THE ANALYSIS OF POPULATION GENETIC STRUCTURE IN ALASKA HARBOR SEALS, *Phoca vitulina*, AS A FRAMEWORK FOR THE IDENTIFICATION OF MANAGEMENT STOCKS

by

Greg M. O’Corry-Crowe, Karen K. Martien, and Barbara L. Taylor

ADMINISTRATIVE REPORT LJ-03-08
"This report is used to ensure prompt dissemination of preliminary results, interim reports, and special studies to the scientific community. The material is not ready for formal publication since the paper may later be published in a modified form to include more recent information or research results. Abstracting, citing, or reproduction of this information is not allowed. Contact author if additional information is required."
THE ANALYSIS OF POPULATION GENETIC STRUCTURE IN ALASKAN HARBOR SEALS, PHOCA VITULINA, AS A FRAMEWORK FOR THE IDENTIFICATION OF MANAGEMENT STOCKS

Greg M. O’Corry-Crowe, Karen K. Martien, and Barbara L. Taylor
Southwest Fisheries Science Center
National Marine Fisheries Service, NOAA
8604 La Jolla Shores Dr.
La Jolla, CA 92037

AUGUST 2003

ADMINISTRATIVE REPORT LJ-03-08
# TABLE OF CONTENTS

EXECUTIVE SUMMARY .......................................................... v

1. INTRODUCTION .................................................................. 1
   1.1 Trends in abundance, and management concerns .................. 3
   1.2 Defining stocks .......................................................... 3
   1.3 Evidence of population structure in Alaskan harbor seals .......... 7
   1.4 Difficulty of resolving population structure in a continuously distributed species . . 7
   1.5 Population genetic structure and the definition of stocks of Alaskan harbor seals . . 8

2. MATERIALS AND METHODS ................................................ 9
   2.1 LABORATORY METHODS ............................................ 9
      2.1.1 Sample collection ............................................... 9
      2.1.2 DNA extraction ................................................ 9
      2.1.3 Mitochondrial DNA amplification and sequencing ............ 9
   2.2 DATA ANALYSIS ..................................................... 13
      2.2.1 Analysis of genetic diversity .................................. 13
      2.2.2 Analysis of genetic differentiation ............................. 13
         2.2.2.1 Definition of initial units ................................. 14
         2.2.2.2 Parameter estimation ..................................... 17
            Boundary Rank .................................................. 17
            UPGMA and Neighbor-joining ................................ 18
            Dispersal rate estimation ..................................... 18
         2.2.2.3 Hypothesis testing ......................................... 22
      2.2.3 Analysis of genetic diversity ................................ 13

3. RESULTS ........................................................................... 23
   3.1 Overall genetic diversity ............................................. 23
   3.2 Population subdivision ................................................. 23
      3.2.1 Parameter estimation .......................................... 23
         Boundary Rank .................................................... 23
         UPGMA and Neighbor-Joining .................................. 23
         Dispersal rate estimation ....................................... 27
         Effect of initial unit definition on estimates of population structure .... 27
   3.2.2 Hypothesis testing ............................................... 27

4. DISCUSSION ....................................................................... 31
   4.1 Population genetic structure in Alaskan harbor seals ............ 31
   4.2 Comparison of the various methods used to resolve population structure .......... 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>The influence of sample size and the configuration of initial strata on estimates of population</td>
<td>34</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of genetic findings to other information of relevance to population structure</td>
<td>35</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Movements</td>
<td>35</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Trends in abundance</td>
<td>37</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Foraging ecology</td>
<td>38</td>
</tr>
<tr>
<td>4.5</td>
<td>Identification of stocks</td>
<td>38</td>
</tr>
<tr>
<td>4.6</td>
<td>Future directions</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td><strong>ACKNOWLEDGMENTS</strong></td>
<td>40</td>
</tr>
<tr>
<td></td>
<td><strong>LITERATURE CITED</strong></td>
<td>41</td>
</tr>
<tr>
<td></td>
<td><strong>APPENDICES</strong></td>
<td>51</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. The distribution of harbor seals, Phoca vitulina, in Alaska. The boundaries of the three currently recognized management stocks are indicated by dashed lines .......................... 2

Figure 2. Distribution of pristine populations (a), versus potential distributions after 50% of total abundance is removed (b,c,d). The area of the oval represents abundance, height represents geographic distance ................................................................. 5

Figure 3. The distribution of 180 sampling sites of harbor seals in Alaska. 50km radius buffer zones are indicated around each site ................................................................. 10

Figure 4. The abundance and distribution of harbor seals hauled out at terrestrial and glacial ice sites during fall molt surveys in Alaska. = 1-50 seals, = 51-100 seals, = 101-200 seals, and = >200 seals. Data provided by J. Bengtson and D. Withrow, Polar Ecosystems Program, National Marine Mammal Laboratory, NMFS, Seattle, Washington. Data from the Pribilof Islands, Glacier Bay and Lake Iliamna are not present on this map .................................................. 15

Figure 5. The 31 initial units defined using information on movement patterns and distribution of harbor seals in Alaska. For details on each stratum see Table 1. The 15 units that were excluded from the final analysis of population genetic structure due to low adjusted sample size are highlighted .................................................. 16

