



Skate egg nursery areas support genetic diversity of Alaska and Aleutian skates in the Bering Sea

Ingrid Spies^{1,*}, James W. Orr², Duane E. Stevenson², Pamela Goddard², Gerald Hoff², Jared Guthridge³, Myles Hollowed¹, Christopher Rooper^{2,4}

¹Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA

²Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA

³Alaska Sea Life Center, PO Box 1329, 301 Railway Ave, Seward, AK 99664, USA

⁴Present address: Stock Assessment and Research Division, Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada

ABSTRACT: Skate egg case nursery sites are specific locations on the ocean floor where some species of skates deposit egg cases to incubate for up to 5 yr until hatching. We examined genetic diversity within and among skate egg nursery sites of the Alaska skate *Bathyraja parmifera* and the Aleutian skate *B. aleutica* in the eastern Bering Sea to gain a better understanding of how skates utilize these areas. Restriction-site associated DNA (RAD) sequencing libraries were used to obtain single nucleotide polymorphism (SNP) datasets for *B. parmifera* (5285 SNPs) and *B. aleutica* (3309 SNPs). We found evidence for significant genetic differentiation among all *B. parmifera* and *B. aleutica* nursery areas, with 1 exception. *B. parmifera* from the spatially proximate Pribilof and Bering Canyons were genetically similar, suggesting that this may represent a large contiguous nursery area. Genetic differences between embryos at distinct developmental stages within nursery areas were not significant. Adult *B. parmifera* taken from the Bering Sea and Aleutian Islands were genetically distinct from embryo collections, implying that additional genetic types of *B. parmifera* may exist that were not represented by the nursery areas sampled in this study. Our data also showed evidence for low effective population sizes and low N_e/N ratios. Results indicate that nursery areas support genetically distinct components of each species, which underscores the importance of skate egg nursery areas for conservation of genetic diversity.

KEY WORDS: *Bathyraja parmifera* · *Bathyraja aleutica* · Elasmobranch · Ecological divergence · Population structure · Genetics · Skate

1. INTRODUCTION

Egg case nursery sites are localized areas where oviparous skates in the Bering Sea and Aleutian Islands lay egg cases in high densities (>50 000 eggs km⁻²) in contact with the ocean bottom or other permanent structure to incubate for up to 5 yr until hatching without parental care (Hoff 2008). Skate egg and juvenile nursery areas fall within the broader definition of elasmobranch nursery areas, i.e. regions of the ocean where elasmobranchs are

encountered in greater numbers than in other regions, may remain or return over extended periods, and use the habitat repeatedly across years (Heupel et al. 2007, Martins et al. 2018). Nursery areas are critical to the survival of many elasmobranch species, although the mechanisms and extent to which each species depends on nursery areas vary (Beck et al. 2001, Hoff 2016, Heupel et al. 2019). For example, nursery areas may be more important for larger and slower-growing elasmobranchs that produce fewer offspring (Heupel et al. 2019).

*Corresponding author: ingrid.spies@noaa.gov

The reproductive behavior of skates associated with their nursery areas is not fully understood. Whether nursery sites are reproductively isolated or whether there is genetic connectivity among nursery sites is not known. It has also been suggested that some proportion of skates may deposit egg cases outside of nursery areas (Matta 2015). Understanding the genetic importance of nursery sites and the genetic stock structure of skate species is essential for best management practices and for determining how to protect essential habitat for skates (Waples 1998, Spies & Punt 2015).

Like other elasmobranchs, skates grow slowly, mature late, and have low fecundity. The Alaska skate *Bathyraja parmifera* is found throughout the Bering Sea to southeastern Alaska and comprises >90% of the skate biomass, estimated at approximately 600 000 t in 2020 (Ormseth 2020) on the eastern Bering Sea shelf. The Aleutian skate *B. aleutica* is found from northern Japan to the Aleutian Islands and southeast Alaska, and is the most abundant skate on the eastern Bering Sea slope, representing approximately one-third of the skate biomass in that region (Stevenson et al. 2008, Hoff 2016, Ormseth 2020). Fertilization in all skates found in the Bering Sea is internal, and embryos are protected by tough proteinaceous egg cases laid by females in pairs on the seafloor (Matta 2006, Hoff 2009). The age at 50% maturity for *B. parmifera* has been estimated at 9 yr and fecundity is estimated between 21 and 37 eggs yr⁻¹ (Gunderson 1997, Matta 2006). The age-structured model used for the stock assessment of *B. parmifera* extends through age 25, as the maximum age observed is at least 20 yr old (Ormseth 2020). Ageing is difficult for this species, and a higher maximum age is expected based on observed population dynamics data.

In the eastern Bering Sea, 26 nursery sites have been identified for 6 species of skates, and many of these sites are used by multiple skate species (Hoff 2008, Rooper et al. 2019). These nursery sites are primarily found along the interface between the continental shelf and slope, where environmental conditions such as depth, current, speed, temperature, and aspect of the slope relative to current direction are favorable (Rooper et al. 2019). For these species of skates, egg case nurseries are directly tied to survival in their early life history, and represent specific habitat needed for developing embryos within egg cases during a protracted development time of at least 3 or 4 yr for some species (Hoff 2008, Rooper et al. 2019). The high density of egg cases present at nursery sites may be important as a 'predator swamping' survival strategy (Hoff 2009).

Multiple genetically diverse nursery areas may help preserve the biodiversity of skates if climate change or other factors reduce the population size (Anderson et al. 2015). Biological diversity stabilizes ecosystem processes and maintains resilience across populations. According to the 'portfolio' theory, when sharks inhabit a range of habitats at early life stages, it may help to reduce variation in adult population sizes (Yates et al. 2012). Similarly, diversity across hundreds of individual populations of sockeye salmon *Oncorhynchus nerka* reduces variability in salmon runs to Bristol Bay, Alaska (Schindler et al. 2010).

The US North Pacific Fishery Management Council has designated 6 skate nursery sites as Habitat Areas of Particular Concern (HAPC; NOAA 2015), but this listing confers no particular protection to skate egg case nursery areas in the Bering Sea. There is no directed fishery for skates, but adults are caught incidentally in substantial numbers, particularly in the Pacific cod *Gadus macrocephalus* hook and line fishery (Ormseth 2020). Trawl fisheries pose risks to skate nursery areas because of their potential to disrupt habitat on the ocean floor and physically remove or damage skate eggs. Climate change is also a potential risk factor; growth and development times for skate embryos are directly related to temperature (Hoff 2008). Nursery sites occur within a narrow range of temperature with low variability, but projections indicate that temperature and temperature variability will increase for the Bering Sea shelf in the near future (Hermann et al. 2016).

Our objectives in this study were to evaluate genetic diversity in the 2 most common skates in the eastern Bering Sea, *B. parmifera* and *B. aleutica*, within and among egg case nursery sites. We also examined a small collection of adults to examine how they were related to egg case nursery areas. Four of the 5 nursery sites we examined are currently listed as HAPC sites (Fig. 1), and a general description of the substrate and temperature observed is presented to provide context for habitat use. We generated 2 single nucleotide polymorphism (SNP) datasets based on low-coverage sequencing of restriction site-associated DNA sequencing (RADseq) libraries (Andrews et al. 2016). We used these SNP datasets to test the following 2 null hypotheses: (1) skate embryos collected from different nursery sites represent a single panmictic population, and (2) relatedness is the same within and among nursery areas. The first hypothesis was designed to test whether genetically distinct nurseries exist to maintain genetic diversity in the species, while the second hypothesis provided information on site fidelity and population structure.

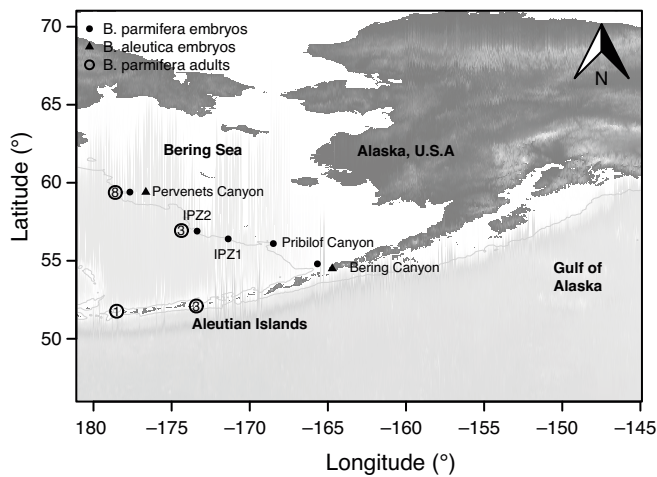


Fig. 1. Skate collection sites. Embryos of Alaska skates *Bathyrja parmifera* (solid circles) and Aleutian skates *B. aleutica* (solid triangles) were collected from egg case nursery sites along the eastern Bering Sea slope: Pervenets Canyon, Inter-Pribilof-Zhemchug 1 (IPZ1), Inter-Pribilof-Zhemchug 2 (IPZ2), Pribilof Canyon, and Bering Canyon. Adult *B. parmifera* were taken from several locations in the eastern Bering Sea and eastern Aleutian Islands, and the number of individuals is shown within open circles. All locations with the exception of IPZ1 have been designated as Habitat Areas of Particular Concern (HAPC). The 500 m isobath is shown as a light grey line

2. MATERIALS AND METHODS

2.1. Sample collection and habitat

Embryos (within skate egg cases) were collected from each nursery site in 2006, 2007, and 2016 using an 83–112 eastern otter trawl net similar to that used for Alaska Fisheries Science Center (AFSC) standard eastern Bering Sea bottom trawl surveys (Lauth & Acuna 2007). Embryos of *Bathyrja parmifera* were collected from the following sites: Bering Canyon,

Pervenets Canyon, Pribilof Canyon, and between Pribilof and Zhemchug Canyons (IPZ1 and IPZ2) (Fig. 1, Table 1). Bering Canyon was sampled twice, in 2007 and 2016, to examine the degree of site fidelity (Table 1). Embryos of *B. aleutica* were collected from Bering and Pervenets Canyons in 2016 (Table 1). Trawls were as short as possible (10 min) to minimize impact to the nursery area. All embryos were preserved in 95% ethanol, and embryo total length was recorded after preservation and converted to developmental stage (2–5) following the methodology of Hoff (2009). An equal number of male and female embryos were collected for each cohort of stages 3–5 (stage 2 embryos were not differentiable by sex). Information on substrate type and temperature were collected at each nursery area. Adults were sampled opportunistically during AFSC bottom trawl surveys conducted from June through August in 2002 and 2006 in the Aleutian Islands and Bering Sea: 8 adults were obtained from Pervenets Canyon, 3 from IPZ2, 3 from the eastern Aleutian Islands, and 1 from the central Aleutian Islands (Fig. 1; Table S1 in the Supplement at www.int-res.com/articles/suppl/m669p121_supp.xlsx). These samples consisted of a piece of tissue preserved in 95% ethanol for genetic analysis, supported by vouchered whole specimens or a photograph for specimens discarded at sea. Embryos from unique combinations of nursery site and year are referred to as collections (Table 1).

