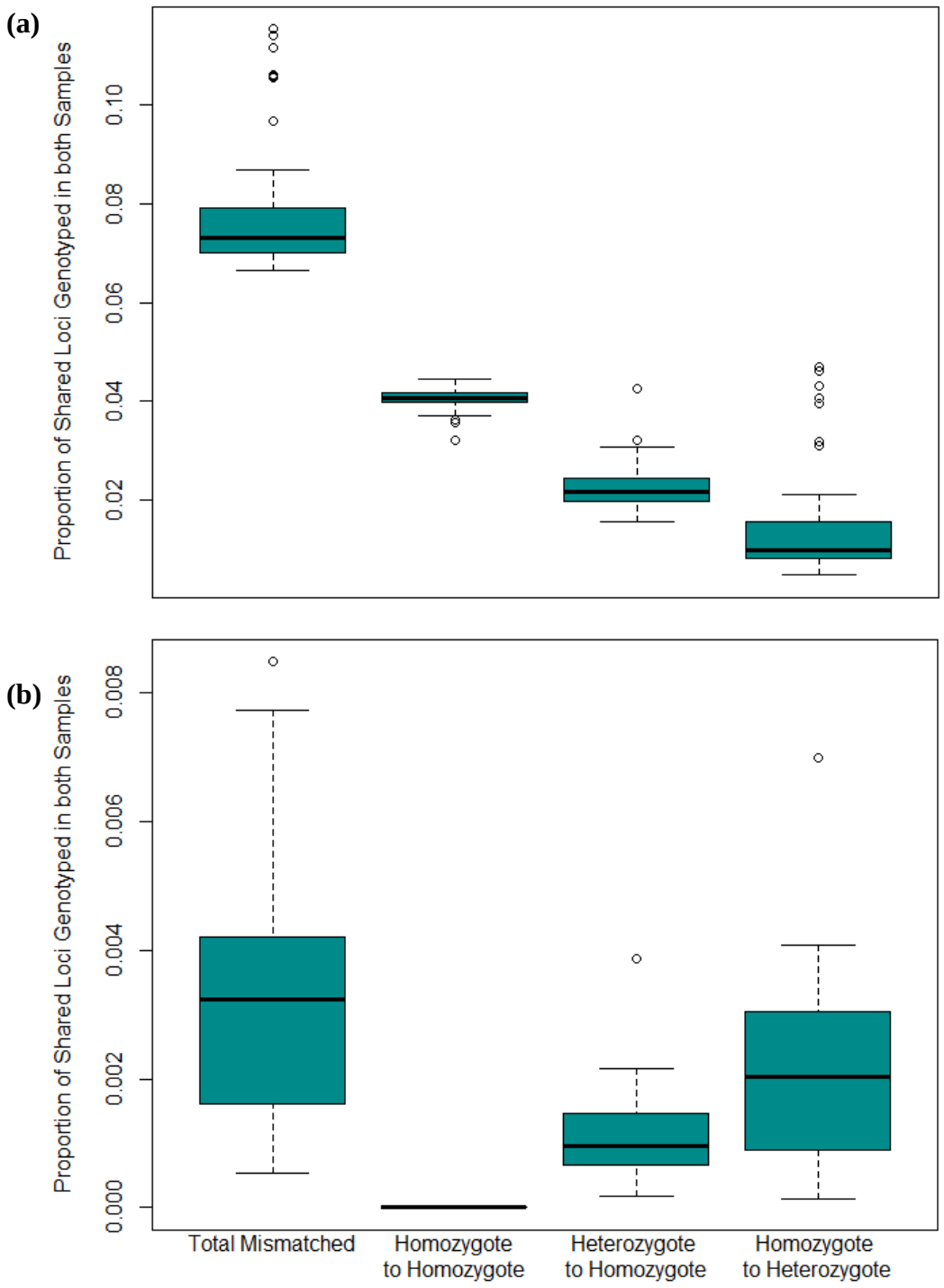


Figure C.4. Genotype discrepancy rates when genotypes were **(a)** called by Stacks, and **(b)** recalled. Boxplots show the overall discrepancy rate (Total Mismatched), and then break down how the genotype changed from the original sample to the subset sample (e.g. the individual was originally called as a heterozygote, but was then called as a homozygote in the subset sample).

