Re-examination of population structure in Arctic ringed seals using DArTseq genotyping

A. R. Lang[1],*, P. Boveng[2], L. Quakenbush[3], K. Robertson[4], M. Lauf[1], K. D. Rode[5], H. Ziel[2], B. L. Taylor[4]

[1]Ocean Associates, Inc., Arlington, VA 22207, USA, under contract to the Southwest Fisheries Science Center, NOAA Fisheries, La Jolla, CA 92037, USA

[2]Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA Fisheries, Seattle, WA 98115, USA

[3]Arctic Marine Mammal Program, Alaska Department of Fish and Game, Fairbanks, AK 99701, USA

[4]Marine Mammal and Turtle Division, Southwest Fisheries Science Center, NOAA Fisheries, La Jolla, CA 92037, USA

[5]Alaska Science Center, US Geological Survey, Anchorage, AK 99508, USA

ABSTRACT: Although Arctic ringed seals *Phoca hispida hispida* are currently abundant and broadly distributed, their numbers are projected to decline substantially by the year 2100 due to climate warming. While understanding population structure could provide insight into the impact of environmental changes on this subspecies, detecting demographically important levels of exchange can be difficult in taxa with high abundance. We used a next-generation sequencing approach (DArTseq) to genotype ~5700 single nucleotide polymorphisms in 79 seals from 4 Pacific Arctic regions. Comparison of the 2 most geographically separated strata (eastern Bering vs. Beaufort Seas) revealed a statistically significant level of genetic differentiation (F_{ST} = 0.001, p = 0.005) that, while small, was 1 to 2 orders of magnitude greater than expected based on divergence estimated for similarly sized populations connected by low (1 % yr^{-1}) dispersal. A relatively high proportion (72 to 88 %) of individuals within these strata could be genetically assigned to their stratum of origin. These results indicate that demographically important structure may be present among Arctic ringed seals breeding in different areas, increasing the risk that declines in the number of seals breeding in areas most negatively affected by environmental warming could occur.

KEY WORDS: *Phoca hispida hispida* · Population structure · Pacific Arctic · Climate warming · DArTseq

1. INTRODUCTION

Over the past several decades, Arctic warming has occurred at twice the rate of the global average (IPCC 2013), resulting in reductions in the extent, thickness, and seasonal duration of sea ice (Stroeve et al. 2012, Lindsay & Schweiger 2015, Wang & Overland 2015). These reductions are expected to have significant, but not uniform, consequences for species that depend on sea ice for important aspects of their life history (Moore & Huntington 2008, Kovacs et al. 2011, Laidre et al. 2015, Descamps et al. 2017). One such species is the ringed seal *Phoca hispida*, which builds subnivean lairs on the ice that are essential for protecting pups from thermal stress and predation (Smith & Stirling 1975, Smith 1976, Gjertz & Lydersen 1986, Lydersen & Smith 1989, Hammill & Smith 1991, Smith et al. 1991, Furgal et al. 1996). Although ringed seals in some areas appear to have adjusted to contemporary ice conditions (Bering and Chukchi Seas; Crawford et al. 2015), ringed seals in other regions have shown declines in body condition,

*Corresponding author: aimee.lang@noaa.gov

reproduction, and pup survival that are thought to be correlated with earlier spring sea ice retreat and declines in snow depth (Ferguson et al. 2005, 2017, Stirling 2005, Harwood et al. 2012a, 2015a, Iacozza & Ferguson 2014).

Ringed seals (subspecies *P. h. hispida*) are currently abundant and broadly distributed throughout the Arctic. While most of what is known about their population size is limited to surveys of only parts of their range (e.g. Bengtson et al. 2005, Conn et al. 2014), combined regional estimates suggest an abundance of around 3 million individuals (Laidre et al. 2015). Despite these large numbers, a status review conducted by NOAA Fisheries concluded that it is likely that the number of Arctic ringed seals will decline substantially by the year 2100 and that seals will no longer persist in substantial portions of their range in the foreseeable future (Kelly et al. 2010b). Following this review, Arctic ringed seals were listed as threatened under the US Endangered Species Act (77 FR 76706), a decision that was vacated in 2016[1] but subsequently reversed[2]. Although the status review considered the extinction risk to the Arctic subspecies as a whole, it was noted that such risk would be elevated if population structure exists (Kelly et al. 2010b). The effects of climate warming have not been uniform across the Arctic, with more marked sea ice declines in some areas than in others (Frey et al. 2014, 2015, Peng & Meier 2017, Wang et al. 2017). As such, ringed seals in some parts of the subspecies' range could be more severely impacted by Arctic warming than those in other areas.

The extent of such impacts will depend in part on the factors driving breeding habitat selection in Arctic ringed seals. If selection of breeding and pupping habitats is most strongly driven by habitat quality, then seals that once used a region that now has little to no ice or snow cover will move to less impacted regions in subsequent years. However, if seals exhibit high natal fidelity (i.e. return to reproduce in the same area where they were born), they may continue to return to an area of suboptimal quality even if high pup mortality and/or low breeding success occurs, as has been observed among Northwest Atlantic harp seals (Stenson & Hammill 2014). If this pattern of behavior persists, the number of seals breeding in those areas most negatively affected by environmental warming could decline, potentially leading to a loss of genetic diversity important to the evolutionary potential of the species.

The drivers behind breeding site selection in ringed seals are not well understood. Tagging studies have shown that ringed seals can range widely during the open water season (Heide-Jørgensen et al. 1992, Teilmann et al. 2000, Born et al. 2004, Freitas et al. 2008, Kelly et al. 2010a, Crawford et al. 2012, Harwood et al. 2012b). During the winter and spring subnivean period, however, Arctic ringed seals maintain much smaller ranges, and most tagged seals, including adult males and females as well as subadults, demonstrated breeding site fidelity across years (Kelly et al. 2010a, Martinez-Bakker et al. 2013, Luque et al. 2014, Harwood et al. 2015b). While this pattern of behavior indicates that the choice of where to reproduce is not random, the limited duration of tagging studies makes it difficult to identify the temporal scale over which breeding site fidelity is maintained and thus what factors influence its development.

If fidelity to breeding sites is driven by the return to natal sites, then genetic differences between regions should build over time. Estrous in ringed seals occurs during lactation (McLaren 1958, Atkinson 1997); thus, the timing and location of breeding and whelping are closely linked. Given this link, natal site fidelity could not only lead to maternally driven structure but may also influence gene flow between regions, as has been seen among ringed seals inhabiting Lake Saimaa, Finland (*P. h. saimensis*), where significant and substantial genetic differentiation between areas within the lake has been revealed at both mitochondrial and nuclear markers (mitochondrial F_{ST} = 0.444, nuclear F_{ST} = 0.107; Valtonen et al. 2012, 2014).

Detecting genetic differences between areas among Lake Saimaa seals is facilitated by the small effective size of this subspecies (~139−150 mature individuals; Sipilä 2016), which allows genetic differences between regions to develop more quickly. However, detecting restricted dispersal in highly abundant taxa, such as the Arctic ringed seal, is much more challenging. For example, F_{ST}, which is commonly calculated as a metric of connectivity between groups, can be estimated under Wright's island model as

$$F_{\text{ST-mtDNA}} \approx 1/(2N_e m - 1)$$

$$F_{\text{ST-SNP}} \approx 1/(4N_e m - 1)$$

where N_e represents the effective population size of the groups being compared (Wright 1965), m represents the fraction of immigrants within a group, and SNP is single nucleotide polymorphism. As described in Lowe & Allendorf (2010), the extent to which gene flow affects evolutionary processes (i.e. genetic con-

---

[1]Alaska Oil and Gas Association vs. National Marine Fisheries Service, case no. 4:14-cv-00029-RRB

[2]Alaska Oil and Gas Association vs. National Marine Fisheries Service, case no. 16-35380

nectivity) depends primarily on the absolute number of dispersers (i.e. $N_e m$), while demographic connectivity, or the degree to which population growth and vital rates are affected by dispersal, is dependent on the relative contribution of net immigration to total recruitment (i.e. m). Both types of connectivity are important, as genetic cohesion acts to maintain the evolutionary potential of the species, while demographic cohesion is needed to avoid depletion in the face of localized threats. Given the relationships described above in Wright's (1965) equation, the effect size (here, the magnitude of F_{ST}) associated with a particular threshold of genetic connectivity ($N_e m$) remains the same irrespective of population size. However, for the same m, the F_{ST} for a small population will be much larger, and much more easily detected, than the F_{ST} for a large population. For example, if we assume that the mature ringed seals that inhabit Lake Saimaa are evenly distributed across the 4 regions compared in the Valtonen et al. (2012) study and that those 4 regions are connected by approximately 1% dispersal per year (see Section 2.3.5 below for details), then the mtDNA F_{ST} would be 0.17. However, in a hypothetical situation in which ~15 000 ringed seals were evenly distributed across 4 regions in the Arctic and connected by the same level of dispersal, the mtDNA F_{ST} would be several orders of magnitude lower ($F_{ST} = 0.001$) and thus much more difficult to detect.

Thus far, genetic analyses have largely failed to detect population structure within Arctic ringed seals. Davis et al. (2008) compared the microsatellite genotypes (n = 11 loci) of seals sampled at 8 sites ranging from Saint Lawrence Island, Alaska, through the Canadian Arctic, Greenland, and Norway and into the White Sea in Russia. Though seals sampled in the White Sea were significantly differentiated from seals sampled at all other sites ($\Phi_{ST} = 0.0180$– 0.0306, p < 0.001), the remaining comparisons yielded very small estimates of differentiation, most of which were not significant. However, sample sizes for some of the areas were relatively low and may have included seals sampled outside of the breeding season. Subsequent analysis by Martinez-Bakker et al. (2013) used a panel of 9 microsatellite loci and included only samples collected during the breeding season at 9 sites ranging from the Chukchi Sea to the eastern Beaufort Sea. Though small but significant differences in mtDNA control region sequences were found between some strata, a geographic pattern among the sites with significant differences was not evident, with no differences detected among some of the most distant sites, while

in other cases neighboring sites showed statistically significant differences. In general, the degree of differentiation between sites was markedly low, and the authors concluded that gene flow among these Arctic sites was relatively high (Martinez-Bakker et al. 2013).

While the failure to identify clear patterns of genetic differentiation among Arctic ringed seals suggests genetic connectivity between ringed seal breeding sites, it is possible that weak, but demographically important, population structure exists but was not detected in these previous studies. Traditional genetic markers, such as those used in the above studies, may have little statistical power to detect small, but biologically significant, levels of genetic differentiation between areas (Waples & Gaggiotti 2006, Lowe & Allendorf 2010). However, recent advances in high-throughput next-generation sequencing (NGS) technologies have provided a cost-effective means to simultaneously discover and genotype large numbers (hundreds to thousands) of SNPs, even in species for which little to no genomic information is available (Baird et al. 2008, Davey et al. 2011, Peterson et al. 2012). The production of these extensive datasets can substantially increase the power and precision of genetic analyses even with limited sample sizes (e.g. Willing et al. 2012, Nazareno et al. 2017), allowing previously undetected patterns of demographic and evolutionary structure to be resolved (Corander et al. 2013, Reitzel et al. 2013, Benestan et al. 2015). In addition, genome-wide scans of diversity using NGS technologies allow outlier loci, putatively under divergent selection, to be identified (Benestan et al. 2016, Gleason & Burton 2016); analysis of these loci can reveal patterns of adaptive variation with important implications for developing conservation and management strategies (Funk et al. 2012, 2018).