2.2. Genomic library construction

RAD sequencing libraries were constructed for 384 embryos of *B. parmifera*, 96 embryos of *B. aleutica* (Table 1), and 23 adult *B. parmifera* (Table S1). Genomic DNA was extracted from skate embryo and adult

Table 1. Collection details for embryos of Alaska skate *Bathyrja parmifera* (AK) and Aleutian skate *B. aleutica* (AL) used in this study, including average measured temperature (Temp) and number of individuals analyzed (N) before [after] sample quality filtering. IPZ: Inter-Pribilof-Zhemchug

Collection location	Latitude (°N)	Longitude (°W)	Species	Collection date	Temp (°C)	N
Bering Canyon	54.8	165.67	AK	3 October 2007	4.5	24 [20]
Bering Canyon	54.8	165.67	AK	29 July 2016	4.5	96 [74]
IPZ1	56.4	171.38	AK	3 July 2016	4.3	96 [86]
IPZ2	56.9	173.37	AK	3 July 2016	3.9	48 [28]
Pervenets Canyon	59.4	177.67	AK	6 August 2006	3.9	24 [18]
Pribilof Canyon	56.1	168.49	AK	8 August 2016	3.8	96 [89]
Bering Canyon	54.5	165.72	AL	3 October 2016	4.5	48 [40]
Pervenets Canyon	59.4	177.65	AL	9 October 2016	3.9	48 [44]

wing tissue using the 96-well format Qiagen DNeasy Blood and Tissue Kit. DNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). We used 500 ng per individual in library construction, and dilute DNA was concentrated using a Multi-Screen-PCR₉₆ filter plate (MilliporeSigma). RAD library construction followed the protocol of Baird et al. (2008) and Drinan et al. (2018), including restriction enzyme digestion with *Sbf*I, ligating adapters, shearing, and PCR. A single library containing 24 embryos of *B. parmifera* from Pervenets Canyon taken in 2006 and 24 embryos from Bering Canyon taken in 2007 (Table 1) was sequenced on an Illumina HiSeq 2500 (100 bp, single read SR100). Ten additional libraries were sequenced to 150 bp single read (SR150) on an Illumina HiSeq 4000 at the University of Oregon Genomics and Cell Characterization Core Facility (GC3F). A maximum of 48 individuals were used per DNA library due to the large size of the skate genome. Known skate genomes are 3–8 times larger than those of teleost fishes. The fugu *Takifugu rubripes* genome is 400 Mb (Kai et al. 2011) and the Atlantic cod *Gadus morhua* genome is 643 Mb (Tørresen et al. 2017), while the size of the little skate *Leucoraja erinacea* genome is 3.2 gigabase pairs (Gbp) (Wang et al. 2012), and the thorny skate *Amblyraja radiata* genome is 2.07 Gbp (https://vgp.github.io/genomeark/Amblyraja_radiata/).

2.3. Data processing and quality checking

Data were downloaded from the GC3F website and processed using Stacks v1.46 (Catchen et al. 2013). Raw data were archived at the Sequence Read Archive (SRA), which is accessible from the National Center for Biotechnology Information (NCBI) at www.ncbi.nlm.nih.gov/bioproject/4306953, BioProject ID PRJNA646172. Data from embryos and adults of *B. parmifera* were processed separately from data from *B. aleutica* to maximize the number of SNPs for each species. Raw data were quality filtered and demultiplexed using the 'process_radtags' subroutine. All sequences were trimmed to 95 bp to accommodate the difference in sequence lengths (100 and 150 bp), and a population map ('popmap') was used to ensure that loci were equally represented in data sequenced on the HiSeq 2500 and the HiSeq 4000.

The 2 most closely related elasmobranch genomes available for alignment at the time of analysis were the whale shark *Rhincodon typus* (Read et al. 2017, Hara et al. 2018) and the white shark *Carcharodon carcharias* (Marra et al. 2019). The white shark ge-

nome was selected for alignment because it contained fewer missing metazoan genes than the whale shark and fewer scaffolds (19 000); fewer scaffolds indicate a better understanding of chromosome structure (Marra et al. 2019). Alignment success of the *B. aleutica* data to the white shark was too low (<0.5%) to acquire an aligned dataset.

A 2-step process was performed to establish a de novo baseline set of SNPs without duplicates, and effectively combine data run on different platforms. A de novo analysis pipeline was selected that increases the number of loci and reduces the SNP calling error rates (Mastretta-Yanes et al. 2015). The 10 individuals with the most sequence data from each collection were used in catalog construction during the first and second steps to reduce the detection of false polymorphisms. First, sequence alignment, SNP identification, and construction of a catalog containing all SNPs were performed using the subroutines 'ustacks,' 'cstacks,' 'sstacks,' and populations with default values except $-m\ 3$ (minimum depth of coverage required to create a stack). Default parameters options included $-M\ 2$, which was selected to avoid large, repetitive sequences and paralogs (Mastretta-Yanes et al. 2015), where $-M$ is the maximum distance in nucleotides allowed between stacks.

The 'fasta' file of all selected loci was converted to a database using Bowtie software v. 2.3.4.1 (Langmead & Salzberg 2012). The Bowtie database was aligned to itself to check for duplicate loci or loci that were inadvertently split ($-v\ 3$, $-k\ 2$), where v is the number of mismatches regardless of quality, and k is the number of valid alignments to report. Identical loci were culled so that all loci in the database were unique using the function 'RemoveDups.R' (<https://github.com/31ingrid/RemoveDups>), which was run in R version 4.0.3 (R Core Team 2020). In the second round of analysis, data from all individuals were aligned to the new database using 'pstacks' with flags ($-m\ 3$, $-bound_high\ 0.05$, $-max_locus_stacks\ 3$), where $bound_high$ is the upper bound for the error parameter epsilon, and max_locus_stacks is the maximum number of stacks allowed per single locus. These were followed by subroutines 'cstacks' ($-g$) and 'sstacks.' Finally, the subroutine 'populations' was run with no restrictions on the number of SNPs per locus ($-m\ 10$, $min_maf\ 0.05$), where min_maf is the minimum minor allele frequency. The $r = 0.5$ parameter was also incorporated, which required that 50% of individuals in a sample collection were genotyped to process a locus for that population.

The SNP with the highest minor allele frequency was then selected from each sequence read. Further

filtering took place in PLINK software v1.90b5.3 format (Purcell et al. 2007) to identify loci with missing call rates, individuals with missing data, individual heterozygosity, and individual observed and expected homozygous genotype counts (*-mind*, *-geno*, *-het*). For datasets of both species, loci genotyped in less than 80% of individuals (*-geno* 0.20) were removed, because this provided sufficient loci genotyped in a high proportion of individuals (Table S2). Individuals with less than 60% genotyping rate were removed (*-mind* 0.4), which also balanced sequencing success while retaining as many samples as possible (Table S2). Individuals with anomalously high heterozygosity were removed if they appeared to be low F_{IS} outliers (<first quantile $-1.5\times$ the interquartile range), as this was indicative of contamination (mixture of more than a single individual). Read depth and heterozygosity were plotted for all samples of *B. parmifera* to check for inconsistencies among sequencing platforms (Fig. A1 in the Appendix), and heterozygosity by collection is shown as a series of boxplots for comparison (Fig. A2).

Prior to examining whether loci conformed to expectations of Hardy-Weinberg equilibrium (HWE) in *B. parmifera*, pairwise F_{ST} and Jost's D were calculated and used to establish a prior grouping of collections. This step was important for the *B. parmifera* data because information on population structure was unknown, and uninformed removal of SNP loci due to HWE could bias marker selection. Prior clustering was not performed for the *B. aleutica* data because only 2 geographically distant nursery areas were sampled, and the expectation of Hardy-Weinberg proportions in each collection was biologically reasonable. The R package 'StAMPP' (Pembleton et al. 2013) was used to perform pairwise comparisons of F_{ST} among all collections, calculated according to Weir & Cockerham (1984) with 1000 bootstraps across loci (Table S3). Pairwise Jost's D was also calculated among collections to provide a second prior estimate of population structure using 'MMOD' in R (Table S3; Winter 2012). Jost's D uses a multiplicative partitioning of diversity based on the effective number of alleles rather than expected heterozygosity (Meirmans & Hedrick 2011) and is considered a useful measure for examining allelic differentiation, particularly when mutation rates are low (Jost 2008). We tested for hierarchical genetic clustering using a method specifically tailored to groups with low levels of differentiation such as marine species (D'Aloia et al. 2020). Clusters were evaluated using the 'pvclust' package with the pairwise matrices of F_{ST} and Jost's D , with Euclidean distance, average cluster method, and support for

each cluster was based on 10 000 bootstrap replicates (Suzuki & Shimodaira 2006). Probabilities were presented as approximately unbiased values (AU) and bootstrap probabilities (BP); AU is considered more accurate than BP (Suzuki & Shimodaira 2006).

Once clusters were established, conformation of genotype proportions to HWE expectations was examined for each cluster of collections using 'hw.test' in the R package 'pegas' (Paradis 2010). A chi-squared test compared expected genotype frequencies calculated from allelic frequencies to actual genotypes. The false discovery rate (FDR) method (Benjamini & Hochberg 1995) was used to correct for multiple comparisons at an alpha level of 0.05. Loci were discarded if they were out of HWE in more than 1 out of 6 clusters identified for *B. parmifera* (following cluster analysis) and if they were out of HWE in both nursery site sample collections of *B. aleutica*.

2.4. Candidate loci for diversifying selection

Candidate loci under natural selection were examined using 2 methods, OutFLANK and BayeScan, with an FDR of 0.05. OutFLANK (Whitlock & Lotterhos 2015; <https://github.com/whitlock/OutFLANK>) estimates the distribution of F_{ST} estimates based on a trimmed set of F_{ST} estimates, and assumes that most loci in the center of the distribution are neutral; loci that fall outside the distribution are considered candidates for selection. The lower and upper 0.05 ends of the distribution were trimmed. BayeScan 2.1 was also used to assess the presence of loci under selection (Foll & Gaggiotti 2008). BayeScan utilizes differences in allele frequencies between populations in a Bayesian framework and provides F_{ST} estimates with a q-value, the FDR analog of the p-value. This used the following collections: Bering 2016, Pribilof 2016, IPZ1, IPZ2, and Adults. Loci were considered as candidates for selection if they were identified as significant using both methods. A second test served as a check for scoring data compatibility in *B. parmifera* between the HiSeq 2500 and HiSeq 4000 platforms; all collections run on the HiSeq2500 were compared with collections run on the HiSeq4000. Loci were removed that were considered under selection among samples analyzed on different platforms.

Candidate loci under selection (Fig. 3) were aligned to the *Amblyraja radiata* annotated genome, which consisted of 49 chromosomes and was published subsequent to analysis. This species is more similar to the genus *Bathyraja* than sharks, as both are classified in the order Rajiformes. Alignment to

the *A. radiata* genome was performed using BLASTn (<https://blast.ncbi.nlm.nih.gov>).