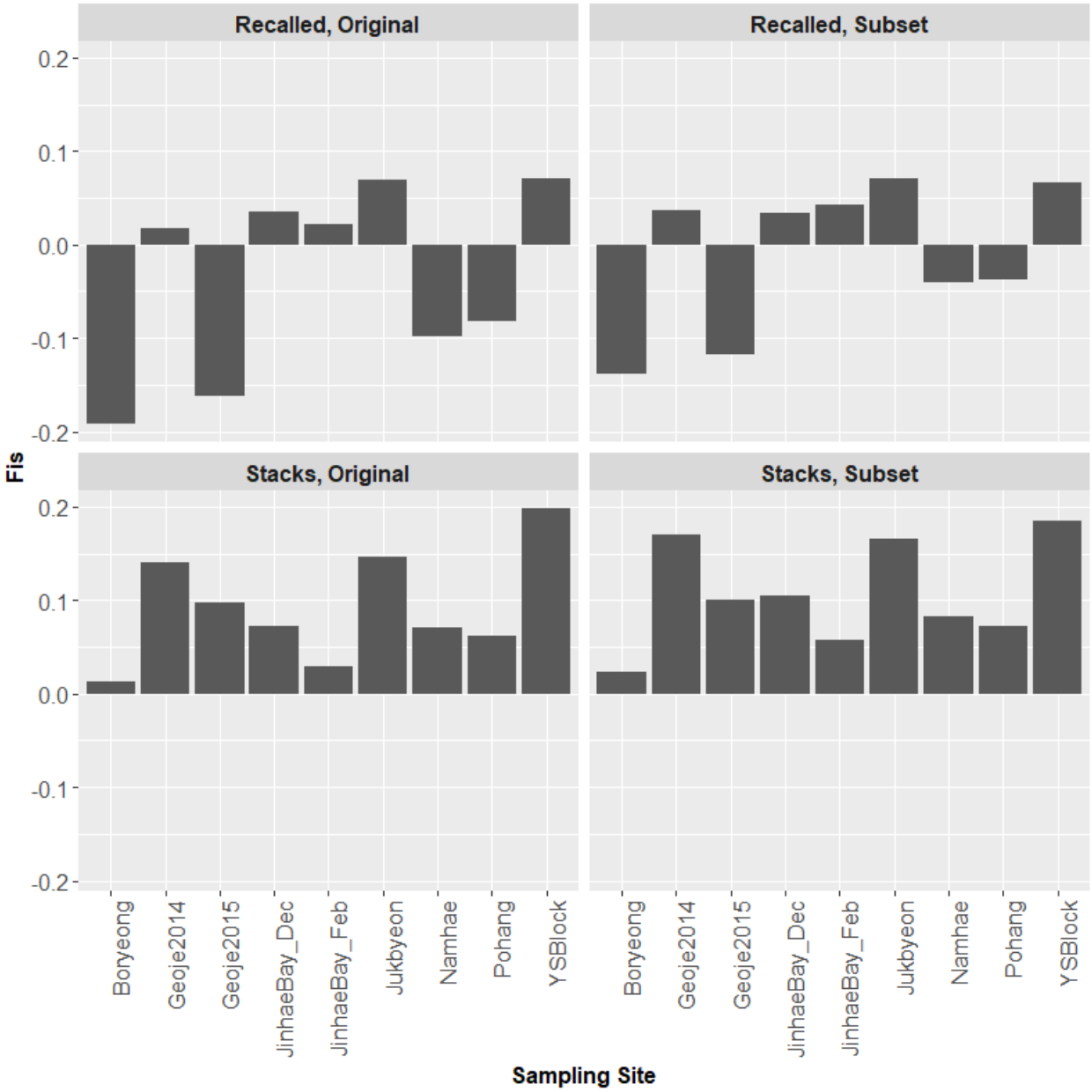


Figure C.5. F_{is} values per sampling site for all four data sets.

C.6. SOURCES

Brieuc, M. S. O., Waters, C. D., Seeb, J. E., & Naish, K. A. (2014). A dense linkage map for Chinook salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal divergence After an ancestral whole genome duplication event. *G3: Genes, Genomes, Genetics, 4*(3), 447-460. doi:10.1534/g3.113.009316

Catchen, J. C., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: building and genotyping loci de novo for short-read sequences. *G3: Genes, Genomes, Genetics, 1*, 171-182.

Catchen, J. C., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124-2140.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources, 15*(1), 28-41. doi:10.1111/1755-0998.12291