In this study, we take advantage of recent advances in NGS to discover and genotype a large panel of SNP markers (n = 5699 loci) in ringed seals (n = 79) sampled during the breeding season in 4 regions of the Pacific Arctic and use this dataset to evaluate population structure in Arctic ringed seals. Compared to previous studies, which relied on ≤11 microsatellite loci, the number and genome-wide placement of the SNPs analyzed should greatly increase the power to detect subtle differences between regions. The results of this study will increase our understanding of genetic and demographic connectivity between regions so that the effect of regional depletions of Arctic ringed seals by hunting or climate warming can be evaluated.

2. MATERIALS AND METHODS

2.1. Samples

Tissue samples (n = 113) were collected from ringed seals in the Pacific Arctic region during the spring breeding season (March through May between 2000 and 2017; Fig. 1) and archived in the Southwest Fisheries Science Center's Marine Mammal and Sea Turtle Research Tissue Collection. The majority of these samples (n = 59) were collected during a biomonitoring program run by the Alaska Department of Fish and Game from seals legally harvested for subsistence by Alaska Natives. The remaining samples were collected from the remains of seals that were killed by polar bears (n = 54). Following collection, the samples were either stored at −80°C or were preserved in a salt-saturated 20% DMSO solution or 100% ethanol and subsequently archived in a −20°C freezer.

Ringed seals are distributed throughout the study area, making it difficult to identify biologically relevant boundaries by which to divide groups for comparison. Thus, for analyses requiring that *a priori* groups be identified for comparison, we stratified the samples into 4 geographic areas. These strata correspond to 4 of the 5 Alaska Native regions that harvest ice seals and are represented by the Ice Seal Committee, which is the tribally authorized Alaska Native

organization that co-manages ice seals in partnership with NOAA Fisheries. The fifth and southernmost Alaska Native region is Bristol Bay. Few samples were available from this area; thus, it was not included as a stratum in our analysis. The 4 strata analyzed include (1) the eastern Bering Sea, (2) the northern Bering Sea, (3) the southeastern Chukchi Sea, and (4) the Beaufort Sea (Fig. 1). The eastern Bering Sea and the Beaufort Sea strata also align with 2 of the 3 biogeographic provinces (Eastern Bering Shelf and Beaufort Sea provinces, respectively) that have been delineated within the study area (Sigler et al. 2011), although the third province (Chirikov–Chukchi Province) includes both the northern Bering and southeastern Chukchi strata used in our analysis. A list of sample information is included in Table S1 in the Supplement at www.int-res.com/articles/suppl/n044p011_supp.pdf.

2.2. DArTseq library preparation and high-throughput sequencing

Many NGS methods rely on the use of restriction enzymes (REs) to reduce genome complexity prior to high-throughput sequencing. DArTseq (RADseq; Davey et al. 2011), which was developed by Diversity Arrays Technology (DArT P/L), is one such method (Jaccoud et al. 2001, Sansaloni et al. 2011, Kilian et
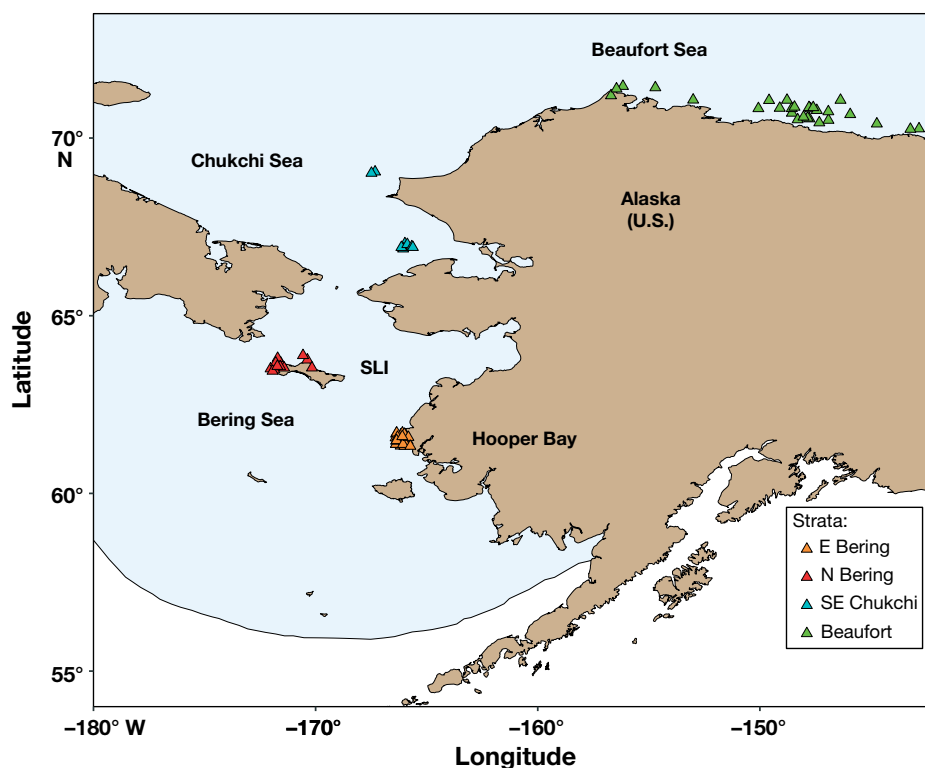

Fig. 1. Collection locations for analyzed samples. The approximate range of ringed seals within this area is shown in light blue. Triangles indicate the collection locations (some of which are approximated) for the analyzed samples. The color of symbols denotes the stratum to which the collected sample belongs (see legend). Places mentioned in the text are shown, including Hooper Bay and Saint Lawrence Island (SLI)

al. 2012). In brief, this method entails a complexity reduction step in which DNA is digested with a combination of REs which act to exclude repetitive regions of the genome while targeting low copy sequences. The resulting libraries are then sequenced on a high-throughput sequencing platform, allowing for high read coverage of regions most likely to be informative in population studies.

DNA was extracted from these samples using the Machinery-Nagel NucleoMag® tissue extraction kit and following the manufacturer's protocol (see www.mn-net.com), with the exception that the proteinase K tissue lysis step was extended to include an overnight digestion at 37°C followed by a 3 h digestion at 55°C. DNA was quantified on a fluorometer using Quant-iT PicoGreen, and DNA integrity was assessed by electrophoresing 100 ng of DNA on a 1% ethidium bromide (EtBr)-stained agarose gel at 70 to 80 V for approximately 1 h alongside a 1 kb DNA size standard. Those extracts that demonstrated the presence of high molecular weight DNA on the gel and that contained at least 500 ng of DNA were retained in the study, and additional samples were selected as needed to replace those that produced insufficient DNA.

A trial RE digest was conducted for a subset (n = 6) of the selected samples to ensure that DNA quality was sufficient for restriction digest to be successful. For this trial, 100 ng of DNA from each sample was digested in a 50 µl reaction that contained 1 µl (10 units) of the HindIII RE, 5 µl of NEB buffer, and (for the remaining volume) DNase/RNase-free distilled water. The digestion was conducted in a thermocycler for 3 h at 37°C. The resulting digests were electrophoresed alongside the original undigested DNA extractions (from the same samples) and 5 µl of lambda DNA-HindIII digest on a 1% agarose gel (pre-stained with EtBr) at 50 V for a total of 2 h.

Although sample selection was initially based on maintaining relatively even coverage of the 4 regions, not all of the samples chosen produced DNA of sufficient quantity and quality for the DArTseq library preparation protocol. In total, 89 samples from the eastern Bering Sea (n = 34), northern Bering Sea (n = 14), southeastern Chukchi Sea (n = 9), and Beaufort Sea (n = 32) were included (Table S1). Approximately 50 to 100 ng of DNA from each of these samples was diluted in 10 to 20 µl TE buffer and then shipped to DArT P/L at the University of Canberra, Australia, for library preparation and sequencing.

Once at the DArT P/L laboratory, DNA quality was re-evaluated. For a subset of samples, the DNA pro-vided was divided in 2 parts prior to digestion/ligation to act as technical replicates to be assessed for scoring consistency (referred to as the reproducibility score, used in filtering below). Double digestion was performed using methylation-sensitive REs as described by Kilian et al. (2012). The only modification to this protocol was that the single PstI-compatible adaptor was replaced with 2 different adaptors corresponding to the PstI and SphI RE overhangs. The PstI-compatible adapter was designed to include the Illumina flowcell attachment sequence, a sequencing primer, and a staggered, varying length barcode region, similar to the sequence reported by Elshire et al. (2011). The SphI-compatible adapter simply comprised the Illumina flowcell attachment region and SphI overhang sequence. Ligated fragments with both a PstI and SphI adaptor were amplified by PCR using an initial denaturation step of 94°C for 1 min, followed by 30 cycles with the following temperature profile: denaturation at 94°C for 20 s, annealing at 58°C for 30 s, and extension at 72°C for 45 s, with an additional final extension at 72°C for 7 min. Equimolar amounts of amplification products from each sample were combined before single-end sequencing for 77 cycles on an Illumina Hiseq2500 to yield approximately 2.5 million reads per sample.

Samples were genetically sexed by amplification and real-time PCR (Robertson et al. 2018). Samples from 1 male and 1 female for which sex had been determined via examination of a stranded animal were included as positive controls in all amplifications. Sex was determined by the amplification pattern: males had 2 products and females had 1.

2.3. Data analysis

2.3.1. Pipeline processing

The FASTQ-formatted sequences generated from the sequencing lane were processed using proprietary diversity array technology (DArT P/L) analytical pipelines. In summary, the primary pipeline filters out poor-quality sequences and corrects low-quality bases from singleton tags using collapsed tags as a template. More stringent selection criteria are applied to the barcode region to ensure that sequences are reliably assigned to the appropriate sample. Identical sequences are then collapsed into fastqcoll files that are processed through the secondary DArT P/L pipeline, which uses proprietary SNP and SilicoDArT (presence/absence of restriction fragments in representation) calling algorithms (DArTsoft14) to produce

genotypes for each sample. Multiple samples were processed from DNA to allelic calls as technical replicates; those loci with reproducibility scores less than 90% across replicates were removed. Calling quality was assured by high average read depth per locus (mean across all markers of 31.2 reads per locus, minimum of 5 reads per locus). The mean call rate across loci was 84.8%. Given that a ringed seal reference genome is not yet available, mapping efficiency of reads against such a reference could not be evaluated. The data were converted to a matrix of SNP loci by individuals, with the contents stored as integers: 0, homozygote, reference state; 1, heterozygote; and 2, homozygote, alternate state.

2.3.2. Additional filtering

For this analysis, only the codominant SNP genotype data (allSnps_singlerow_dartseq.csv) received from DArT P/L were analyzed. The genotype data were imported into R (v3.5.0) and then filtered to remove SNPs that had average read counts less than 10. SNPs with high read depth, which can represent false heterozygotes caused by copy number variations or paralogous sequences, were also re‐moved. For this criterion, we first calculated the average read depth (d) across all loci and then removed those loci with average read depths greater than $d + 4 \times \sqrt{d}$, which has been shown to be effective in reducing the number of false heterozygotes (Li 2014). We then used the package dartR 1.1.11 (Gruber et al. 2018) to remove loci that (1) had reproducibility scores, which were calculated based on the technical replicates described above, less than 100%; (2) had been called in less than 90% of individuals; (3) were monomorphic; and (4) were secondary SNPs (e.g. SNPs that were found on the same sequence fragment as 1 or more other SNPs), in which case only the locus with the highest polymorphic information content was retained. After this filtering was completed, we removed individuals that were missing >20% of the data.