2.5. Inbreeding coefficient, basic statistics

The inbreeding coefficient, F_{IS} , was calculated for each skate collection using the function 'genedivFis' in the R package 'genepop' (Rousset 2008). F_{IS} was estimated for each collection overall and separately for each embryonic developmental stage. Significance of F_{IS} was determined using $\chi^2 = n \times F_{IS}^2$, which follows the chi-squared distribution with a single degree of freedom, and n represents the number of individuals (Nei 1987). Rarefied allelic richness was calculated over all loci using the R package 'hierfstat' and reported as the mean over loci by population (Goudet 2005).

2.6. Tests for genetic differentiation

Pairwise F_{ST} and Jost's D were performed on *B. parmifera* (for a second time based on the final dataset) and *B. aleutica*. Analysis of molecular variance (AMOVA) between nursery collections (Table 1), between stages within collections, between individuals within stages, and within collections was calculated using the R packages 'poppr' and 'ade4' (Kamvar et al. 2014, Dray & Dufour 2007). AMOVA testing was performed on embryo collections of *B. parmifera* performed in 2 ways. First, 5 groups were tested ($N = 5$), retaining Pribilof and Bering Canyons as distinct groups: IPZ1, IPZ2, Pervenets Canyon, Pribilof Canyon, and Bering Canyon (years combined). Second, 4 groups were tested ($N = 4$), by combining Pribilof and Bering Canyon collections: IPZ1, IPZ2, Pervenets Canyon, and Pribilof/Bering Canyons (years combined). All stages and collections of embryos of *B. parmifera* were included, and collections from Bering Canyon were combined. Collections of *B. aleutica* consisted of a single stage, so they were excluded from this analysis. The observed variance component (σ) at all hierarchical levels, and the p-value to determine significance were calculated.

Isolation-by-distance was evaluated among 2016 embryo collections from Bering Canyon, Pribilof Canyon, IPZ1, and IPZ2. Other embryo collections taken from Bering Canyon in 2006 and Pervenets Canyon in 2007 were excluded due to small sample sizes. Distances among nursery areas were calculated as the shortest geographic distance, estimated using Google Earth (earth.google.com/web). A Mantel test with

1000 permutations was applied to the linearized matrix of pairwise F_{ST} and geographical distance matrix.

The R package 'adegenet' was also used to conduct discriminant analysis of principal components (DAPC) to visualize the genetic relationship among collections for *B. parmifera*, with and without the adult collection (Jombart 2008, Jombart & Ahmed 2011). Cross validation ('xvalDAPC') was used to identify a reduced number of principal components that discriminated observations reliably and consistently.

2.7. Kinship

Two methods were used to determine relatedness, that of the software KING 2.1.3 (Manichaikul et al. 2010) and COLONY (Jones & Wang 2010). COLONY v.2.0.6.5 considers sibships among all sampled individuals jointly, while KING considers the relationship for a single pair of individuals in isolation. If 2 or more siblings (sharing the same father, or mother, or both) were present in a collection, COLONY used information (about their inferred common parent or parents) to find probability of full or half siblings (Jones & Wang 2010). The number of SNPs used in COLONY was reduced to 231 to increase computation time by selecting a minor allele frequency of 40% or greater ($-maf$ 0.4) and a genotyping rate of at least 95% ($-geno$ 0.05). COLONY was run with polygamy for males and females, medium precision for full likelihood, and no inbreeding. KING estimated 2 statistics: the proportion of SNPs with zero identity-by-state (IBS0) between a pair of individuals (e.g. AA and GG) and the kinship coefficient. Higher kinship coefficients indicate greater relatedness, and a negative kinship coefficient indicates no relatedness. All statistics were estimated within and among collections of *B. parmifera* and *B. aleutica* using full SNP datasets. Kinship coefficients estimated by KING and sibship probabilities from COLONY among nursery sites were used as a threshold probability for sibship within nursery sites, as the level of genetic differentiation among nursery areas renders the possibility of sibship among nursery areas extremely unlikely, except between Bering and Pribilof Canyons.

2.8. Assignment testing

A mixed stock analysis in RUBIAS was performed to assign adult *B. parmifera* to nursery area of origin (Moran & Anderson 2019). RUBIAS is a Bayesian hierarchical genetic stock assignment approach that

accounts for population structure among baseline populations using a mixture set of unknown origin and a reference set of known origin (Anderson et al. 2008). Posterior density curves and 95% credible intervals for mixture proportions were created from Markov chain Monte Carlo output. The accuracy of the individual assignment analysis was evaluated by testing self-assignment of simulated individuals of known origin. Known simulated proportions for each reporting group were compared with the numbers estimated by RUBIAS to test the accuracy of the individual assignment analysis. A Z-score was computed from each individual's log-likelihood and compared with a simulated normal density to test whether adults originated from nursery areas outside the reference set, and a Kolmogorov-Smirnov test was used to compare the Z-score with the normal density. RUBIAS required unscored loci to be removed from individual samples in mixture sets of unknown origin, which required the removal of 12 loci.

2.9. Effective population size

We used the program LDNe (Waples & Do 2008) to estimate effective population sizes, with a random mating model and a jackknife method for estimating confidence intervals. This method uses the amount of linkage disequilibrium within a population to estimate effective population size and corrects for bias due to sample size. We selected a minor allele frequency of 0.05 to ensure that the results were not influenced by the presence of rare alleles. Collections of *B. parmifera* from Bering and Pribilof Canyons taken in 2016 were combined for this analysis.

3. RESULTS

3.1. Sample collection and habitat

The substrate at Bering Canyon was soft sand, while the substrate at nursery site IPZ1 was boulders and sand. The substrate at the other nursery sites (Pribilof Canyon, IPZ2, and Pervenets Canyon) was sandy with bits of small to medium gravel. Observed temperatures at nursery areas fell within a range of $<1^{\circ}\text{C}$ ($3.8\text{--}4.5^{\circ}\text{C}$), with the warmest temperatures found at the lowest latitude within Bering Canyon (Table 1). Embryos of *Bathyraja parmifera* ranged in size from 23 to 241 mm total length, and all embryo collections contained at least 2 developmental stages, ranging from Stage 2 through Stage 5 (Table 2). Collections of *B. aleutica* taken from Bering and Pervenets Canyons in 2016 were all Stage 5 and 142–256 mm long (Table 2).

3.2. Data quality filtering in *B. parmifera*

A total of 34 048 SNPs were identified following the initial stacks analysis of *B. parmifera*. Two individuals were removed in the stacks pipeline due to insufficient data or missing barcodes. For the other 405 individuals, 33 342 SNPs remained following selection of the single SNP with the highest minimum allele frequency where multiple SNPs were identified within the same sequence read. After filtering for missing genotypes, there were 5289 SNPs (Table S2). During filtering for individuals with $<60\%$ sequencing success, 73 individuals were removed (Table S2). Two additional samples were removed: one was a duplicated sample (iden-

Table 2. Collections of samples by location, species, and year, F_{IS} over all samples for each year and location, the number of alleles (N_A), observed, the developmental stages observed, the number of individuals of each stage (N), F_{IS} by stage in order corresponding to the stages, allelic richness (A_R), and expected and observed heterozygosity (H_e/H_o). The 2016 collections from Bering and Pribilof canyon are also presented jointly (Ber.+Prib.) to provide information on N_A , A_R , and F_{IS} . AK: Alaska skate; AL: Aleutian skate; –: not applicable. Details about the adult samples are shown in Table S1 in the Supplement

Canyon	Sp.	Year	F_{IS}	N_A	Stage(s)	N	F_{IS} (by stage)	A_R	H_e/H_o
Bering	AK	2007	0.0112	8447	3,5	16,4	0.0138, 0.0021	1.6878	0.225/0.222
Bering	AK	2016	0.0354	8582	2,4,5	20, 14, 40	0.0651, 0.0956, 0.0009	1.6900	0.232/0.225
IPZ1	AK	2016	0.1120	8596	2,3,4,5	22,12,37,15	0.0933, 0.0051, 0.1676, 0.0835	1.6896	0.231/0.206
IPZ2	AK	2016	0.2008	8472	2,5	1, 27	–, 0.2118	1.6817	0.226/0.185
Pervenets	AK	2006	–0.0053	8422	3,4	15, 3	–0.0026, –0.0174	1.6886	0.226/0.234
Pribilof	AK	2016	0.0138	8600	2,3,5	22, 45, 22	0.0226, 0.0089, 0.014	1.6903	0.232/0.230
Ber.+Prib.	AK	2016	0.0235	8627	–	–	–	1.6901	0.233/0.228
Adult	AK	–	0.3657	9230	–	–	–	1.6747	0.220/0.147
Bering	AL	2016	0.0874	6615	5	47	–	1.9961	0.268/0.247
Pervenets	AL	2016	0.0566	6613	5	46	–	1.9954	0.266/0.254

tified as an identical twin in KING), and the other displayed anomalously high heterozygosity, with an outlier value for individual F_{IS} based on previously stated criteria, resulting in a final dataset for *B. parmifera* that contained 315 embryos and 15 adults and 95.4 % genotyping success rate (Table 1; Table S1).

The read depth and heterozygosity for individuals run on the HiSeq2500 fell in the middle of the range of values seen in individuals sequenced on the HiSeq4000 (Fig. A1), indicating read depth and heterozygosity were not skewed or deficient for the Pervenets 2006 and Bering 2007 collections. Samples sequenced by the HiSeq2500 had a read depth of 64.9 on average and mean heterozygosity of 0.34, while samples sequenced on the HiSeq4000 had an average read depth of 90.8 and mean heterozygosity of 0.31. OutFLANK identified 3 candidate loci under positive selection in the dataset for *B. parmifera* due to differences in scoring among platforms (Table S4). These loci were removed, leaving 5285 loci in the dataset used to test for clustering.

Estimates of pairwise F_{ST} (prior to filtering for HWE) indicated that embryos of *B. parmifera* from IPZ1 and IPZ2 were significantly differentiated (the latter highly so) from all other nursery areas except Bering 2007 and Pervenets 2006 (Table S3). F_{ST} between IPZ2 vs. Bering, Pribilof, and IPZ1 collections was 0.0027, 0.0027, and 0.0020 respectively, and all were significant at $\alpha = 0.05$ following FDR p-value correction (Table S3). F_{ST} between IPZ1 vs. Bering, Pribilof, and IPZ2 were 0.0003, 0.0002, and 0.002, respectively, and were all significant as well (Table S3). Adult *B. parmifera* were also significantly different from all embryo collections. Collections from Bering Canyon taken in 2007 and 2016 were not significantly different, and neither Bering Canyon sample was significantly different from Pribilof Canyon. Genetic differentiation measured by F_{ST} among embryos of *B. aleutica* from Bering and Pervenets Canyons was 0.001 and significant ($p = 0.017$, data not shown in a table). Values of Jost's D followed the relative magnitude of F_{ST} , with the largest values among the IPZ2 collection and the adult collection vs. other collections (Table S3).