The remaining filtering steps were conducted in R using the strataG package (Archer et al. 2017). These steps included removing (1) loci with minor allele frequencies ≤0.05, (2) any loci that were not bi-allelic, (3) loci that had expected heterozygosity greater than 0.6, (4) one of each pair of samples that shared 80% or more of their alleles (i.e. samples that were likely collected from the same individual), and (5) loci that were out of Hardy-Weinberg equilibrium in 3 or more of the *a priori* defined geographic strata.

2.3.3. Identification of outlier loci and analysis of relatedness

We used the R package OutFLANK v0.2 (Whitlock & Lotterhos 2015) to identify SNPs that might be under diversifying selection. This approach is based on identifying loci with atypical values of F_{ST} as outliers. OutFLANK uses an improved method for deriving the null distribution of population differentiation for presumed neutral loci and results in fewer false positives than some other methods. We ran OutFLANK with a 5% left and right trim of the null distribution of F_{ST} and a false discovery rate (q-value) of 5%.

The R package related 1.0 (Pew et al. 2015) was used to estimate pairwise relatedness (r) between all pairs of genotyped individuals based on the allele frequencies of the full sample set. The 'compareestimators' function was then used to generate simulated genotypes for 200 dyads representing 4 categories of known relatedness (unrelated, half sibling, full sibling, and parent–offspring) based on the allele frequencies of a subset of the SNP data (n = 500 randomly chosen loci). The distribution of r values for simulated individuals within each category was plotted against those generated from the sampled individuals to visualize how well our dataset was able to discriminate between relationship categories. Although related allows for coefficients to be generated under 7 different estimators, here we present only the results based on the Lynch & Ritland (1999) estimator, which had the highest correlation coefficient between the observed and expected relatedness values.

2.3.4. Comparisons between geographic strata

The R package strataG (Archer et al. 2017) was used to generate summary statistics describing the genetic diversity retained in each stratum, including the observed and expected heterozygosity and the number of alleles. This package was also used to estimate genetic differentiation (F_{ST}) between pairs of the *a priori* defined geographic strata. Significance of comparisons was evaluated via bootstrapping across 1000 permutations. A correction for multiple tests was not applied when interpreting the results of the pairwise comparisons, as each comparison was testing an independent hypothesis. Corrections for multiple tests effectively reduce the critical value (α), or Type I error rate, at the expense of the Type II error rate (Perneger 1998); as such, inappropriate application of correction factors can have serious conservation management implications.

To evaluate the impact of sample size on our comparisons, we randomly subsampled the 2 largest strata (the eastern Bering Sea stratum and the Beaufort Sea stratum) to create 20 datasets in which the sample sizes were equivalent to (1) the sample size of the southeastern Chukchi Sea stratum (n = 8), (2) the sample size of the northern Bering Sea stratum (n = 14), (3) a sample size of 20, and (4) a sample size of 25. Pairwise comparisons were then conducted for these datasets (60 in total), and the proportion of significant tests was counted.

Pairwise comparisons between regions were repeated after subdivision of strata by sex to explore whether sex-based differences in structure could be detected.

2.3.5. Calculating expected effect size

When interpreting the results of genetic analyses to delineate management units, Palsbøll et al. (2007) advocate placing the focus on the amount of genetic divergence rather than on rejecting panmixia, to address the case where statistical power is too low to reject panmixia among demographically independent units and where the statistical power is sufficient to reject panmixia even among demographically correlated groups. This approach involves, first, determining the amount of population genetic divergence that corresponds to the relevant dispersal rate and, second, comparing the empirical results with this criterion, as done by Martien et al. (2012).

Here, we used this approach and Wright's (1965) equation (see Section 1) to calculate the expected magnitude of F_{ST} between demographically independent stocks of Arctic ringed seals. Limited information is available on the dispersal rate at which groups become demographically independent, although simulations based on cetaceans have shown that if the dispersal rate between 2 groups is less than a few percent per year, 1 group could eventually be extirpated if the groups are managed as a single unit and human-caused mortality is disproportionately focused on only one of the groups (Taylor 1997). Generation length for ringed seals was estimated to be approximately 18 yr by Pacifici et al. (2013), based on a maximum life span of 46 yr and an age of first reproduction of 6.9 yr. No information on the generation length of ringed seals within our study area was available, but to identify a lower bound on generation time, we also calculated the effect size based on a maximum life span of 15 yr (the average age of females in Lydersen & Gjertz

1987) and an age of first reproduction of 4 yr. Based on these parameters, a 1% dispersal rate per year translates into a per generation rate of 0.07 to 0.18 (*m* in the equation noted in Section 1). Although the abundance of ringed seals in the strata we compared is not known, aerial surveys of the eastern Chukchi Sea indicated that ~231 000 seals were present in May and June of 1999 and 2000 (Bengtson et al. 2005), while aerial surveys of the eastern Bering Sea during April and May of 2012 indicated that ~170 000 seals were present (Conn et al. 2014). Using these abundance estimates to represent the census size (N_c) of our strata and further assuming that ringed seals have an $N_e{:}N_c$ ratio at the low to middle range of that typical of mammals (0.2; Frankham 1995), we calculated an input N_e value that ranged from 34 000 to 46 200 seals.

2.3.6. Clustering analyses

Two methods were used to evaluate the number of genetic clusters (K) that could be detected in our dataset. Model-based Bayesian clustering was conducted using the program STRUCTURE 2.2.3 (Pritchard et al. 2000). This analysis was based on a model of admixture with correlated allele frequencies. Given that the inclusion of related individuals may lead to false inferences in STRUCTURE analyses (Anderson & Dunham 2008, Rodríguez-Ramilo & Wang 2012), 1 of the individuals in the single putative paternal half-sibling pair that was identified (described in Section 3.1) was removed prior to conducting this analysis. In addition, the number of sampled individuals per stratum has been shown to have a strong impact on the ability of STRUCTURE to correctly estimate the number of clusters (Fogelqvist et al. 2010, Puechmaille 2016). Given that the sample sizes representing the eastern Bering Sea stratum (n = 28) and the Beaufort Sea stratum (n = 29) were similar and markedly larger than those representing the other 2 strata (n = 14, northern Bering Sea; n = 8, southeastern Chukchi Sea), the STRUCTURE analysis was run using both the full SNP dataset (all 4 strata) and only the 2 strata with similar and larger sample sizes. Both the 2- and 4-strata analyses were initially run with no *a priori* information on geographic origin of the samples included. To evaluate whether the performance of STRUCTURE improved when information on the geographic location of sampling was included *a priori*, the 2-strata analysis was also run with the inclusion of a location prior (i.e. LOCPRIOR = 1).

For all 3 analyses, the model was run using values of *K* ranging from 1 to 10, with 3 replicate runs for each value of *K* and a burn-in of 200 000 steps followed by 1 000 000 Markov chain Monte Carlo iterations. The output from replicate runs was summarized using STRUCTURE HARVESTER (Earl 2012). The most parsimonious number of clusters present within the dataset was inferred based on the *K* value that had the highest lnPr(*X*|*K*) or, for *K* > 2, the ad hoc statistic delta *K* (Evanno et al. 2005). Membership scores for each value of *K* were aligned and averaged over replicates using CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007), and barplots were generated to visualize the results using the R package pophelper (Francis 2017).

While STRUCTURE forms genetic clusters of individuals by minimizing departure from Hardy-Weinberg and linkage disequilibria, discriminant analysis of principal components (DAPC) is a non-model based method that optimizes genetic separation between groups and minimizes variation within groups (Jombart et al. 2010). We ran DAPC using the R package adegenet v2.1.3 (Jombart 2008) and following the recommendations of the online tutorial provided by the developers (Jombart & Collins 2015). First, the function 'find.clusters' was run to identify the most likely number of clusters in the data in the absence of *a priori* sampling information. This function converts the genotypic data into uncorrelated principal components (PCs) and ranks the clustering solutions for different values of *K* using the Bayesian information criterion (BIC). For the ringed seal dataset, BIC increased in a roughly linear fashion with increasing values of *K*, and the contribution of additional PCs to the total variance did not asymptote (Fig. S1a). The lowest BIC was found at *K* = 1, indicating that no distinct genetic clusters were detected within the SNP dataset (Fig. S1b).

Given this result, we proceeded by running the DAPC (function 'dapc' in adegenet) with sampling locality provided as a prior. We initially ran the DAPC with the number of PCs retained limited to the number of samples divided by 3, which is the maximum number recommended by the developers. We then used the function 'optim.a.score' with 100 simulations to identify the optimal number of PCs to retain without leading to overfitting of the discriminant functions, which can result in spurious discrimination of any set of clusters. The *a*-score is a measure of the proportion of successful re-assignments to the *a priori* defined clusters corrected for the number of retained PCs. The DAPC was then re-run using the optimal number of PCs and retaining 3 discriminant

functions (the number of putative populations minus 1, as recommended). The function 'scatter.dapc' in adegenet was then used to plot the results of the DAPC analysis when sampling locality was used as a prior. Inertial ellipses were drawn to encompass 67% of the variance (the default value) within each stratum.

2.3.7. Assignment test

We used the R package assigner v0.5.7 (Gosselin et al. 2016) to evaluate whether our dataset could be useful in assigning individuals to the group in which they were sampled. Given the small sample sizes for the northern Bering Sea (n = 14) and particularly the southeastern Chukchi Sea (n = 8) stratum, we restricted these analyses to using only the eastern Bering Sea (n = 28) and the Beaufort Sea (n = 29) strata, which had relatively high sample sizes and also represent the 2 most geographically disparate strata. Although the sample sizes were similar for these 2 strata, we created 10 subsets of the data by randomly subsampling 28 individuals from each stratum to avoid any bias created by uneven sample sizes. Assigner employs the gsi_sim algorithm (Anderson et al. 2008) to conduct a self-assignment analysis using a leave-one-out (LOO) approach. Each individual is sequentially removed from the set of baseline samples, and the remaining individuals are used to calculate F_{ST} at each locus. The loci are ranked according to the F_{ST} values, and the chosen number of loci are retained based on that ranking. These loci are then used to assign the LOO individual to the stratum that has the highest probability of producing its genotype. This approach (training, holdout, LOO [THL] method) was used to avoid high-grading bias (Anderson 2010) while maximizing the sample size of individuals retained in the training dataset from which the baseline allele frequencies for each stratum are calculated. Assignment success is then estimated as the proportion of individuals that are correctly assigned to their sampling location of origin. The analysis was run both with and without imputing missing genotype data; imputation was conducted using the random forest algorithm based on 100 trees. For each subsample, we ran the THL analysis for 10 iterations and with the number of ranked SNPs ranging from 10% of the total number to the full dataset.

To further evaluate the results of the assignment test, we created 20 subsets of the data in which the individuals included in the original analysis were randomly assigned to 1 of 2 strata without replace-

ment. The analysis described above was then repeated with the proportion of ranked SNPs ranging from 50 to 80% of the total. The assignment success based on these randomized datasets was then compared to the empirical results.

3. RESULTS

The analyses presented here utilized only the codominant SNP data provided by DArT P/L, which included 100 281 SNPs genotyped. The species of 2 of the 89 samples genotyped were determined via mitochondrial DNA control region sequencing to have been misidentified, and these 2 samples (1 from the southeastern Chukchi Sea stratum and 1 from the Beaufort Sea stratum) were removed from the analysis prior to filtering. A summary of the number of loci and/or samples removed at each filtering step is shown in Table 1. Our final SNP dataset included 79 individuals with 80% or greater locus coverage for 5699 bi-allelic SNP loci. Measures of SNP diversity were similar among strata (Table 2), with average expected heterozygosity ranging from 0.250 to 0.263 and average observed heterozygosity ranging from 0.219 to 0.228.

3.1. Identification of outlier loci and relatedness analysis

The OutFLANK analysis did not identify any loci as being putatively under selection given a q-value threshold of 0.05. The OutFLANK approach generally has a low rate of false positives and has been noted as being only suited for identification of loci under strong spatially diversifying selection (Whitlock & Lotterhos 2015).

The mean relatedness coefficient calculated for all pairs of genotyped individuals was −0.013 (SD = 0.0144). All except 1 pair of individuals had coefficients that fell at or below 0.066. The coefficients estimated from the sampled individuals fell within the range of those calculated from the simulated genotypes of unrelated individuals (−0.108 to 0.122, median = −0.002; Fig. S2), where unrelated refers to individuals that do not fall in the other 3 categories (i.e. parent–offspring, full

sibling, half sibling). The one exception to this pattern was a pair of individuals that were sampled during the same year in Hooper Bay, Alaska, and had a relatedness coefficient of 0.237. This value is similar to what would be expected between half-sibling pairs, and it fell in the middle of the distribution calculated for the simulated half-sibling pairs. Review of the harvest data associated with these 2 samples revealed that they were part of the same cohort (one was estimated to be 10 mo old, the other 12 mo old, both males). Given that they did not share the same mtDNA haplotype, these 2 individuals appear to be paternal half siblings that are the offspring of a single male mating with 2 different females in the same year (2007).

3.2. Comparisons between geographic strata

Overall differentiation was low but significant (F_{ST} = 0.001, p = 0.028; χ^2 p-value = 0.012). Pairwise comparisons between the *a priori* defined geographic strata revealed low but statistically significant levels of genetic differentiation between the eastern Bering Sea stratum and the Beaufort Sea stratum (F_{ST} = 0.001, p = 0.005; Table 3). With the exception of the comparisons

Table 1. Number of putative single nucleotide polymorphisms (SNPs) retained after each filtering step. Samples flagged for removal included n = 2 samples identified in the mtDNA analysis as bearded seals, n = 2 from stranded individuals that may not have died during the breeding season, n = 3 samples collected from regions outside the study area (Okhotsk Sea), and n = 2 samples identified as duplicates (i.e. samples collected from the same animal). Secondary loci are defined as SNPs found on the same sequence fragment as 1 or more other SNPs. When secondary loci were identified, only the locus with the highest polymorphic information content was retained. *d*: average read depth across all loci; HWE: Hardy-Weinberg equilibrium

Criterion	No. of individuals	No. of loci removed	No. of loci remaining
Unfiltered dataset[a]	90	0	100 281
Avg. read count <10	90	18 613	81 668
Avg. read count $> d + 4 \times \sqrt{d}$[b]	90	16 483	65 185
Remove flagged individuals	81	13 346	51 839
Reproducibility <1.0	81	5700	46 139
Call rate ≤0.9	81	17 031	29 108
Secondary loci	81	8584	20 524
Individual coverage <0.8	79	148	20 376
Minor allele frequency <0.05	79	14 596	5780
Heterozygosity >0.60	79	0	5780
Out of HWE in 3 or more strata	79	81	5699

[a]Four samples failed to pass the DArT quality assurance/quality control screening process and were removed prior to any filtering;
[b]This step is intended to remove loci with excessive coverage, which may represent paralogs or repetitive elements.

Table 2. Single nucleotide polymorphism diversity indices for samples after subdivision into strata. n: no. of samples; P: proportion of loci genotyped within the stratum, H_e: mean expected heterozygosity across loci; H_o: mean observed heterozygosity across loci

Stratum	n	P	H_e	H_o
E Bering	28	0.982	0.260	0.219
N Bering	14	0.987	0.257	0.228
SE Chukchi	8	0.984	0.250	0.223
Beaufort	29	0.979	0.263	0.224
All	79	0.991	0.267	0.223

involving the northern Bering Sea, the magnitude of the remaining pairwise comparisons was similar (F_{ST} = 0.001) but not significant.

When individuals were randomly subsampled from the eastern Bering Sea and the Beaufort Sea strata and then compared to each other, only 2 of 20 comparisons were statistically significant with a sample size of 8, while 4 of 20 were statistically significant with a sample size of 14 (Table S2). When sample sizes were increased to 20 per stratum, 14 of 20 were statistically significant, and at 25 samples per stratum, all but one test (of 20) were significant.

When the samples were subdivided by sex, the comparison between the eastern Bering Sea and the Beaufort Sea remained statistically significant for both males and females (females, F_{ST} = 0.003, p = 0.016; males, F_{ST} = 0.002, p = 0.016). No other sex-specific comparisons were significant (Table 3).

3.3. Calculating expected effect size

If we assume a generation length between 7 and 18 yr, a dispersal rate of 1% per year, and effective population sizes ranging from 34 000 to 46 200 seals, then we would expect a maximum $F_{ST\text{-}SNP}$ value between 1×10^{-4} (minimum 3×10^{-5}). The empirical estimates of F_{ST} based on the SNP data (F_{ST} = 0.001) are approximately 1 to 2 orders of magnitude higher than these expected effect sizes for all pairwise comparisons, except for 2 (both of which include the northern Bering Sea).

3.4. Clustering analyses

The results of the STRUCTURE analysis using the full SNP dataset (with 1 of the samples comprising the putative half-sibling pair removed) are shown in Fig. 2a. The ΔK method (Evanno et al. 2005) identified $K = 2$ as the most likely number of clusters (Fig. S3). However, at $K = 2$, there was no clear geographic pattern to the assignments, although some individuals did assign strongly to one of the putative clusters (Fig. 2a). However, in general the number of individuals that were assigned to each cluster was roughly even, and many individuals were relatively admixed, both of which are suggested to be indicative of a lack of population structure (Pritchard et al. 2010). Given these results, and that the mean estimated lnP(K) was highest for $K = 1$ (albeit only slightly), which cannot be evaluated using the ΔK method, we consider it most likely that there is a single cluster. When a priori information on the geographic location of sampling was incorporated, the results were similar for $K = 2$, and runs for $K = 3$ and higher tended to lump all individuals in a single cluster, with the exception of those samples with the highest amounts of missing data (Fig. 2b).

No further resolution was obtained when we repeated the STRUCTURE analyses using only the eastern Bering Sea and the Beaufort Sea strata (Fig. 2c), which contained similar sample sizes and thus should not be subject to biases due to uneven sampling. Although the ΔK method identified $K = 5$ as the optimal clustering solution, the mean estimated lnP(K) remained highest for $K = 1$ (Fig. S3), and no geographic pattern was apparent when the barplots showing individual assignments were examined (Fig. 2c).

As noted in Section 2.3.6, running the K-means clustering method in the absence of information on the geographic origin of samples found that the lowest BIC was for a single cluster (Fig. S1b). We then ran the DAPC model with the samples stratified a priori by geographic location and with the number of PCs set to the number of samples divided by 3, as recommended by the developers. Optimizing the a-score for this initial DAPC model indicated that 22 PCs should be retained. However, the a-score was low (a-score maximum = 0.148), indicating weak or unstable discrimination of individuals into the a priori defined groups. A clear peak in the a-score across the range of PCs tested was not observed (Fig. S1c). The 3 discriminant functions that were retained (the number of groups minus 1) explained 33.2% of the variance. The scatter plots showed only slight separation between the mean values representing each group, and the ellipses encompassing 67% of the variance within groups showed substantial overlap (Fig. 3).
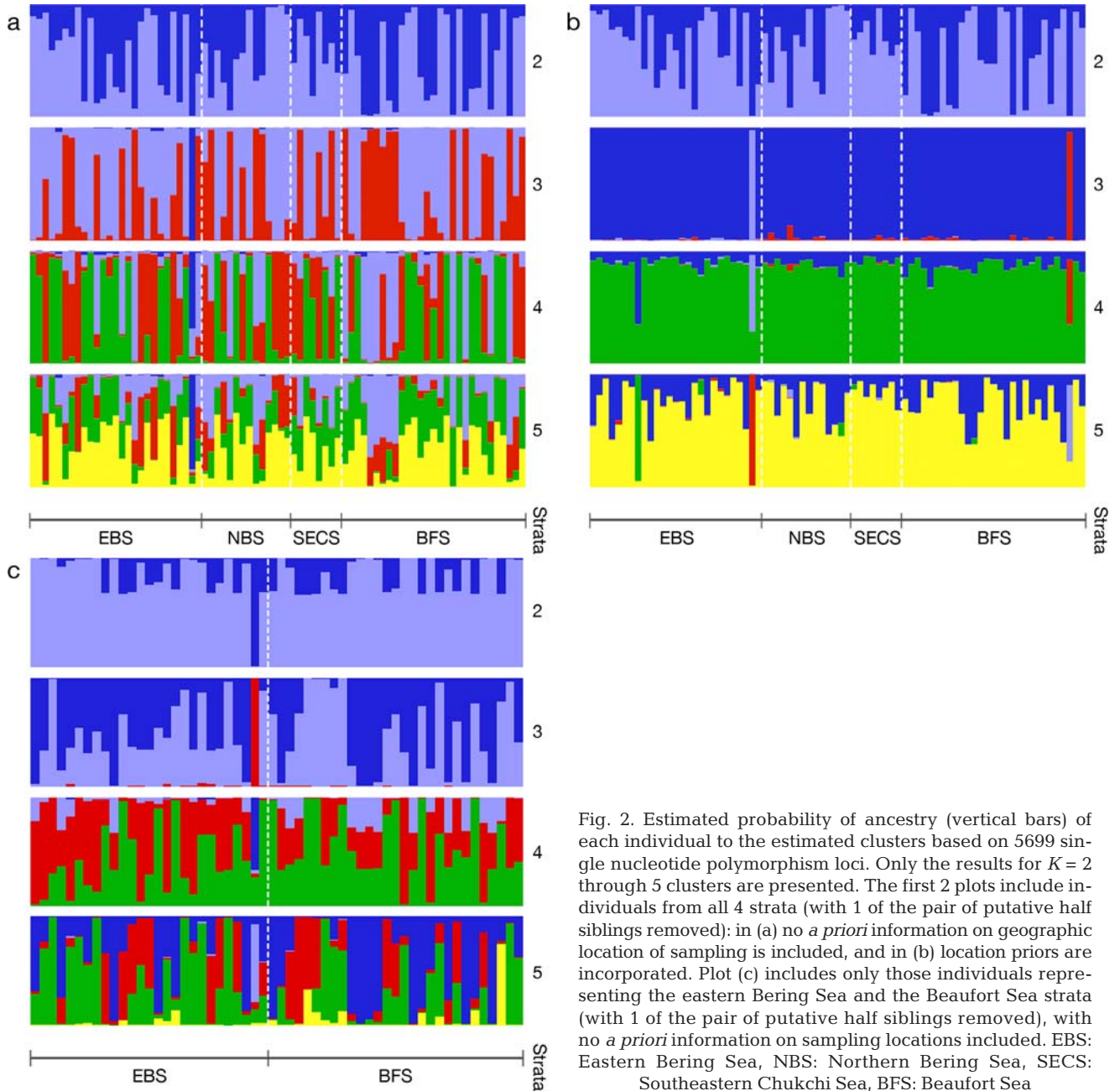
Fig. 2. Estimated probability of ancestry (vertical bars) of each individual to the estimated clusters based on 5699 single nucleotide polymorphism loci. Only the results for $K = 2$ through 5 clusters are presented. The first 2 plots include individuals from all 4 strata (with 1 of the pair of putative half siblings removed): in (a) no *a priori* information on geographic location of sampling is included, and in (b) location priors are incorporated. Plot (c) includes only those individuals representing the eastern Bering Sea and the Beaufort Sea strata (with 1 of the pair of putative half siblings removed), with no *a priori* information on sampling locations included. EBS: Eastern Bering Sea, NBS: Northern Bering Sea, SECS: Southeastern Chukchi Sea, BFS: Beaufort Sea