Dendrograms provided >95% support for adult and IPZ2 collections as distinct from other groups, while Bering 2016 and Pribilof 2016 collections clustered together with >95% support (Fig. 2). IPZ1 clustered separately from other nursery areas with 100% support based on F_{ST} , but its relationship was less certain under Jost's D . Placement of the Pervenets 2006 and Bering 2007 collections was incon-

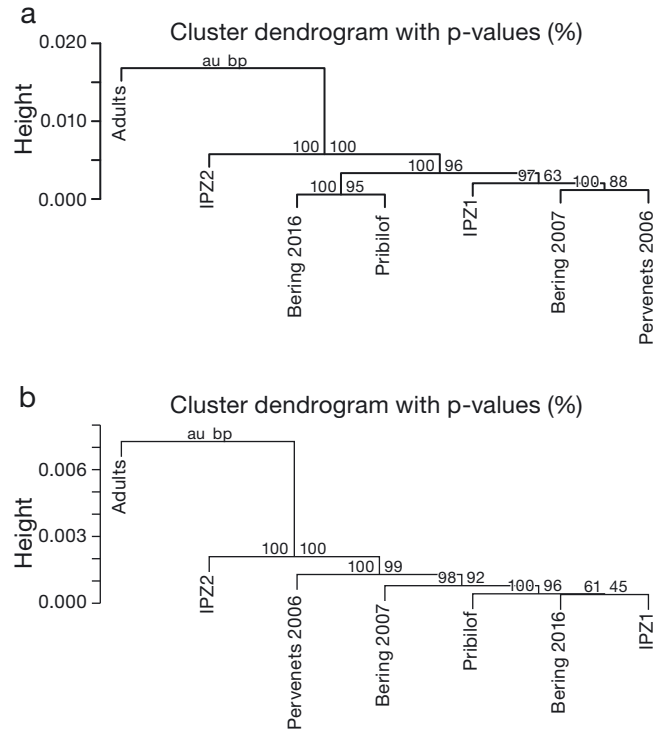


Fig. 2. Hierarchical clusters of *Bathyraja parmifera* based on (a) pairwise F_{ST} and (b) Jost's D for 5285 single nucleotide polymorphisms (SNPs) not filtered for Hardy-Weinberg equilibrium (HWE). Numbers represent probabilities for cluster support based on 10 000 bootstrap replicates, and are presented as approximately unbiased p-values (AU; left side) and bootstrap probability values (BP; right side). AU is considered more accurate than BP, as the latter is a simple statistic computed by bootstrap resampling. IPZ: Inter-Pribilof-Zhemchug

sistent, clustering together with low support (AU = 85 %, BP = 45 %) in the F_{ST} dendrogram but clustering independently with high support (AU > 95 %) in the Jost's D dendrogram. The Pervenets 2006 and Bering 2007 collections contained fewer individuals than other collections, 18 and 20 respectively, while other collections consisted of 28–89 individuals. Therefore, based on their discordant placement and low bootstrap support in the F_{ST} and Jost's D dendrograms, the Pervenets 2006 and Bering 2007 collections were considered somewhat unreliable due to low sample sizes (Fig. 2). Based on this analysis, we defined the following clusters when examining loci for HWE: IPZ1, IPZ2, Adults, Pervenets 2006, Bering 2007, and a sixth cluster consisting of the Bering 2016 and Pribilof 2016 collections combined. After removing 483 SNPs out of HWE in more than 1 of the 6 clusters, 4807 SNPs remained in the *B. parmifera* dataset.

3.3. Data quality filtering in *B. aleutica*

Stacks identified 41 009 SNPs in the dataset for *B. aleutica* at a stack depth of 10 or greater following data filtering and specified constraints, and 25 227 remained after retaining the locus with the highest minimum allele frequency per read. There were 23 903 loci remaining after HWE filtering, and 3309 SNPs remained in the final dataset after filtering for missing genotypes. PLINK identified 10 individuals with >60% missing data, which were pruned from the dataset (Table S2). Three individuals were removed due to identical genotypes identified by KING. This produced a final dataset for *B. aleutica* with 84 individuals, 3309 SNP loci, 89.6% genotyping success, and a mean stack depth of 77 (Table 1).

3.4. Candidate loci for diversifying selection

After removing loci that were out of HWE, 5 loci were identified as candidates for positive selection by OutFLANK (Table S4). BayeScan identified 3 candidate loci under positive selection, 1 of which was also identified by OutFLANK (69495_52, Fig. 3; Tables S4 & S5). The most extreme allele frequencies of the 3 SNPs identified by BayeScan (including the one also identified by OutFLANK) occurred in the Adult and IPZ2 collections (Table S5). OutFLANK identified 26 loci as candidates for positive selection in the dataset for *B. aleutica*. One of these was also the only locus identified by BayeScan (Table S4). Up to 10 matches to the *Amblyraja radiata* genome were presented for each short read, with percent identity >80%, as well as with E-values <10⁻¹⁴, and gene name and function (Table S6). Two short reads (25 904) in *B. parmifera* and sequence 162 908 in *B. aleutica* produced many matches, while 69 495 had no matches and 34 342 produced only one (Table S6).

3.5. Inbreeding coefficient, basic statistics

None of the F_{IS} estimates was considered significant following FDR correction for multiple tests (Benjamini & Hochberg 1995). For *B. parmifera*, the inbreeding coefficient, F_{IS} , was positive in collections from all nursery areas except Pervenets (Table 2). This trend did not change when individual stages were considered separately. The highest F_{IS} was observed in the adult samples (0.3657, Table 2), indicating more admixture than relatedness among samples, although these values were not significantly different from

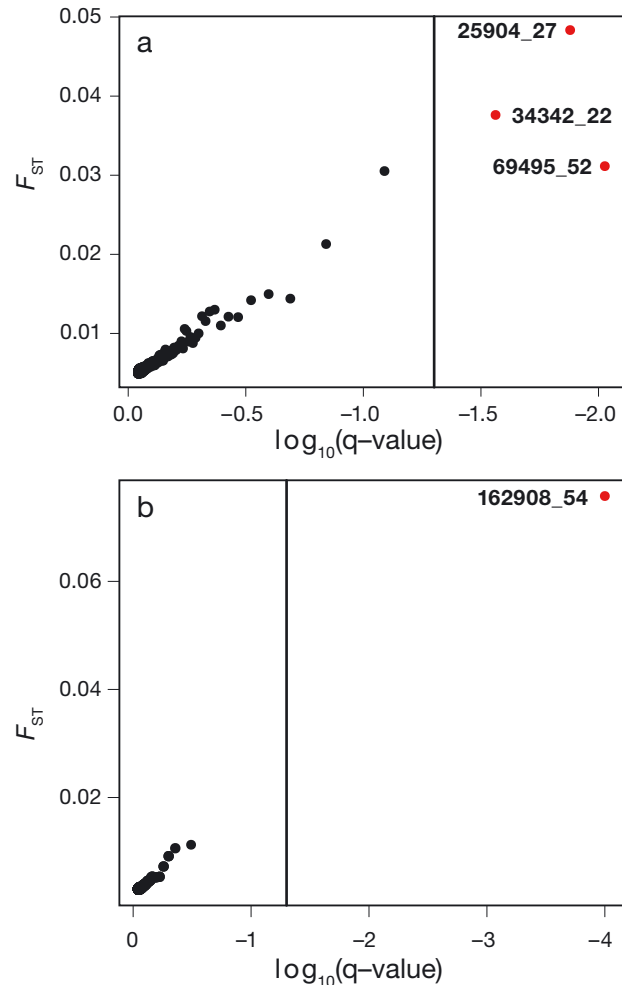


Fig. 3. BayeScan results, showing $\log_{10}(\text{q-value})$, i.e. the false discovery rate (FDR) analog of the p-value vs. F_{ST} and SNPs that were candidates for positive selection in (a) *Bathyraja parmifera* and (b) *B. aleutica*. Vertical line represents the threshold for FDR log scale, FDR = 0.05; red dots represent SNP candidates for positive selection, and numbers are SNP identification

zero. The same trends in F_{IS} were reflected by heterozygosity; adult *B. parmifera* and IPZ2 had notably lower levels of heterozygosity than other collections of *B. parmifera* (Table 2; Fig. A2). Similarly, the lowest rarefied allelic richness, A_R , was observed in the adult collection ($A_R = 1.6747$), followed by the IPZ2 collection ($A_R = 1.6817$) (Table 2). F_{IS} was also positive (although not significantly different from zero) in *B. aleutica* collections, and in the range observed among *B. parmifera* embryo collections, 0.0874 for Bering Canyon and 0.0566 for Pervenets Canyon (Table 2). Allelic richness was slightly higher for these 2 collections than *B. parmifera*, both greater than 1.99, but not comparable with *B. parmifera* because the datasets were analyzed with separate pipelines.

Table 3. Below diagonal: pairwise F_{ST} with significance evaluated with $n = 1000$ bootstraps following false discovery rate (FDR) p-value adjustment (Benjamini & Hochberg 1995) at $\alpha = 0.05$. Above diagonal: Jost's D . **Bolded** values below diagonal indicate statistical significance following FDR correction. Pairwise estimates were performed between Bering (2016), Bering (2006), Pribilof, Inter-Pribilof-Zhemchug 1 and 2 (IPZ1, IPZ2), Adults, and Pervenets (2007), as well as the 2016 Bering and Pribilof collections combined (Bering/Pribilof). Comparisons between Bering 2016 and Pribilof 2016 vs. the combined Bering/Pribilof group were considered redundant and were not calculated (dashes)

	Bering/ Pribilof	Bering 2016	Pribilof 2016	IPZ1	IPZ2	Bering 2007	Adults	Pervenets
Bering/Pribilof		–	–	0.0002	0.0011	0.0002	0.0032	0.0001
Bering 2016	–		–0.0002	0.0000	0.0007	–0.0004	0.0030	–0.0001
Pribilof 2016	–	–0.0001		–0.0000	0.0007	–0.0004	0.0025	–0.0005
IPZ1	0.0004	0.0003	0.0001		0.0012	0.0004	0.0033	0.0002
IPZ2	0.0029	0.0027	0.0025	0.0021		0.0011	0.0039	0.0010
Bering 2007	0.0003	0.0002	0.0005	–0.0003	0.0007		0.0037	0.0001
Adults	0.0093	0.0090	0.0086	0.0074	0.0049	0.0068		0.0035
Pervenets	0.0000	–0.0002	0.0002	–0.0010	0.0002	0.0003	0.0062	

3.6. Tests for genetic differentiation

Pairwise F_{ST} and Jost's D were calculated a second time with the final dataset containing 4802 SNP loci after filtering for HWE (Table 3). Pairwise tests also included 2016 Pribilof and Bering collections combined, which showed significant differences with all other collections except Bering 2007 and Pervenets 2006. Given their small sample sizes, these collections may not be representative of the nursery sites from which they were taken. We also calculated F_{ST} among cohorts with 1000 bootstrap replicates from the same nursery area (Table S7), and none was significant among stages 2, 4, and 5 from Bering Canyon (2016), among stages 2, 3, and 5 from Pribilof Canyon (2016), or among stages 2, 3, 4, and 5 from IPZ1 (2016).