3.5. Assignment test

Assignment success was maximized when missing genotype data were imputed using the random forest algorithm; thus, only those results are reported here. When overall assignment success was defined as the proportion of samples that were correctly assigned back to their sampling origin with a probability of 50% or greater, the overall assignment success was greater than 70% once 1140 or more loci were analyzed and was maximized at 80% correctly

assigned when 2850 loci were analyzed (Fig. 4). When the threshold for assignment of an individual to a given stratum was raised to a probability of 70% or greater, the values remained similar, with up to 79% being assigned correctly (Table 4). Assignment success was somewhat higher for the Beaufort Sea stratum (reaching a high of 88% correct assignments) than for the eastern Bering Sea stratum (72% correct assignments) (Fig. 4). When all loci were included in the analysis, assignment success was 72 to 76% with the 50% probability

a

DA eigenvalues

b

DA eigenvalues

c

DA eigenvalues

- N Bering
- SE Bering
- Beaufort
- SE Chukchi
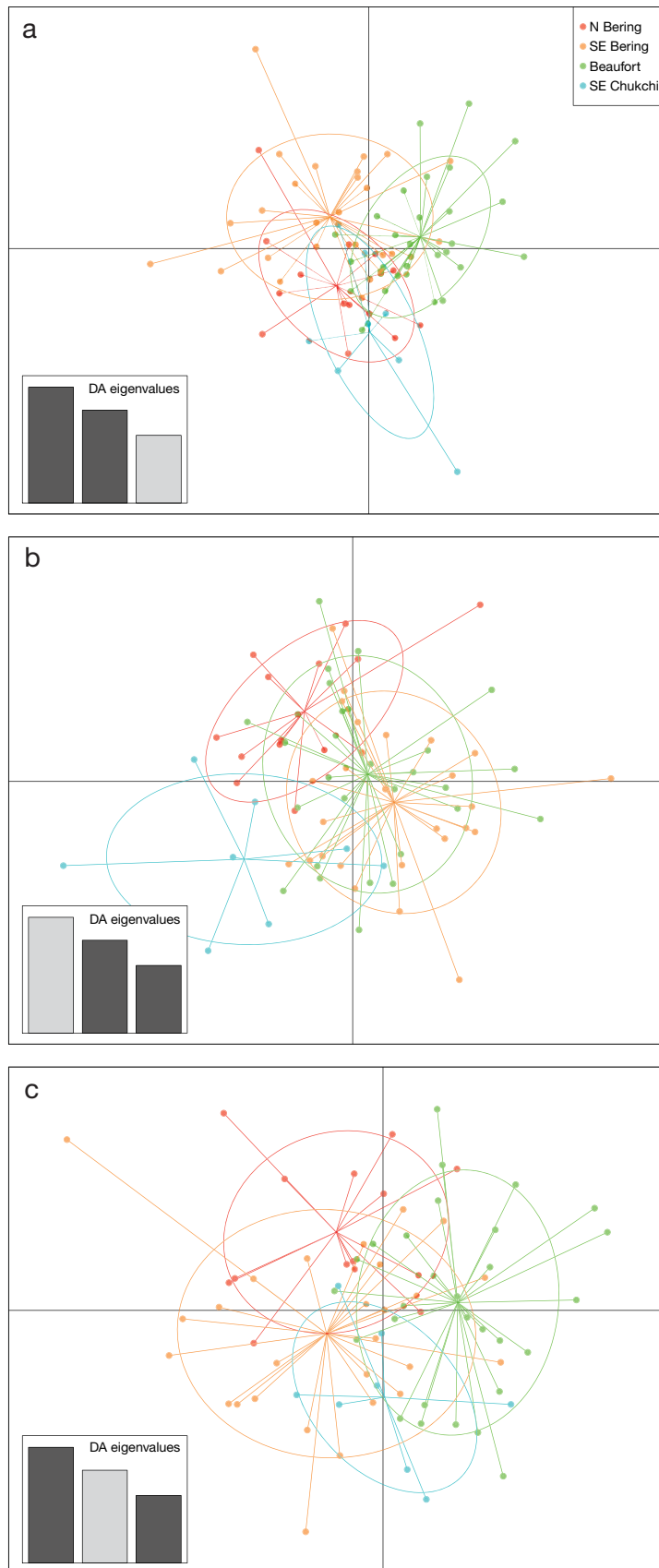
threshold and 71 to 75 % with the 70 % probability threshold.

When the assignment test was repeated using the simulated datasets in which individuals were randomly grouped in 1 of 2 strata, the maximum assignment success was 55.5 %, which is similar to what would be expected at random when assigning between 2 groups (Fig. 5).

4. DISCUSSION

Early studies of many commercially exploited marine fish and mollusc species found little to no genetic differentiation, which was typically attributed to a high degree of gene flow in marine environments and to the large effective sizes of many populations (Cano et al. 2008). In several cases, however, the advent of high-throughput sequencing and the associated ability to evaluate population structure at thousands (vs. tens) of loci has uncovered evidence of population structure (e.g. Benestan et al. 2015, Momigliano et al. 2017). Similarly, early studies of subspecies and stocks of spinner dolphins *Stenella longirostris* in the eastern tropical Pacific, which historically numbered in the low millions (Wade et al. 2007), found little evidence of genetic differences using traditional genetic markers (mtDNA and microsatellites), despite the presence of morphological differences (Dizon et al. 1994, Galver 2002). Using >3700 nuclear SNP loci identified through a genotyping by sequencing approach, however, Leslie & Morin (2016) were able to show genetic structure that supported the morphologically distinguished groups.

The issue with detecting population structure in taxa with high abundance is that the rate at which genetic differences evolve between groups is inversely proportional to

Fig. 3. Results of running a discriminant analysis (DA) of principal components on the single nucleotide polymorphism genotype data using sampling locality as a prior and retaining 22 principal components. Points represent individual genotypes, color coded by their original sampling locality. The inertial ellipses represent 67 % of the variance within each stratum: (a) first 2 discriminant factors, (b) second and third discriminant factors, and (c) first and third discriminant factors
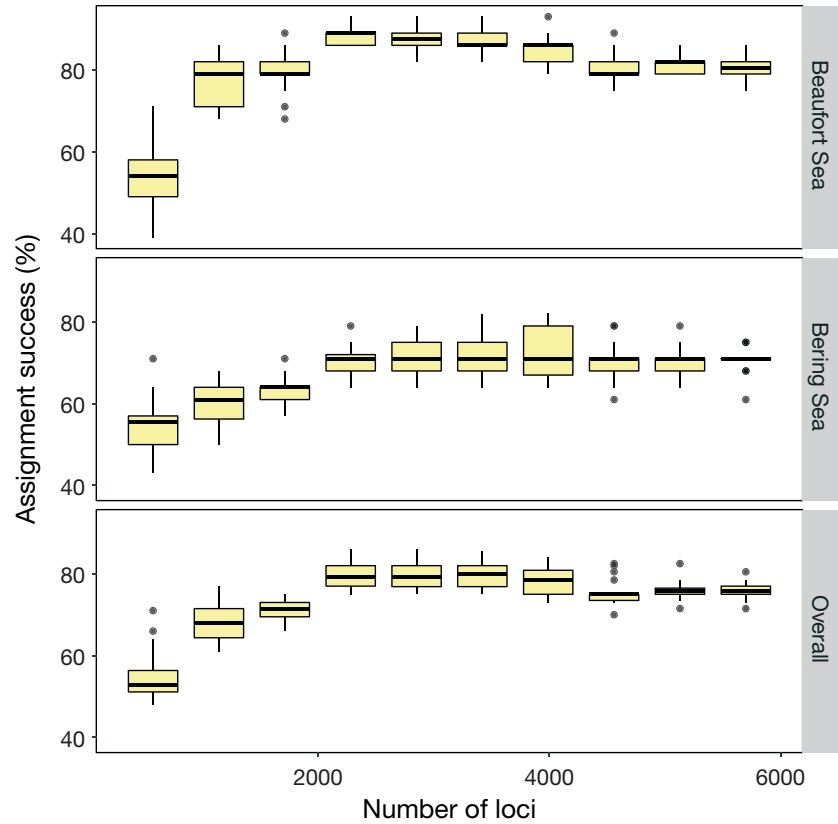
Fig. 4. Proportions of samples assigned correctly to their stratum of sampling origin (based on a 50% threshold) as the number of loci increased from 10% of all loci to the full dataset. For analysis based on only a proportion of the loci, loci were selected based on their ranking as derived from computing F_{ST} values between strata. Missing data were imputed based on the random forest algorithm

effective population size. In such scenarios, we would expect to find small effect sizes (i.e. F_{ST}) even when gene flow is limited (Taylor & Dizon 1996). While previous genetic studies have failed to identify clear patterns of population structure among Arctic ringed seals (Davis et al. 2008, Martinez-Bakker et al. 2013), those studies used a relatively low number of genetic markers (≤11 microsatellite loci) to compare geographic strata. As such, some question remained as to whether those datasets were powerful enough to detect small, but potentially biologically relevant, genetic differences between regions in this abundant subspecies.

To address this uncertainty, we increased the number of nuclear loci genotyped by several orders of magnitude (to ~5700 SNPs in 79 individuals). Low levels of divergence ($F_{ST} < 0.001$) between regions were detected in the SNP analysis, and only the comparison of the 2 most distant strata (the eastern Bering Sea and the Beaufort Sea

strata) was statistically significant. No loci were identified as being putatively under selection in our

Table 3. Pairwise comparisons between strata with all samples included and after subdividing by sex. n: sample size. **Bold**: significant values (p < 0.05)

Comparison	F_{ST}	F_{ST} p
All samples		
E Bering Sea (n = 28) vs. N Bering Sea (n = 14)	0.000	0.490
E Bering Sea (n = 28) vs. SE Chukchi Sea (n = 8)	0.001	0.154
E Bering Sea (n = 28) vs. Beaufort Sea (n = 29)	**0.001**	**0.005**
N Bering Sea (n = 14) vs. SE Chukchi Sea (n = 8)	0.001	0.263
N Bering Sea (n = 14) vs. Beaufort Sea (n = 29)	0.000	0.316
SE Chukchi Sea (n = 8) vs. Beaufort Sea (n = 29)	0.001	0.207
Females only		
E Bering Sea (n = 12) vs. N. Bering Sea (n = 8)	0.002	0.204
E Bering Sea (n = 12) vs. SE Chukchi Sea (n = 5)	0.004	0.096
E Bering Sea (n = 12) vs. Beaufort Sea (n = 11)	**0.003**	**0.016**
N Bering Sea (n = 8) vs. SE Chukchi Sea (n = 5)	−0.001	0.693
N Bering Sea (n = 8) vs. Beaufort Sea (n = 11)	0.000	0.452
SE Chukchi Sea (n = 5) vs. Beaufort Sea (n = 11)	0.001	0.263
Males only		
E Bering Sea (n = 16) vs. N. Bering Sea (n = 6)	−0.002	0.877
E Bering Sea (n = 16) vs. SE Chukchi Sea (n = 3)	0.001	0.356
E Bering Sea (n = 16) vs. Beaufort Sea (n = 18)	**0.002**	**0.016**
N Bering Sea (n = 6) vs. SE Chukchi Sea (n = 3)	0.006	0.108
N Bering Sea (n = 6) vs. Beaufort Sea (n = 18)	0.000	0.522
SE Chukchi Sea (n = 3) vs. Beaufort Sea (n = 18)	0.002	0.271

Table 4. Results of running gsi_sim in the R package assigner. Values are averaged after 20 iterations. Due to limited sample sizes representing the northern Bering Sea and southeastern Chukchi Sea strata, only the eastern Bering Sea and the Beaufort Sea strata were analyzed here. For the upper half of the table, a sample is considered to be assigned correctly if it had a probability of assignment to its stratum of origin >50%. For the bottom half of the table, a sample was considered to be assigned correctly if it had a probability of assignment to its stratum of origin >70%

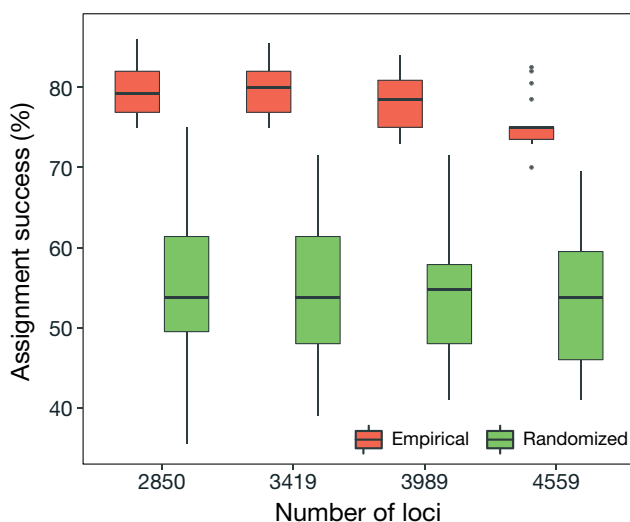| No. of loci	Beaufort Sea		Eastern Bering Sea		Overall
	No. of samples assigned correctly	Proportion of samples assigned correctly (%)	No. of samples assigned correctly	Proportion of samples assigned correctly (%)	Proportion of samples assigned correctly (%)
Stratum of origin >50%					
570	304	54	310	55	55
1140	430	77	336	60	68
1710	443	79	352	63	71
2280	494	88	396	71	79
2850	491	88	402	72	80
3419	486	87	404	72	79
3989	476	85	401	72	78
4559	448	80	394	70	75
5129	456	81	394	70	76
5699	453	81	397	71	76
Stratum of origin >70%					
570	288	51	299	53	52
1140	423	76	321	57	66
1710	439	78	345	62	70
2280	490	88	390	70	79
2850	485	87	394	70	78
3419	486	87	400	71	79
3989	471	84	396	71	77
4559	443	79	386	69	74
5129	451	81	388	69	75
5699	447	80	395	71	75

Fig. 5. Assignment success (proportions of samples that assigned to their stratum of origin at 50% or greater) in the empirical data (red) versus those generated from analysis of datasets in which individuals are randomly assigned to 1 of 2 strata of origin (green)

analysis, although additional exploration of the dataset could be valuable, as the approach we used is known to be conservative (Whitlock & Lotterhos 2015). The lack of evidence for local adaptation, together with the low magnitude of genetic differentiation identified between strata, suggests that gene flow between regions has been sufficient to avoid the loss of essential genetic variability needed to maintain the evolutionary potential of the subspecies. It is important to note, however, that large portions of the Arctic ringed seal's range were not analyzed (e.g. the Canadian Arctic and the Okhotsk Sea); thus, additional structure of evolutionary significance may have been identified with broader geographic sampling.

It is less clear how the low, and in 1 comparison statistically significant, levels of observed divergence relate to demographic connectivity, which depends on the relative contribution of dispersal to population dynamics and acts on ecological, rather than evolutionary, time scales (Waples & Gaggiotti 2006, Lowe & Allendorf 2010). The tipping point where populations become demographically independent usually occurs when the underlying genetic signal is weak, and in highly abundant species, genetic differentiation can persist despite a relatively low migration rate (Waples & Gaggiotti 2006). To provide some context in which our estimates of genetic divergence can be interpreted, we calculated the expected level of nuclear differentiation between areas when dispersal is low (1% per year). The empirical estimates of F_{ST} based on the SNP data ($F_{ST} = 0.001$ for all but 2 of the pairwise comparisons) are approximately 1 to 2 orders of magnitude higher than this expected effect size. Clearly, these estimates of differentiation rely on a number of assumptions, in terms of both the life history values used as well as the assumptions inherent in Wright's model (see Whitlock & McCauley 1999 for critique). However, while the actual rate could differ substantially, this exercise suggests that our results are consistent with what would be expected if some breeding sites used by Arctic ringed seals are connected by dispersal low enough to result in demographic independence.

Attempting to put the small but significant difference detected between the eastern Bering Sea and the Beaufort Sea in the pairwise comparisons of SNP data into the context of some of our other findings is challenging given the limitations of many genetic analytical approaches to discriminate between populations when divergence is low. For example, simulation testing of clustering methods, such as STRUCTURE (Pritchard et al. 2000, Evanno et al. 2005), has shown that such methods may have trouble accurately identifying the number of genetic clusters, and assigning individuals to those clusters, at low levels of genetic differentiation ($F_{ST} < 0.01$) (Latch et al. 2006, Waples & Gaggiotti 2006, Kalinowski 2011), an issue which may persist even when thousands of loci are analyzed (Benestan et al. 2015). The DAPC, which maximizes between-group variability while minimizing that within groups, has been able to at least partially separate out weakly differentiated stocks in some scenarios (e.g. Benestan et al. 2015, Leslie & Morin 2016, Dussex et al. 2018). When used with the ringed seal data, however, the DAPC was not able to identify genetic clusters in the absence of *a priori* information on geographic location of sampling and revealed substantial overlap of clusters even when geographic stratifications were incorporated.

Some support for the general pattern of population structure detected in our pairwise comparisons can be derived from the assignment test results. Assignment tests have been shown to lack power when differentiation is low, although their performance also depends on sample sizes and the number and variability of loci genotyped (Paetkau & Strobeck 1994, Berry et al. 2004, Paetkau et al. 2004, Hall et al. 2009, Lowe & Allendorf 2010). Here, we found that for the 2 most well-sampled groups, individuals could be assigned back to their stratum of origin with moderately high success (>70%) using the ~1100 loci that demonstrated the greatest differences between groups, with maximum assignment success (72% of the northern Bering Sea and 88% of the Beaufort Sea) being achieved when between 2280 and 3419 loci were used. While these patterns in assignment success suggest that a relatively large number of loci are needed to have a moderately high probability of correctly assigning samples to groups, the results indicate that some features (here, allele frequencies) differ between at least the 2 most well-sampled strata, which would be expected if population structure exists between those groups. Benestan et al. (2015) used similar methodology to conduct a LOO assignment test on American lobsters, which, like Arctic ringed seals, have high abundance and weak but significant levels of genetic differentiation ($F_{ST} \sim 0.001$). Similar to our results, they found that assignment success peaked at around 3000 loci. They also found that average assignment success increased with the number of individuals sampled, with mean assignment success remaining <75% when fewer than 30 individuals were sampled from each group.

4.1. Integrating other lines of evidence

In some cases where the genetic signal is small and/or its interpretation ambiguous, evaluating the genetic results in the context of other lines of evidence has proven informative. Within our study area, however, there are limited data from additional lines of evidence with which to corroborate the relevance of our genetic findings. The observations of ringed seals demonstrating fidelity to breeding sites over multiple years are congruent with what would be expected if natal fidelity is occurring, although the lack of known returns of seals to breed in regions where they were first identified as young of the year limits the inferences about population structure that can be drawn from the tagging data. In addition, the finding of a pair of putative paternal half siblings that were sampled in the same year at the same site is intriguing. Given the size of the population and that males are thought to maintain breeding territories throughout a given breeding season, it seems improbable that 2 seals sired by the same male in the same season would be sampled so close in proximity the following season if they were choosing wintering sites at random. While this finding suggests that at least some male yearling seals show an affinity for remaining in or returning to their natal site during their first winter post-weaning, ringed seals do not begin breeding until they are at least 6 or 7 yr old, and tagging data have shown that subadult seals overwinter near the ice edge, while breeding adults winter farther north in heavy and shorefast ice (Crawford et al. 2012). Thus, evidence that seals may show fidelity to their natal sites during their first year is not necessarily indicative of seals returning to their natal sites as breeding adults.

Some evidence that population structure may exist within the range of the Arctic ringed seal comes from recent studies identifying morphometric and life history differences among seals inhabiting the Canadian Arctic (Ferguson et al. 2018, 2019). Seals from sites located in the northern portion of the area studied (the Canadian High Arctic) were larger (both in

total length and girth) than their counterparts in the more southern waters of Hudson Bay. The northern seals grew more slowly, reaching asymptotic size 5 to 7 yr later than the southern seals; had lower fecundity and longer life spans; and exhibited sexual dimorphism that was not present among the southern seals. Within this area, latitudinal differences in diet and foraging behavior have also been observed (Yurkowski et al. 2016), and Ferguson et al. (2018, 2019) suggest that the divergent life history patterns that have emerged between the northern and southern seals are a response to differences in local environmental conditions. Of note, however, is that changes in growth rates and the average age of maturity of females have been observed to occur within areas over short time scales (between 1975–1984 and 2003–2012) in ringed seals in Alaska due to environmental conditions (Crawford et al. 2015), indicating there is flexibility within the species for both morphology and maturity.

4.2. Study design considerations

All samples analyzed in this study were collected opportunistically, either from harvested seals or research programs focused on other objectives (i.e. not on evaluating population structure). We only used samples that were collected during the breeding season to focus on the period when genetic differences by strata, if present, would be most pronounced in the dataset. This restriction, however, reduced the number of samples available for the study. Given these limitations, we increased power to detect demographically important genetic differentiation by increasing the number of loci genotyped, which, at least for microsatellite loci, has been shown to improve statistical power more than a larger sample size (Hale et al. 2012, Landguth et al. 2012). Note, however, that while the magnitude of differentiation (F_{ST} = 0.001) was similar among all comparisons except for 2 of those involving the northern Bering Sea, significant differentiation in the pairwise comparisons only occurred among the 2 strata with the highest (and roughly similar) sample sizes. In addition, comparisons based on random subsamples of individuals from those 2 strata did not consistently detect significant differentiation until the sample sizes reached 20 or higher. While the failure to detect significant differences in most of our comparisons could be a reflection of insufficient sampling, simulation-based studies have shown that with large numbers of SNPs (>1500 loci), accurate estimates of F_{ST}

can be derived with as few as 2 samples (Nazareno et al. 2017).