AMOVA results consistently showed significant variation among embryo collections of *B. parmifera* from different nursery areas. The p-values decreased from 0.023 to 0.007 when Bering and Pribilof samples were combined into a single group, and the percentage of total variance attributed to differences among collections increased from 0.08 to 0.12, providing support for grouping samples from Bering and Pribilof Canyons. In neither case was significant variation observed among stages within collections (Table 4). Higher-level comparisons, among individuals within

stages and within collections were not significant in either case ($p = 1.000$).

Adult *B. parmifera* were distinct from embryos in the DAPC, in which 200 principal components were selected using cross-validation (Fig. 4a). A single adult (*Bparm209*) taken from Pervenets Canyon clustered with embryos from the Pervenets, Bering, and Pribilof Canyon nursery areas (Fig. 4a). The remaining adults, including 7 from Pervenets Canyon, 3 from IPZ2, 3 from the eastern Aleutian Islands, and 1 adult from the central Aleutian Islands, were located in a separate half of the plot, indicating relatively

Table 4. Results of AMOVA testing on embryo collections of *Bathyraja parmifera* performed 2 ways: (a) 5 groups ($N = 5$): Inter-Pribilof-Zhemchug 1 and 2 (IPZ1, IPZ2), Pervenets Canyon, Pribilof Canyon, and Bering Canyon (years combined), and (b) 4 groups ($N = 4$): IPZ1, IPZ2, Pervenets Canyon, and Pribilof combined with Bering Canyon (years combined). The results include the observed variance component (σ) at several hierarchical levels, the percent of the total variance (%), and the p-value to determine significance, based on 1000 permutations of the data

Source of variation	df	σ	%	p
(a)				
Among collections	4	0.7277	0.08	0.023
Among stages within collections	11	–0.2966	–0.03	0.667
Among individuals within stages	299	–76.4037	–8.28	1.000
Within collections	315	998.1724	108.24	1.000
Total	629	922.1998	100	
(b)				
Among collections	3	1.0918	0.12	0.007
Among stages within collections	12	–0.3922	–0.04	0.687
Among individuals within stages	299	–76.4037	–8.28	1.000
Within collections	315	998.1724	108.21	1.000
Total	629	922.4682	100	

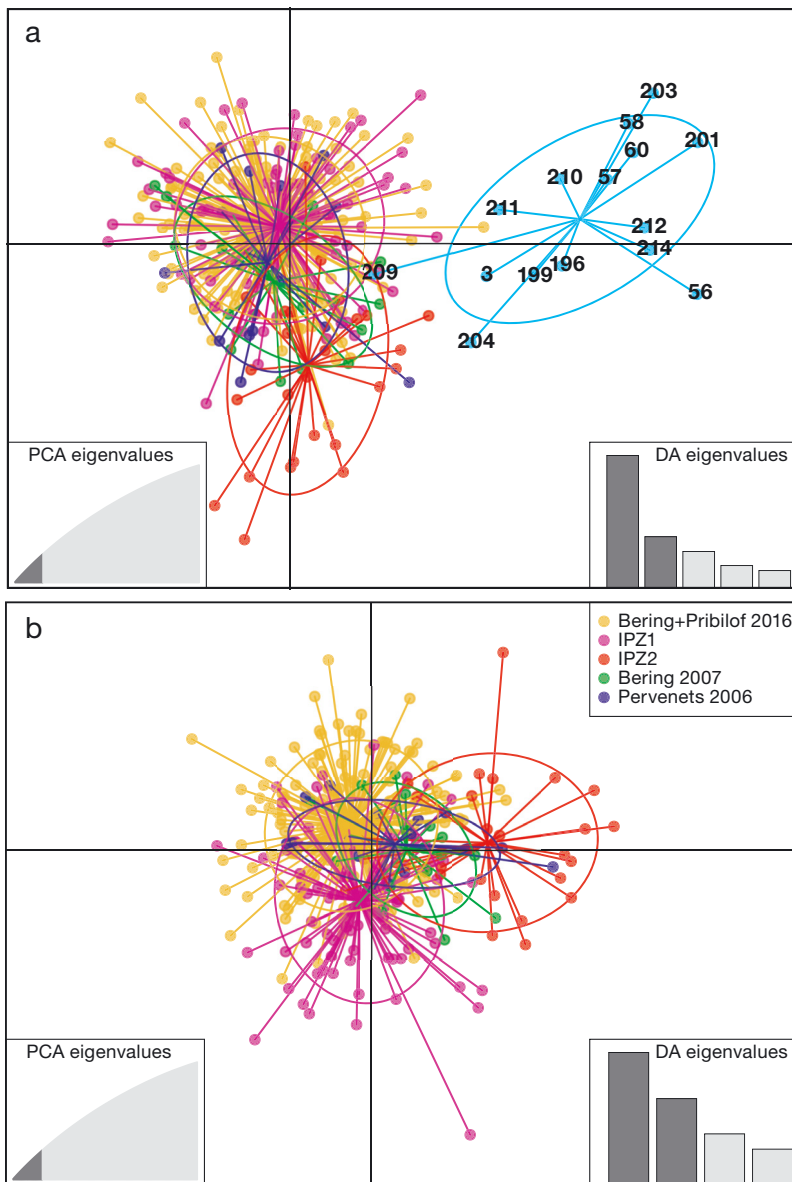


Fig. 4. Discriminant analysis of principal components (DAPC) for all year and nursery site combinations of *Bathyraja parmifera*, (a) including and (b) excluding adult samples. Adult samples are labeled with an individual number (see Table S1 in the Supplement). IPZ: Inter-Pribilof-Zhemchug

large differentiation between adult and embryo samples. Removing adults from the DAPC plot (Fig. 4b; 100 principal components) highlighted differentiation between IPZ1 and IPZ2, as well as a bifurcation in the samples from IPZ2, with some clustering with embryos from other sites and others situated distally. Embryo collections from the other canyons (Bering, Pervenets, and Pribilof) clustered together, including temporally distinct samples from Bering Canyon.

The test for isolation-by-distance was not significant ($p = 0.309$). This was in part due to high levels of

differentiation between the IPZ2 nursery area collection and all other collections, including embryos from the closest nursery area, IPZ1, 134 km away (Table S8).

3.7. Kinship

Kinship coefficients for *B. parmifera* were negatively correlated with IBS0, meaning that higher kinship was associated with higher probability of identity-by-state (Fig. 5). The range and mean of kinship coefficients differed among *B. parmifera* and *B. aleutica* collections (Fig. 5). The highest mean kinship coefficient for *B. parmifera* embryos was among Pervenets Canyon samples (0.004), followed by Bering 2007 (−0.013) and then Bering/Pribilof 2016 combined (−0.032) (Fig. 5a). The lowest mean kinship coefficient for *B. parmifera* was among adult individuals (−0.786), consistent with random sampling of that collection. The IPZ1 and IPZ2 collections had the lowest kinship coefficients of all nursery areas (−0.182 and −0.319). Mean relatedness among all *B. parmifera* nursery areas (−0.286) was lower than mean relatedness within all nursery areas except IPZ2, indicating some level of admixture in that sample.

The mean kinship coefficients for embryos of *B. aleutica* were also negatively correlated with IBS0 and were within the range of values observed for *B. parmifera*: −0.270 for Bering Canyon and −0.065 for Pervenets Canyon (Fig. 5b). Relatedness between *B. aleutica* from Bering and Pervenets Canyons was lower than relatedness within either nursery area (−0.355) (Fig. 5b).

The highest kinship coefficient calculated with KING 2.1.3 among nursery collections was as high as the highest kinship coefficient within nursery areas (0.15). A similar result was found using COLONY; the highest probability of full siblings among nursery areas was 0.703, and the probability of full siblings within nursery areas was 0.703 or lower. Given the large population size of the *B.*

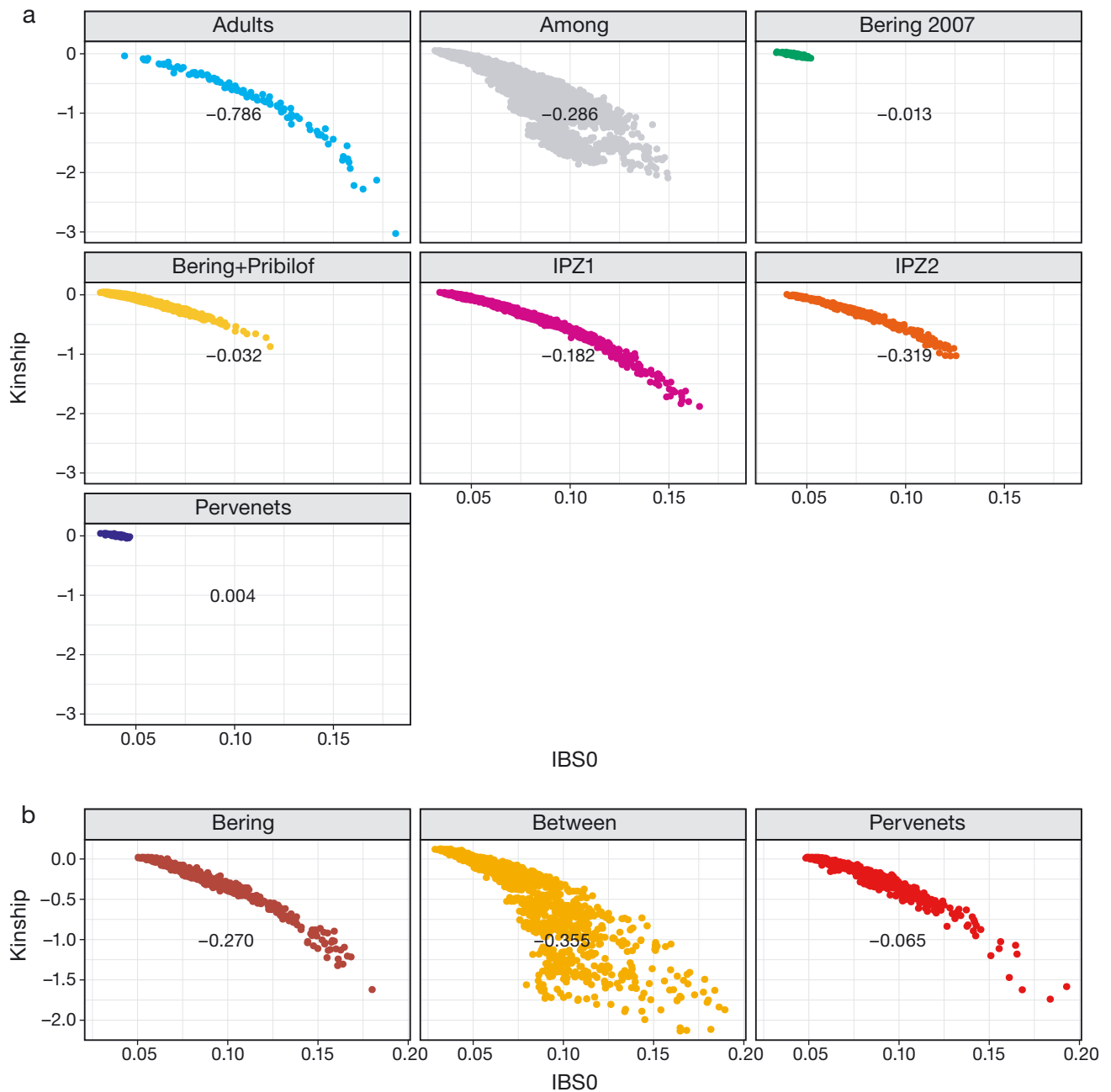


Fig. 5. Relationship between the proportion of SNPs with zero identity-by-state (IBS0) and the kinship coefficient for (a) *Bathyraraja parmifera* within each embryo collection, within adult samples, and among embryo collections (excluding adults), and (b) for *B. aleutica* within Bering Canyon and Pervenets Canyon collections, and between collections. The mean kinship coefficient for each group is shown. The 2016 collections from Bering and Pribilof Canyons were combined. IPZ: Inter-Pribilof-Zhemchug

parmifera population in the Bering Sea, we ruled out the possibility that related individuals were sampled from different nursery areas. Therefore, sibship was considered unlikely within any of the *B. parmifera* nursery areas because no pairwise comparisons were higher within nursery areas than among nursery areas.