Simulations have shown that in addition to sample size, sampling scheme can also affect inference of population structure (Schwartz & McKelvey 2009, Koen et al. 2013, Oyler-McCance et al. 2013, Landguth & Schwartz 2014). While the strata analyzed were represented by at least 8 samples genotyped at ~5700 loci, the sampling was uneven and patchy, and large portions of the range of the Arctic subspecies were not represented. Sample collection efforts occurred over a 12 yr period, during which time the sea ice conditions fluctuated. These fluctuations may have affected the distribution and behavior of the seals, but the limited number of samples collected from any area in a given year was too small to evaluate potential temporal changes in genetic diversity or structure. Most of the samples representing the 2 northernmost strata were collected from seals killed by polar bears and were broadly distributed throughout a larger area, while all of the Bering Sea samples were from harvested seals and were collected in a few communities (Hooper Bay in the eastern Bering Sea stratum and Savoonga and Gambell, which are about 60 km apart, on Saint Lawrence Island in the northern Bering Sea stratum). If adult seals show site fidelity to breeding areas, irrespective of whether such areas represent their natal sites, then more spatially concentrated sampling of the 2 southern strata relative to the northern strata could have increased the probability of sampling related individuals (i.e. half siblings). However, as long as each stratum is sampled randomly, the inclusion of close relatives by chance should not introduce a bias into our sample sets. While the 1 putative paternal half-sibling pair that was detected in our dataset was found in the eastern Bering Sea stratum, the individuals were sampled in different months of the same year (i.e. they were not sampled together). Repeating the pairwise comparisons after removing 1 of the half siblings did not change the results substantively (Table S3).

Our results, although robust, highlight the need for dedicated sampling efforts to (1) increase the sample sizes representing different areas, (2) fill in geographic gaps between sampling sites, and (3) broaden the temporal scale of sampling to evaluate the stability of the genetic signal.

5. SUMMARY AND FUTURE WORK

As the Arctic continues to warm, ringed seals face a variety of potential threats associated with declin-

ing habitat quality. Given that the magnitude of these threats to ringed seals is unlikely to be uniform across their range, understanding connectivity between areas, and the mechanisms underlying it, is important in predicting the overall risk to the subspecies. For example, if connectivity is low and fidelity to breeding areas is largely driven by the return to natal areas, the continued return of seals to reproduce and breed in areas with low snow cover and/or unstable ice would likely lead to low pup survival due to the collapse of subnivean lairs and the subsequent exposure of pups to cold, increased predation risk, and potential early separation from their mothers. Under this scenario, eventually the sites most negatively impacted by climate warming would no longer be used by seals for pupping and breeding in the spring, although seals might continue to use the area for foraging during the non-breeding season. However, if connectivity between breeding sites is high and selection of sites is driven by prey distribution, it is plausible that seals could choose to reproduce and breed in areas with high prey availability but poor ice stability and/or low snow cover. In such a case, overall declines in abundance associated with reduced pup survival might occur, but the breeding sites would continue to be used, acting as a population sink.

Our results, as well as those of previous studies (Davis et al. 2008, Martinez-Bakker et al. 2013), indicate that interchange between ringed seal breeding areas likely occurs. This connectivity, if it continues, may protect the seals from the loss of evolutionary potential and adaptive capacity as they navigate habitat alterations and loss in the face of environmental warming. However, the magnitude of nuclear differentiation we observed between the southernmost (eastern Bering Sea) and northernmost (Beaufort Sea) strata, while small, is markedly higher than that estimated assuming 1% dispersal per year, and the statistically significant differences between these areas suggest a lack of panmixia across our study area. In combination with the relatively high rate at which samples from these 2 strata could be genetically assigned to their region of origin, these results suggest that subtle, but demographically important, structure is present among breeding areas used by Arctic ringed seals in the Bering, Chukchi, and Beaufort seas, consistent with most seals returning to their natal sites to breed as adults. If such return continues despite reduced habitat quality at breeding sites most negatively affected by environmental warming, declines in the number of seals breeding in those areas could occur.

Our results suggest that a comprehensive analysis of samples collected during the breeding season and throughout the range of the Arctic subspecies, in combination with the large-scale genotyping approach used here, is needed to more fully assess demographic structure across the range and to thereby better understand the impacts of future environmental changes on ringed seals. One advantage of the approach used here is that it provides a resource from which SNP genotyping assays could be designed that would facilitate expanding the spatial as well as temporal scope of the current study. In particular, these data could be mined to identify a panel of the most informative SNPs that could be used with low-quality and/or historic samples, which might greatly expand the number of samples. Another genetic approach that has proven useful for characterizing population structure in weakly differentiated populations is kinship analysis (Kane & King 2009, Saenz-Agudelo et al. 2009, Palsbøll et al. 2010). Such an approach is challenging in large populations (Hellberg 2009) given the number of samples that need to be collected and genotyped. However, concentrated sampling of seals in a portion of the range that is continued over multiple breeding seasons could provide insight into whether natal fidelity occurs, in which case, with sufficient sampling one would expect to find individuals of breeding age that represent parent–offspring pairs.

LITERATURE CITED

Anderson EC (2010) Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. Mol Ecol Res 10:701–710

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program Structure. Mol Ecol Res 8:1219–1229

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

Archer FI, Adams PE, Schneiders BB (2017) STRATAG: an R package for manipulating, summarizing and analysing population genetic data. Mol Ecol Res 17:5–11

Atkinson S (1997) Reproductive biology of seals. Rev Reprod 2:175–194

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

Benestan L, Gosselin T, Perrier C, Sainte-Marie B, Rochette R, Bernatchez L (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). Mol Ecol 24:3299–3315

Benestan L, Quinn BK, Maaroufi H, Laporte M and others (2016) Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). Mol Ecol 25:5073–5092

Bengtson JL, Hiruki-Raring LM, Simpkins MA, Boveng PL (2005) Ringed and bearded seal densities in the eastern Chukchi Sea, 1999–2000. Polar Biol 28:833–845

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? Mol Ecol 13:551–561

Born EW, Teilmann J, Acquarone M, Riget FF (2004) Habitat use of ringed seals (*Phoca hispida*) in the North Water area (north Baffin Bay). Arctic 57:129–142

Cano JM, Shikano T, Kuparinen A, Merilä J (2008) Genetic differentiation, effective population size and gene flow in marine fishes: implications for stock management. J Integr Field Sci 5:1–10

Conn PB, Ver Hoef JM, McClintock BT, Moreland EE and others (2014) Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. Methods Ecol Evol 5:1280–1293

Corander J, Majander KK, Cheng L, Merilä J (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. Mol Ecol 22:2931–2940

Crawford JA, Frost KJ, Quakenbush LT, Whiting A (2012) Different habitat use strategies by subadult and adult ringed seals (*Phoca hispida*) in the Bering and Chukchi seas. Polar Biol 35:241–255

Crawford JA, Quakenbush LT, Citta JJ (2015) A comparison of ringed and bearded seal diet, condition and productivity between historical (1975–1984) and recent (2003–2012) periods in the Alaskan Bering and Chukchi seas. Prog Oceanogr 136:133–150

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:409–510

Davis CS, Stirling I, Strobeck C, Coltman DW (2008) Population structure of ice-breeding seals. Mol Ecol 17:3078–3094

Descamps S, Aars J, Fuglei E, Kovacs KM and others (2017) Climate change impacts on wildlife in a High Arctic archipelago–Svalbard, Norway. Glob Change Biol 23:490–502

Dizon AE, Perrin WF, Akin PA (1994) Stocks of dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern tropical Pacific: a phylogeographic classification. NOAA Tech Rep NMFS 119

Dussex N, Taylor HR, Stovall WR, Rutherford K, Dodds KG, Clarke SM, Gemmell N (2018) Reduced representation sequencing detects only subtle regional structure in a heavily exploited and rapidly recolonizing marine mammal species. Ecol Evol 8:8736–8749

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Res 4:359–361

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLOS ONE 6:e19379

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Ferguson SH, Stirling I, McLoughlin P (2005) Climate change and ringed seal (*Phoca hispida*) recruitment in western Hudson Bay. Mar Mamm Sci 21:121–135

Ferguson SH, Young BG, Yurkowski DJ, Anderson R, Willing C, Nielsen O (2017) Demographic, ecological, and physiological responses of ringed seals to an abrupt decline in sea ice availability. PeerJ 5:e295

Ferguson SH, Zhu X, Young BG, Yurkowski DJ, Thiemann GW, Fisk AT, Muir DC (2018) Geographic variation in ringed seal (*Pusa hispida*) growth rate and body size. Can J Zool 96:649–659

Ferguson SH, Yurkowski DJ, Young BG, Willing C, Zhu X, Muir DCG, Fisk AT, Thiemann GW (2019) Do intraspecific life history patterns follow interspecific predictions? A test using latitudinal variation in ringed seals. Pop Ecol 61:371–382

Fogelqvist J, Niittyvuopio A, Ågren J, Savolainen O, Lascoux M (2010) Cryptic population genetic structure: the number of inferred clusters depends on sample size. Mol Ecol Resour 10:314–323

Francis RM (2017) POPHELPER: an R package and web app to analyse and visualize population structure. Mol Ecol Resour 17:27–32

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. Genet Res 66:95–107

Freitas C, Kovacs KM, Ims RA, Fedak MA, Lydersen C (2008) Ringed seal post-moulting movement tactics and habitat selection. Oecologia 155:193–204

Frey KE, Maslanik JA, Kinney JC, Maslowski W (2014) Recent variability in sea ice cover, age, and thickness in the Pacific Arctic region. In: Grebmeier J, Maslowski W (eds) The Pacific Arctic region. Springer, Dordrecht

Frey KE, Moore GWK, Cooper LW, Grebmeier JM (2015) Divergent patterns of recent sea ice cover across the

Bering, Chukchi, and Beaufort seas of the Pacific Arctic region. Prog Oceanogr 136:32–49

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. Trends Ecol Evol 27:489–496

Funk WC, Forester BR, Converse SJ, Darst C, Morey S (2018) Improving conservation policy with genomics: a guide to integrating adaptive potential into US Endangered Species Act decisions for conservation practitioners and geneticists. Conserv Genet 20:115–134

Furgal CM, Kovacs KM, Innes S (1996) Characteristics of ringed seal, *Phoca hispida*, subnivean structures and breeding habitat and their effects on predation. Can J Zool 74:858–874

Galver LM (2002) The molecular ecology of spinner dolphins, *Stenella longirostris*: genetic diversity and population structure. PhD thesis, University of California San Diego, La Jolla, CA

Gjertz I, Lydersen C (1986) Polar bear predation on ringed seals in the fast-ice of Hornsund, Svalbard. Polar Res 4: 65–68

Gleason LU, Burton RS (2016) Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*. Mol Ecol 25:3557–3573

Gosselin T, Anderson EC, Bradbury IR (2016) assigner: assignment analysis with GBS/RAD data using R. R package version 0.4.1. https://github.com/thierrygosselin/assigner

Gruber B, Unmack PJ, Berry OF, Georges A (2018) dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol Ecol Resour 18:691–699