3.8. Assignment testing

The Kolmogorov-Smirnov test was used to examine a mixture dataset including all adults (except *Bparm209*) against a reference dataset that included 3 groups: IPZ1, IPZ2, and a third group containing embryos from Bering, Pervenets, and Pribilof Can-

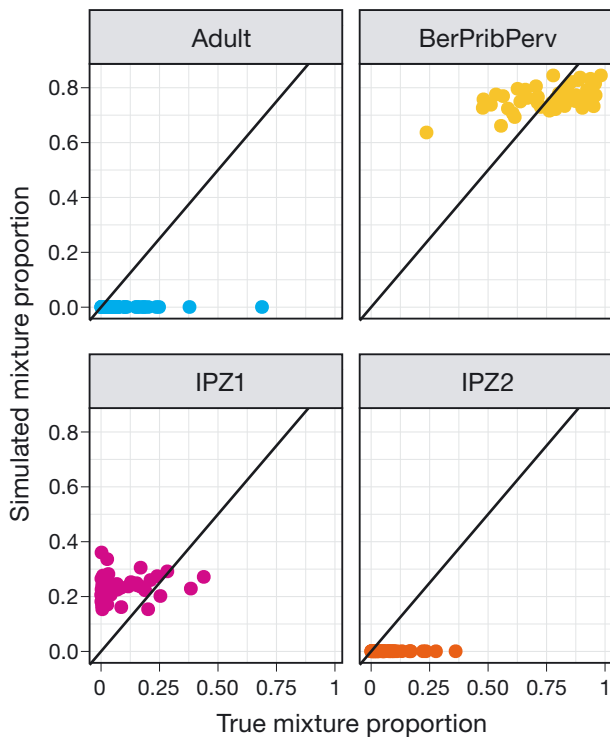


Fig. 6. Comparison of true mixture proportion and simulation mixture contributions for each reference group. Adult: *Bathyrāja parmifera* adults (excluding individual *Bparm209*); BerPribPerv: all embryos from Bering, Pribilof, and Pervenets Canyons; IPZ1 and IPZ2: Inter-Pribilof-Zhemchug 1 and 2

yons. The Kolmogorov-Smirnov test was significant ($p < 0.0001$), providing evidence that the adults in the mixture set likely originated from nursery area(s) outside the reference set.

However, simulated re-assignment success was low for the adult samples and for IPZ2 (Fig. 6). A second test assigned adult *Bparm209* to a reference dataset that included IPZ1, IPZ2, embryos from Bering, Pribilof, and Pervenets combined, and all adults except itself. *Bparm209* was assigned with highest probability to the Bering-Pribilof-Pervenets embryo group, 79.1%, with a credible interval of 0.187–1.000 (Fig. 7, Table 5). Low assignment success of simulated adult and IPZ2 samples implies uncertainty in this result.

3.9. Effective population size

Effective population size was estimable among collections of *B. parmifera* for IPZ1 and the 2016 Bering and Pribilof collections (combined). The estimate of N_e was 1257 (95% CI: 1190–1332) for IPZ1 and 890 (95% CI: 890–905) for Bering/Pribilof. Effective pop-

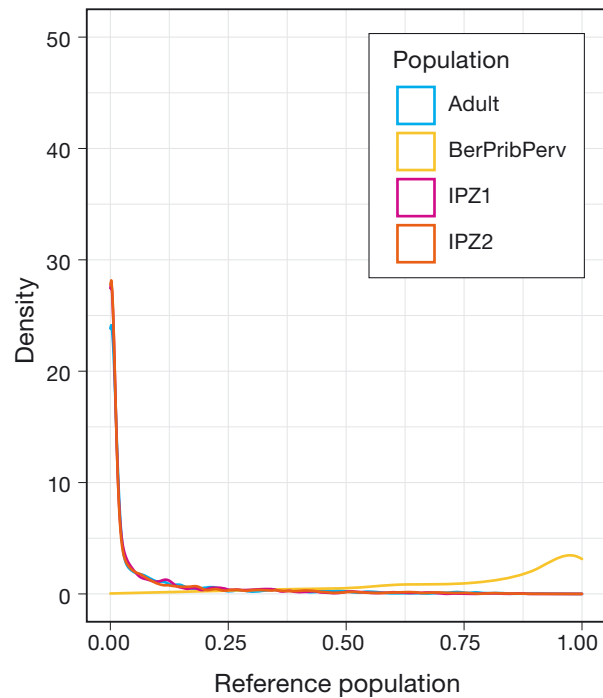


Fig. 7. Posterior density of the probability of assignment of adult *Bathyrāja parmifera* 209 to reference spawning populations. This adult was assigned to the group of embryos from Bering, Pribilof, and Pervenets Canyons and not the adult collection. See Fig. 6 for abbreviations

Table 5. Probability (p) of assignment of adult *Bparm209* to reference populations, with low and high 95% credible intervals (CI). Reference sets included the following: *Bathyrāja parmifera* adults (excluding *Bparm209*); embryos from Bering, Pribilof, and Pervenets Canyons combined; embryos from Inter-Pribilof-Zhemchug 1 (IPZ1); and embryos from Inter-Pribilof-Zhemchug 2 (IPZ2)

Assignment to reference population	p	Low CI	High CI
Adult	0.069	<0.001	0.555
Bering, Pribilof, Pervenets	0.791	0.187	1.000
IPZ1	0.070	<0.001	0.537
IPZ2	0.070	<0.001	0.549

ulation size was estimable for both collections of *B. aleutica*. It was higher for the Pervenets Canyon population of *B. aleutica* ($N_e = 3722$, 95% CI: 2673–6119) than Bering Canyon ($N_e = 236$, 95% CI: 230–243).

4. DISCUSSION

Our understanding of the life history of skates of the genus *Bathyrāja* in the Bering Sea is hindered by a lack of information on the genetic population struc-

ture of skate egg nursery areas. Egg case nursery areas are essential habitat for eggs to develop for up to 5 yr before hatching, but whether females deposit eggs randomly or in specific nursery areas is unknown. These questions are important for understanding the conservation value of skate egg case nursery areas and for determining appropriate management units.

This represents the first study to examine the genetic connectivity among skate egg nursery areas in the Bering Sea. We examined thousands of SNPs to understand whether genetic diversity differed significantly among the nursery areas and to compare relatedness within and among nursery areas. Our most important finding was that skate egg nursery sites in the eastern Bering Sea do not represent a single panmictic population. Embryos of *B. parmifera* from IPZ1 and particularly IPZ2 were genetically distinct from other collections, and differences were not associated with geographic distance or isolation-by-distance (Table S8). In contrast, embryos of *B. parmifera* from the Bering and Pribilof Canyon nursery sites sampled 229 km apart appeared to represent a large interconnected region for egg case deposition by *B. parmifera*. We also observed significant genetic differences among embryos of *B. aleutica* from Pervenets and Bering Canyons (840 km apart). Embryos of different developmental stages within nursery sites were genetically similar, in contrast to significant differences in F_{ST} observed among several egg case nursery sites (Table 3, Fig. 4; Table S7). The F_{ST} values among nursery sites were relatively low in magnitude (0.000–0.003) at distances from 130–530 km, but low F_{ST} values may still be statistically significant and biologically meaningful (e.g. Waples 1998, Knutsen et al. 2011). Genetic differentiation among adults of other species of skates and rays was higher than observed here. Chevolut et al. (2006) observed genetic differentiation among adult thornback rays *Raja clavata* (global estimate of $F_{ST} = 0.042$) in the eastern North Atlantic between the Mediterranean, the Azores, and continental shelf of Europe at distances of 500–1300 km. O'Connell et al. (2019) found evidence for fine-scale population structure ($F_{ST} = 0.00–0.12$) in winter skates *Leucoraja ocellata* but not in little skates *L. erinacea* between Georges Bank and the Mid-Atlantic Ocean at distances on the order of 300 km.

Another key finding was that relatedness within the Bering, Pribilof, and Pervenets Canyon *B. parmifera* collections and within the *B. aleutica* collections was higher than among embryo collections, indicating that in general, skate embryos taken from the

same nursery area were more related than skate embryos taken from different nursery areas (Fig. 5). Higher relatedness within nursery areas indicates some level of site fidelity, at least by females, which could be further tested through analysis of mitochondrial DNA. While relatedness was typically higher within than among nursery collections, an exception was found at IPZ2, which had the lowest mean relatedness. The IPZ2 nursery area also displayed the highest F_{IS} and largest spread of heterozygosity of all embryo nursery collections, and a diffuse pattern in the DAPC plot (Fig. 4), suggesting mixed usage by unrelated groups of *B. parmifera* (Table 2; Fig. A2).