Hale ML, Burg TM, Steeves TE (2012) Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLOS ONE 7:e45170

Hall LA, Palsbøll PJ, Beissinger SR, Harvey JT and others (2009) Characterizing dispersal patterns in a threatened seabird with limited genetic structure. Mol Ecol 18: 5074–5085

Hammill MO, Smith TG (1991) The role of predation in the ecology of the ringed seal in Barrow Strait, Northwest Territories, Canada. Mar Mamm Sci 7:123–135

Harwood LA, Smith TG, Auld JC (2012a) Fall migration of ringed seals (*Phoca hispida*) through the Beaufort and Chukchi Seas, 2001–02. Arctic 65:35–44

Harwood LA, Smith TG, Melling H, Alikamik J, Kingsley MCS (2012b) Ringed seals and sea ice in Canada's western Arctic: harvest-based monitoring 1992–2011. Arctic 65:377–390

Harwood LA, Smith TG, George JC, Sandstrom SJ, Walkusz W, Divoky GJ (2015a) Change in the Beaufort Sea ecosystem: diverging trends in body condition and/or production in five marine vertebrate species. Prog Oceanogr 136:263–273

Harwood LA, Smith TG, Auld JC, Melling H, Yurkowski DJ (2015b) Seasonal movements and diving of ringed seals, *Pusa hispida*, in the western Canadian Arctic, 1999–2001 and 2010–11. Arctic 68:193–209

Heide-Jørgensen MP, Stewart BS, Leatherwood S (1992) Satellite tracking of ringed seals *Phoca hispida* off northwest Greenland. Ecography 15:56–61

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Annu Rev Ecol Evol Syst 40: 291–310

Iacozza J, Ferguson SH (2014) Spatio-temporal variability of snow over sea ice in western Hudson Bay, with reference to ringed seal pup survival. Polar Biol 37:817–832

IPCC (2013) Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:e25

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using *adegenet* 2.0.0. MRC Centre for Outbreak Analysis and Modelling, Imperial College London

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94

Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106:625–632

Kane NC, King MG (2009) Using parentage analysis to examine gene flow and spatial genetic structure. Mol Ecol 18: 1551–1552

Kelly BP, Badajos OH, Kunnasranta M, Moran JR, Martinez-Bakker M, Wartzok D, Boveng P (2010a) Seasonal home ranges and fidelity to breeding sites among ringed seals. Polar Biol 33:1095–1109

Kelly BP, Bengtson JL, Boveng P, Cameron M and others (2010b) Status review of the ringed seal (*Phoca hispida*). US Dep Commer, NOAA Tech Memo NMFS-AFSC-212

Kilian A, Wenzl P, Huttner E, Carling J and others (2012) Diversity arrays technology: a generic genome profiling technology on open platforms. In: Pompanon F, Bonin A (eds) Data production and analysis in population genomics. Springer, New York, NY

Koen EL, Bowman J, Garroway CJ, Wilson PJ (2013) The sensitivity of genetic connectivity measures to unsampled and under-sampled sites. PLOS ONE 8:e56204

Kovacs KM, Lydersen C, Overland JE, Moore SE (2011) Impacts of changing sea-ice conditions on Arctic marine mammals. Mar Biodiv 41:181–194

Laidre KL, Stern H, Kovacs KM, Lowry L and others (2015) Arctic marine mammal population status, sea ice habitat loss, and conservation recommendations for the 21st century. Conserv Biol 29:724–737

Landguth EL, Schwartz MK (2014) Evaluating sample allocation and effort in detecting population differentiation for discrete and continuously distributed individuals. Conserv Genet 15:981–992

Landguth EL, Fedy BC, Oyler-McCance SJ, Garey AL and others (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Mol Ecol Resour 12:276–284

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE Jr (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv Genet 7:295–302

Leslie MS, Morin PA (2016) Using genome-wide SNPs to detect structure in high-diversity and low-divergence populations of severely impacted eastern tropical Pacific spinner (*Stenella longirostris*) and pantropical spotted dolphins (*S. attenuata*). Front Mar Sci 3:253

Li H (2014) Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 30:2843–2851

Lindsay R, Schweiger A (2015) Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations. Cryosphere 9:269–283

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? Mol Ecol 19:3038–3051

Luque SP, Ferguson SH, Breed GA (2014) Spatial behaviour of a keystone Arctic marine predator and implications of climate warming in Hudson Bay. J Exp Mar Biol Ecol 461:504–515

Lydersen C, Gjertz I (1987) Population parameters of ringed seals (*Phoca hispida* Schreber, 1775) in the Svalbard area. Can J Zool 65:1021–1027

Lydersen C, Smith TG (1989) Avian predation on ringed seal *Phoca hispida* pups. Polar Biol 9:489–490

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152:1753–1766

Martien KK, Baird RW, Hedrick NM, Gorgone AM and others (2012) Population structure of island-associated dolphins: evidence from mitochondrial and microsatellite markers for common bottlenose dolphins (*Tursiops truncatus*) around the main Hawaiian Islands. Mar Mamm Sci 28:E208–E232

Martinez-Bakker ME, Sell SK, Swanson BJ, Kelly BP, Tallmon DA (2013) Combined genetic and telemetry data reveal high rates of gene flow, migration, and long-distance dispersal potential in Arctic ringed seals (*Pusa hispida*). PLOS ONE 8:e77125

McLaren IA (1958) The biology of the ringed seal (*Phoca hispida* Schreber) in the eastern Canadian Arctic. Bull Fish Res Board Can 118:1–97

Momigliano P, Harcourt R, Robbins WD, Jaiteh V, Mahardika GN, Sembiring A, Stow A (2017) Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). Heredity 119:142–153

Moore SE, Huntington HP (2008) Arctic marine mammals and climate change: impacts and resilience. Ecol Appl 18:S157–S165

Nazareno AG, Bemmels JB, Dick CW, Lohmann LG (2017) Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. Mol Ecol Res 17:1136–1147

Oyler-McCance SJ, Fedy BC, Landguth EL (2013) Sample design effects in landscape genetics. Conserv Genet 14:275–285

Pacifici M, Santini L, Di Marco M, Baisero D and others (2013) Generation length for mammals. Nat Conserv 5:87–94

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. Mol Ecol 3:489–495

Paetkau D, Slade R, Burdens M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol 13:55–65

Palsbøll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16

Palsbøll PJ, Peery MZ, Berube M (2010) Detecting populations in the 'ambiguous' zone: kinship-based estimation of population structure at low genetic divergence. Mol Ecol Resour 10:797–805

Peng G, Meier WN (2017) Temporal and regional variability of Arctic sea-ice coverage from satellite data. Ann Glaciol 59:191–200

Perneger TV (1998) What's wrong with Bonferroni adjustments. Br Med J 316:1236–1238

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model organisms. PLOS ONE 7:e37135

Pew J, Muir PH, Wang J, Frasier TR (2015) related: an R package for analysing pairwise relatedness from codominant molecular markers. Mol Ecol Resour 15:557–561

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pritchard JK, Wen X, Falush D (2010) Documentation for *structure* software: Version 2.3. http://pritch.bsd.uchicago.edu/

Puechmaille SJ (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour 16:608–627

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. Mol Ecol 22:2953–2970

Robertson KM, Lauf ML, Morin PA (2018) Genetic sexing of pinnipeds: a real-time, single step qPCR technique. Conserv Genet Resour 10:213–218

Rodríguez-Ramilo ST, Wang J (2012) The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. Mol Ecol Resour 12:873–884

Saenz-Agudelo P, Jones GP, Thorrold SR, Planes S (2009) Estimating connectivity in marine populations: an empirical evaluation of assignment tests and parentage analysis under different gene flow scenarios. Mol Ecol 18:1765–1776

Sansaloni C, Petroli C, Jaccoud D, Carling J, Detering F, Grattapaglia D, Killian A (2011) Diversity arrays technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. BMC Proc 5:P54

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conserv Genet 10:441

Sigler MF, Renner M, Danielson SL, Eisner LB and others (2011) Fluxes, fins, and feathers: relationships among the Bering, Chukchi, and Beaufort seas in a time of climate change. Oceanography 24:250–265

Sipilä T (2016) *Pusa hispida* ssp. *saimensis*. The IUCN Red List of Threatened Species 2016:e.T41675A66991678. http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T41675A66991678.en (accessed 12 Feb 2018)

Smith TG (1976) Predation of ringed seal pups (*Phoca hispida*) by the arctic fox (*Alopex lagopus*). Can J Zool 54:1610–1616

Smith TG, Stirling I (1975) The breeding habitat of the ringed seal (*Phoca hispida*). The birth lair and associated structures. Can J Zool 53:1297–1305

Smith TG, Hammill MO, Taugbøl G (1991) A review of the developmental, behavioural and physiological adaptations of the ringed seal, *Phoca hispida*, to life in the Arctic winter. Arctic 44:124-131

Stenson G, Hammill M (2014) Can ice breeding seals adapt to habitat loss in a time of climate change? ICES J Mar Sci 71:1977−1986

Stirling I (2005) Reproductive rates of ringed seals and survival of pups in northwestern Hudson Bay, Canada, 1991−2000. Polar Biol 28:381−387

Stroeve JC, Serreze MC, Holland MM, Kay JE, Malanik J, Barrett AP (2012) The Arctic's rapidly shrinking sea ice cover: a research synthesis. Clim Change 110: 1005−1027

Taylor BL (1997) Defining 'population' to meet management objectives for marine mammals. In: Dizon AE, Chivers SJ, Perrin WF (eds) Molecular genetics of marine mammals: incorporating the proceedings of a workshop on analysis of genetic data to address problems of stock identity as related to management of marine mammals. Soc Mar Mammal Spec Publ 3, Allen Press, Lawrence, KS, p 49−65

Taylor BL, Dizon AE (1996) The need to estimate power to link genetics and demography for conservation. Conserv Biol 10:661−664

Teilmann J, Born EW, Acquarone M (2000) Behaviour of ringed seals tagged with satellite transmitters in the North Water polynya during fast-ice formation. Can J Zool 77:1934−1946

Valtonen M, Palo JU, Ruokonen M, Kunnasranta M, Nyman T (2012) Spatial and temporal variation in genetic diversity of an endangered freshwater seal. Conserv Genet 13:1231−1245

Valtonen M, Palo JU, Aspi J, Ruokonen M, Kunnasranta M, Nyman T (2014) Causes and consequences of fine-scale population structure in a critically endangered freshwater seal. BMC Ecol 14:22

Wade PR, Watters GM, Gerrodette T, Reilly SB (2007) Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. Mar Ecol Prog Ser 343:1−14

Wang M, Overland JE (2015) Projected future duration of the sea-ice-free season in the Alaskan Arctic. Prog Oceanogr 136:50−59

Wang M, Yang Q, Overland JE, Stabeno P (2017) Sea-ice cover timing in the Pacific Arctic: the present and projections to mid-century by selected CMIP5 models. Deep Sea Res II 152:22−34

Waples RS, Gaggiotti OE (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419−1439

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:S24−S36

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. Heredity 82:117−125

Willing EM, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. PLOS ONE 7:e42649

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395−420

Yurkowski DJ, Ferguson S, Choy ES, Loseto LL and others (2016) Latitudinal variation in ecological opportunity and intraspecific competition indicates differences in niche variability and diet specialization of Arctic marine predators. Ecol Evol 6:1666−1678