While we predicted that adult samples would be closely related to the sampled embryos, we had the unexpected result that only 1 of the 15 adults of *B. parmifera* taken from the Bering Sea and Aleutian Islands was genetically similar to embryos from sampled egg case nursery sites (Fig. 4). This adult male, which was caught in a single haul with 4 other males to which it was not related, had high assignment probability to the Bering and Pribilof Canyon nursery areas (Figs. 6 & 7; Table S1). It should be noted that although the probability of reassignment was high (80%) for embryos from Bering and Pribilof Canyons, it was much lower for adult samples and embryos from IPZ1 and IPZ2 (Fig. 7), indicating that assignment test results may be unreliable and larger sample sizes may be needed. The genetic distinctiveness of the adults (F_{ST} : 0.005–0.009) in our collection indicated that the nursery areas we sampled are not a complete representation of the genetic diversity in this species. Adult samples were less related to each other than were embryos within collections from different nursery areas (Fig. 5), which is consistent with random sampling over a large area. Adult samples were also associated with positive F_{IS} (Table 2), which can be the result of low heterozygosity, potentially resulting from the Wahlund effect or subpopulation structure. Whether the genetic distinctiveness of the adult skates examined could be due to long-distance migration from natal habitat is not known. Limited dispersal appears to be the rule in *L. erinacea*, *L. ocellata*, and big skates *Beringraja binoculata*, with a few individuals performing long-distance migrations that may be related with mating or egg deposition (King & McFarlane 2010, Farrugia et al. 2016, O'Connell et al. 2019). A tagging study of 1238 *R. binoculata* recovered 75% of individuals within 21 km of their tagging location, and only a small proportion of skates (1.5%) performed long-range migrations up to 2350 km (King & McFarlane 2010). These long-range migrations were primarily per-

formed by females that were maturing or had recently matured (King & McFarlane 2010).

Finally, results indicated low effective population sizes and ratios of effective population size to census size (N_e/N): for *B. parmifera*, $N_e = 1257$ at IPZ1 and 890 at Bering/Pribilof nursery areas, and for *B. aleutica*, $N_e = 3722$ in Pervenets Canyon and 236 in Bering Canyon. These estimates of N_e strongly suggest low N_e/N ratios in *B. parmifera* given the estimate of over 200 million individuals of this species in the Bering Sea, although precise estimates of N_e/N ratios are not directly calculable (Ormseth 2020). The N_e/N ratio is typically small in marine fish populations (10^{-3} or lower), whereas sharks exhibit high N_e/N ratios, approaching 1 (Portnoy et al. 2009, Dudgeon & Ovenden 2015). Our results are supported by estimates of low N_e/N ratio in thornback rays, between 9×10^{-5} and 6×10^{-4} (Chevolot et al. 2008). It is not understood whether the causes of the low N_e/N ratios in thornback rays and species of *Bathyraja* examined here result from skewed sex ratios, fluctuations in population size, or variance in reproductive success (Chevolot et al. 2008). Complex reproductive patterns may be a likely explanation, resulting in unequal sex ratios or high variance in reproductive success; polyandry has been observed in manta rays *Manta birostris* and southern stingrays *Dasyatis americana* (Yano et al. 1999, Chapman et al. 2003, Chevolot et al. 2008). The presence of multiple male participants appears to be a conserved behavior during courtship of batoids (McCallister et al. 2020). High juvenile mortality due to protracted egg incubation time could also provide an explanation for low N_e/N ratios. The effective population size (N_e) is an important parameter for conservation because it represents the size of an idealized population with the same genetic diversity as the population of interest. Higher N_e/N ratios are characteristic of more vulnerable populations, as removal of fewer individuals reduces the overall effective population size. Effective population sizes greater than 500 will prevent loss of genetic variation, and N_e values lower than 1000 may result in accumulation of deleterious alleles (Franklin 1980, Lande 1994, Dudgeon & Ovenden 2015). Although *B. parmifera* appear to have a low N_e/N ratio, the effective population sizes by nursery area for *B. parmifera* and *B. aleutica* are also on the low end of the threshold for conservation. Similarly, low effective population sizes were observed in thornback rays in the Irish Sea ($N_e = 283$) (Chevolot et al. 2008). Samples consisting of multiple cohorts are not expected to have a large effect on the effective population size because they are genetically similar; however, mixed groups at IPZ2 may down-

wardly bias estimates of effective population size (Waples et al. 2014). It should be noted that small sample sizes are a potential reason for failure to estimate effective population size, and may provide a negative bias in effective population size; therefore, future efforts should use larger sample sizes (Marandel et al. 2019).

Future work is needed to resolve new and unanswered questions resulting from this study. Although we identified candidate markers under selection for both species, further work is needed to understand whether there is a genetic component to local adaptation in these species. Also, sequence read depth was over the 40× coverage threshold considered sufficient to minimize errors in SNP calling (Catchen et al. 2011), but the lack of resolution in the relationship between the Bering 2007 and Pervenets 2006 collections was apparent in clustering dendrograms (Fig. 2) and could be a result of low sample sizes (20 and 18, respectively). While this study answered several questions about the genetic population structure of *B. parmifera* and *B. aleutica*, it raised other questions. The presence of comingling genetically distinct adult skates remains a mystery, as does their nursery area of origin (Fig. 4a). Tagging studies of these species would greatly improve our understanding of skate behavior, migration, and reproduction. This study also highlighted the need to understand causes for high variability in reproductive success in these species leading to low effective population size estimates and remarkably low N_e/N ratios. Finally, additional sampling from all *B. parmifera* nursery sites will help us estimate the sizes of populations served by these nursery sites.

All egg case nursery sites that we sampled are currently listed as HAPCs, except IPZ1 (NOAA 2015). Given the genetic distinctiveness of embryos from IPZ1 and its low effective population size, we suggest that this site also merits consideration as an HAPC. Further, our research indicates that strengthening this designation to include protection from mobile bottom fishing gear and its known impacts on seafloor habitat and ecosystems is important for all egg case nursery areas (Kaiser et al. 2002, Thrush & Dayton 2002, Jennings et al. 2005). Commercial trawling within skate egg case nursery areas could potentially destroy large numbers of eggs that are crushed either during retention in the net or by the trawl footrope. For those egg cases that survive the mechanical aspects of being captured as bycatch and discarded during trawl operations, they will likely be displaced from their nursery area and may have increased mortality outside those specific habitats.

5. CONCLUSIONS

This study revealed new information on the population structure and use of nursery habitats by Alaska skates *Bathyraja parmifera* and Aleutian skates *B. aleutica*. Data showed that skate egg nursery sites in the eastern Bering Sea do not represent a single panmictic population, and that egg case deposition by females is non-random. Adult samples indicated that more nursery areas may exist, although our data did not provide evidence for the number or nature of those sites. Genetic differentiation did not follow an isolation-by-distance pattern. Nursery sites were genetically distinct, with the exception of the adjacent Bering and Pribilof Canyons, and may support relatively low effective population sizes, on the order of 1000. Results of this study also indicated very low N_e/N ratios in skates of the genus *Bathyraja*, similar to a study of thornback rays, which may be attributable to complex mating behavior. Multiple genetically diverse nursery areas may work in tandem to maintain genetic diversity in these species, similar to a metapopulation complex or the portfolio effect. Further complexity in samples from IPZ2 suggests that several distinct groups may utilize this nursery area. The results of this work justify the importance of the conservation of nursery areas as essential habitat for maintaining genetic diversity in *B. parmifera* and *B. aleutica*, and may serve to inform future studies of the population dynamics of these and other skates in the Bering Sea.

Acknowledgements. We thank Beth Matta, Olav Ormseth, and Florian Devloo-Delva, and an anonymous reviewer for thoughtful comments on the manuscript; Liz Dawson and Rebecca Haehn for assistance with DNA extraction; Krista Nichols for use of laboratory equipment; Isadora Jimenez-Hidalgo for guidance on laboratory techniques; Dan Drinan, Mary Fisher, Carolyn Tarpey, Per Erik Jorde, and Robin Waples for technical and bioinformatics assistance; and Katherine Maslenikov for curation. Thank you to all survey crew and captains who assisted in sample collections at sea and to Mike Canino for encouraging the use of next generation sequencing methodologies. This project was funded by the North Pacific Research Board.

LITERATURE CITED

- Anderson EC, Waples RS, Kalinowski ST (2008) An improved method for predicting the accuracy of genetic stock identification. *Can J Fish Aquat Sci* 65:1475–1486
- Anderson SC, Moore JW, McClure MM, Dulvy NK, Cooper AB (2015) Portfolio conservation of metapopulations under climate change. *Ecol Appl* 25:559–572
- Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat Rev Genet* 17:81–92
- Baird NA, Etter PD, Atwood TS, Currey MC and others (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE* 3:e3376
- Beck MW, Heck KL, Able KW, Childers DL and others (2001) The identification, conservation, and management of estuarine and marine nurseries for fish and invertebrates: a better understanding of the habitats that serve as nurseries for marine species and the factors that create site-specific variability in nursery quality will improve conservation and management of these areas. *BioScience* 51: 633–641
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci *de novo* from short-read sequences. *Genes Genomes Genet* 1:171–182
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22:3124–3140
- Chapman DD, Corcoran MJ, Harvey GM, Malan S, Shivji MS (2003) Mating behavior of southern stingrays, *Dasyatis americana* (Dasyatidae). *Environ Biol Fishes* 68:241–245
- Chevolut M, Ellis JR, Hoarau G, Rijnsdorp AD, Stam WT, Olsen JL (2006) Population structure of the thornback ray (*Raja clavata* L.) in British waters. *J Sea Res* 56:305–316
- Chevolut M, Ellis JR, Rijnsdorp AD, Stam WT, Olsen JL (2008) Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata*. *Heredity* 101: 120–126
- D'Aloia CC, Andrés JA, Bogdanowicz SM, McCune AR, Harrison RG, Buston PM (2020) Unraveling hierarchical genetic structure in a marine metapopulation: a comparison of three high-throughput genotyping approaches. *Mol Ecol* 29:2189–2203
- Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 22:1–20
- Drinan DP, Gruenthal KM, Canino MF, Lowry AD, Fisher MC, Hauser L (2018) Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (*Gadus macrocephalus*). *Evol Appl* 11:1448–1464
- Dudgeon CL, Ovenden JR (2015) The relationship between abundance and genetic effective population size in elasmobranchs: an example from the globally threatened zebra shark *Stegostoma fasciatum* within its protected range. *Conserv Genet* 16:1443–1454
- Farrugia TJ, Goldman KJ, Tribuzio C, Seitz AC (2016) First use of satellite tags to examine movement and habitat use of big skates *Beringraja binoculata* in the Gulf of Alaska. *Mar Ecol Prog Ser* 556:209–221
- Foll M, Gaggiotti O (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180: 977–993
- Franklin IR (1980) Evolutionary change in small populations. In: Soulé ME, Wilcox BA (eds) *Conservation biology: an evolutionary ecological perspective*. Sinauer Associates, Sunderland, MA, p 135–140
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical *F*-statistics. *Mol Ecol Notes* 5:184–186
- Gunderson DR (1997) Trade-off between reproductive effort and adult survival in oviparous and viviparous fishes. *Can J Fish Aquat Sci* 54:990–998
- Hara Y, Yamaguchi K, Onimaru K, Kadota M and others (2018) Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat Ecol*

- Evol 2:1761–1771
- ✦ Hermann AJ, Gibson GA, Bond NA, Curchitser EN and others (2016) Projected future biophysical states of the Bering Sea. *Deep Sea Res II* 134:30–47
- ✦ Heupel MR, Carlson JK, Simpfendorfer CA (2007) Shark nursery areas: concepts, definition, characterization and assumptions. *Mar Ecol Prog Ser* 337:287–297
- ✦ Heupel MR, Kanno S, Martins AP, Simpfendorfer CA (2019) Advances in understanding the roles and benefits of nursery areas for elasmobranch populations. *Mar Freshw Res* 70:897–907
- Hoff GR (2008) A nursery site of the Alaska skate (*Bathyraja parmifera*) in the eastern Bering Sea. *Fish Bull* 106:233–244
- ✦ Hoff GR (2009) Embryo developmental events and the egg case of the Aleutian skate *Bathyraja aleutica* (Gilbert) and the Alaska skate *Bathyraja parmifera* (Bean). *J Fish Biol* 74:483–501
- ✦ Hoff GR (2016) Identification of multiple nursery habitats of skates in the eastern Bering Sea. *J Fish Biol* 88:1746–1757
- Jennings S, Freeman S, Parker R, Duplisea DE, Dinmore TA (2005) Ecosystem consequences of bottom fishing disturbance. *Am Fish Soc Symp* 41:73–90
- ✦ Jombart T (2008) adegenet: an R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405
- ✦ Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071
- ✦ Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* 10:551–555
- ✦ Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Mol Ecol* 17:4015–4026
- ✦ Kai W, Kikuchi K, Tohari S, Chew AK and others (2011) Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in teleosts and mammals. *Genome Biol Evol* 3:424–442
- ✦ Kaiser MJ, Collie JS, Hall SJ, Jennings S, Poinier IR (2002) Modification of marine habitats by trawling activities: prognosis and solutions. *Fish Fish* 3:114–136
- ✦ Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281
- ✦ King JR, McFarlane GA (2010) Movement patterns and growth estimates of big skate (*Raja binoculata*) based on tag-recapture data. *Fish Res* 101:50–59
- ✦ Knutsen H, Olsen EM, Jorde PE, Espeland SH, Andre C, Stenseth NC (2011) Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Mol Ecol* 20:768–783
- ✦ Lande R (1994) Risk of population extinction from fixation of new deleterious mutations. *Evolution* 48:1460–1469
- ✦ Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
- Lauth B, Acuna E (2007) Results of the 2006 eastern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate resources. US Dep Commer NOAA Tech Memo NMFS-AFSC-176
- ✦ Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics* 26:2867–2873
- ✦ Marandel F, Lorange P, Berthelé O, Trenkel VM, Waples RS, Lamy JB (2019) Estimating effective population size of large marine populations, is it feasible? *Fish Fish* 20:189–198
- ✦ Marra NJ, Stanhope MJ, Jue NK, Wang M and others (2019) White shark genome reveals ancient elasmobranch adaptations associated with wound healing and the maintenance of genome stability. *Proc Natl Acad Sci USA* 116:4446–4455
- ✦ Martins APB, Heupel MR, Chin A, Simpfendorfer CA (2018) Batoid nurseries: definition, use and importance. *Mar Ecol Prog Ser* 595:253–267
- ✦ Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol Ecol Resour* 15:28–41
- Matta ME (2006) Aspects of the life history of the Alaska skate, *Bathyraja parmifera*, in the eastern Bering Sea. PhD dissertation, University of Washington, Seattle, WA
- ✦ Matta ME (2015) Reproductive biology of the Alaska skate *Bathyraja parmifera*, with comments on an intersexual individual. *J Fish Biol* 87:664–678
- ✦ McCallister M, Mandelman J, Bonfil R, Danylchuk A, Sales M, Ajemian M (2020) First observation of mating behavior in three species of pelagic myliobatiform rays in the wild. *Environ Biol Fishes* 103:163–173
- ✦ Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Mol Ecol Resour* 11:5–18
- ✦ Moran BM, Anderson EC (2019) Bayesian inference from the conditional genetic stock identification model. *Can J Fish Aquat Sci* 76:551–560
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, NY
- NOAA (National Oceanic and Atmospheric Administration) (2015) Fisheries of the exclusive economic zone off Alaska; skates management in the Bering Sea and Aleutian Islands Management Area; habitat areas of particular concern. *Fed Regist* 80:1378–1379
- ✦ O'Connell KA, Di Santo V, Maldonado J, Molina E, Fujita MK (2019) A tale of two skates: comparative phylogeography of North American skate species with implications for conservation. *Copeia* 107:297–304
- Ormseth O (2020) Assessment of the skate stock complex in the Bering Sea and Aleutian Islands. NPFMC Bering Sea and Aleutian Islands Stock Assessment and Fishery Evaluation. <https://apps-afsc.fisheries.noaa.gov/refm/docs/2020/BSAIs skate.pdf>
- ✦ Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420
- ✦ Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol Ecol Resour* 13:946–952
- ✦ Portnoy DS, McDowell JR, McCandless CT, Musick JA, Graves JE (2009) Effective size closely approximates the census size in the heavily exploited western Atlantic population of the sandbar shark, *Carcharhinus plumbeus*. *Conserv Genet* 10:1697–1705
- ✦ Purcell S, Neale B, Todd-Brown K, Thomas L and others (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575
- R Core Team (2020) R: a language and environment for statistical computing (version 4.0.3). R Foundation for Statistical Computing, Vienna
- ✦ Read TD, Petit RA III, Joseph SJ, Alam MT and others (2017) Draft sequencing and assembly of the genome of the world's largest fish, the whale shark: *Rhincodon typus* Smith 1828. *BMC Genomics* 18:532

- Rooper CN, Hoff GR, Stevenson DE, Orr JW, Spies IB (2019) Skate egg nursery habitat in the eastern Bering Sea: a predictive model. *Mar Ecol Prog Ser* 609:163–178
- Rousset F (2008) Genepop'007: a complete reimplementa-tion of the Genepop software for Windows and Linux. *Mol Ecol Resour* 8:103–106
- Schindler DE, Hilborn R, Chasco B, Boatright CP, Quinn TP, Rogers LA, Webster MS (2010) Population diversity and the portfolio effect in an exploited species. *Nature* 465: 609–612
- Spies IB, Punt AE (2015) 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
- Stevenson DE, Orr JW, Hoff GR, McEachran JD (2008) Emerging patterns of species richness, diversity, popula-tion density, and distribution in the skates (Rajidae) of Alaska. *Fish Bull* 106:24–39
- Suzuki R, Shimodaira H (2006) Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22:1540–1542
- Thrush SF, Dayton PK (2002) Disturbance to marine benthic habitat by trawling and dredging: implications for marine biodiversity. *Annu Rev Ecol Syst* 33:449–473
- Tørresen OK, Star B, Jentoft S, Reinar WB and others (2017) An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC Genomics* 18:95
- Wang Q, Arighi CN, King BL, Polson SW and others (2012) Community annotation and bioinformatics workforce development in concert — Little Skate Genome Annotation Workshops and Jamborees. *Database (Oxford)* 2012: bar064
- Waples RS (1998) Separating the wheat from the chaff: pat-terns of genetic differentiation in high gene flow species. *J Hered* 89:438–450
- Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilib-rium. *Mol Ecol Resour* 8:753–756
- Waples RS, Antao T, Luikart G (2014) Effects of overlapping generations on linkage disequilibrium estimates of effec-tive population size. *Genetics* 197:769–780
- Weir BS, Cockerham CC (1984) Estimating F -statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST} . *Am Nat* 186(Suppl 1):S24–S36
- Winter DJ (2012) MMOD: an R library for the calculation of population differentiation statistics. *Mol Ecol Resour* 12: 1158–1160
- Yano K, Sato F, Takahashi T (1999) Observations of the mating behavior of the manta ray, *Manta birostris*, at the Ogasawara Islands, Japan. *Ichthyol Res* 46: 289–296
- Yates PM, Heupel MR, Tobin AJ, Simpfendorfer CA (2012) Diversity in young shark habitats provides the potential for portfolio effects. *Mar Ecol Prog Ser* 458:269–281

Appendix. Additional figures summarizing read depth and heterozygosity for all samples

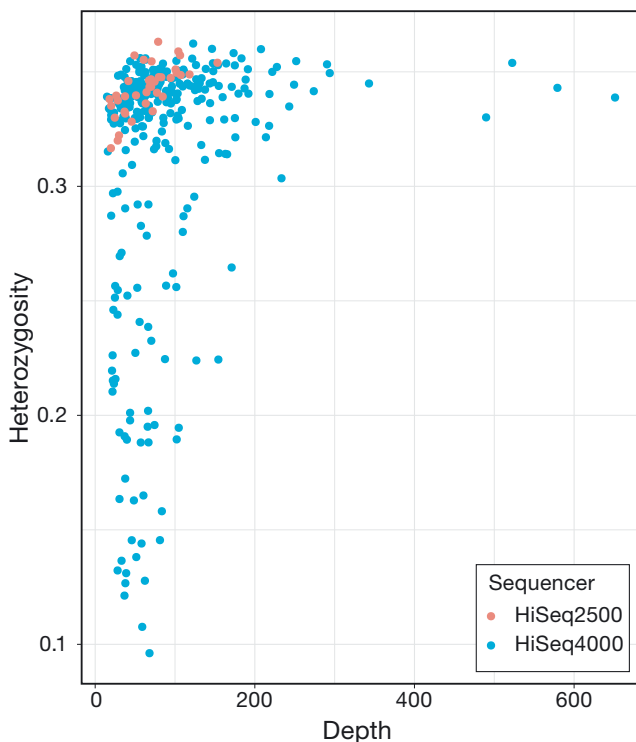


Fig. A1. Relationship between read depth and heterozygosity for all samples in the final dataset, with colors indicating which sequencer was used, (HiSeq 2500 or HiSeq 4000)

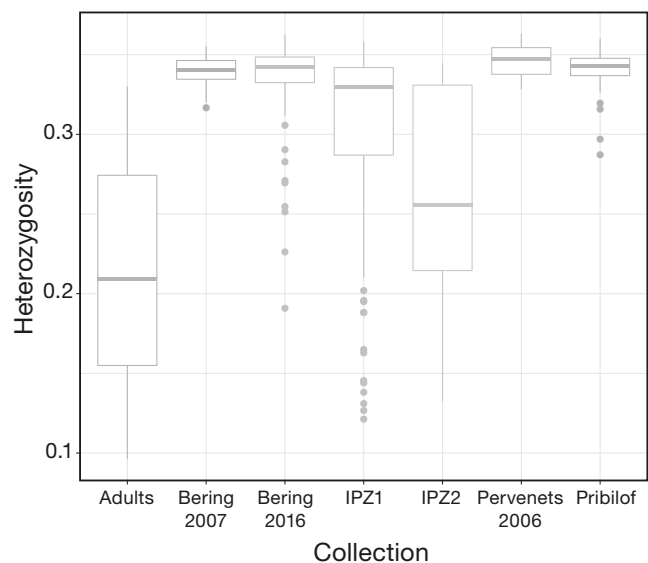


Fig. A2. Boxplots of heterozygosity for each collection of *Bathyraja parmifera*, : adults, and embryos from egg case nursery sites Bering 2007, Bering 2016, IPZ1, IPZ2, Pervenets 2006, and Pribilof. The middle of the boxplot represents the median and the range of the boxes represent the 25 and 75 percentiles