Figure 6. The connectivity matrix used in the Boundary Rank analysis of population genetic structure. A: Allowed connections among the 15 strata that were used in the analysis of population structure. B: detail of the connectivity matrix for Southeast Alaska, C: detail of the connectivity matrix for the Gulf of Alaska. Sampling sites are indicated in red, abundance is indicated in green, strata are outlined in blue, connections are indicated by dashed lines, and the strata excluded based on limited samples are highlighted in jade ................................................................. 19-21

Figure 7. Dendrogram of the outcome of the BR analysis. The analysis began with 16 units, as shown by the legend at the bottom of the tree, and proceeded until all 16 were clustered into a single unit. At each step, the two most similar units were combined. The depth of the nodes in the tree indicate the step in the analysis at which the units connected by that node were coalesced ................................................................. 24

Figure 8. Dendrogram of the outcome of the UPGMA cluster analysis based on \( \chi^2/dof \) of the 16 initial units of harbor seals in Alaska .................................................. 25

Figure 9. Neighbor-joining (N-J) tree based on \( \chi^2/dof \) of the estimated relationships among the 16 initial units of harbor seals in Alaska .................................................. 26

Figure 10. Graphic representation of the stage in the Boundary Rank analysis when all units were significantly differentiated from their neighbors at \( \chi^2 \) permutation \( p \)-value of < 0.1 .................................................. 28
LIST OF TABLES

Table 1. Summary of the number of sampling sites, the number of samples (n), the proportion of individuals with unique haplotypes, haplotypic diversity (H) and the adjusted sample size (n_a) within the 31 initial harbor seal units in Alaska. Units with n<4 are highlighted ........................................ 11

Table 2. Sample size and the number of sampling locations in relation to harbor seal abundance. The Alaska range is divided into 12 large regions. Minimum abundances are based on raw counts of seals hauled out during aerial and land-based surveys .................................................. 12

Table 3. Number of seals moving per annum between units identified by Boundary Rank. Estimates were made by dividing the estimates generated by the program Migrate (Beerli and Felsenstein, 2001) by an average harbor seal generation time of 8 years. 95% likelihood intervals are given in parentheses .............................................. 29

Table 4. Population subdivision among 16 initial harbor seal strata with large adjusted sample size. The values are p values from permutation chi-square test. The dark shaded cells are comparisons where p<0.05, the light shaded cells are comparisons where 0.05<p<0.1, while the unshaded cells are comparisons where p>0.1 ......................................................... 30
EXECUTIVE SUMMARY

Background

In the 1995 Alaska Marine Mammal Stock Assessment Reports, the National Marine Fisheries Service (NMFS) defined three stocks of harbor seal (*Phoca vitulina*) in Alaska, based primarily on broad-scale geographic differences in trends in abundance. NMFS, however, recognized that considerable uncertainty about Alaskan harbor seal stock structure remained and in the fall of 1994 initiated genetic studies of harbor seal stock structure in Alaska. This report details the findings from these studies, which conclude that current evidence supports a minimum of 12 stocks of harbor seals in Alaska.

Harbor seals (*Phoca vitulina*) occupy a near-continuous distribution in the coastal and continental shelf waters of Alaska from Dixon Entrance in the southeast, west throughout the Gulf of Alaska and the Aleutian Archipelago to Kuskokwim Bay in the Bering Sea (Figure ES-1). This important marine predator occupies a diverse range of habitats, hauls out at thousands of discrete coastal sites and represents a significant marine resource to a range of users. Harbor seals have declined dramatically in some parts of their Alaska range over the past few decades while in other parts their numbers have increased or remained stable over similar time periods. These declines and differences in trend suggest areas with independent population dynamics, and therefore, highlight the need for the definition of biologically meaningful management units, also known as stocks. The spatial scale of these stocks is important for interpreting direct and indirect human-caused mortality in relation to abundance, population trend and other aspects of harbor seal biology.

Figure ES-1. The distribution of harbor seals, *Phoca vitulina*, in Alaska. The boundaries of the three currently recognized management stocks are indicated by dashed lines.
Management Objectives

The stated management goals of the Marine Mammal Protection Act (MMPA) are that population stocks should not be “permitted to diminish beyond the point at which they cease to be a significant functioning element in the ecosystem of which they are a part....and [that] they should not be permitted to diminish below their optimum sustainable population.” The former objective has been interpreted as maintaining the species range, while the latter as not allowing stocks to decline below 50% of their historical population size “keeping in mind the carrying capacity of the habitat”. Reduction of the range and local depletion are also undesirable because they could violate the goal of sustained use by Alaska Natives. Range contraction or fragmentation may mean that hunters that traditionally hunted in one area may have to travel farther and spend more time to obtain the same number of seals.

To attain the objective of maintaining the species range, the definition of stocks should be based on semi-isolated groups or sub-populations of seals that seldom exchange individuals. If that is not done, there is a risk of local depletion or extirpation. Therefore, an understanding of population structure and patterns of dispersal is central to identifying biologically meaningful stocks.

Genetic analysis of population structure and dispersal patterns

Genetic analysis is a long-established method of analyzing population structure and dispersal patterns and of defining units of conservation. Because of its maternal mode of inheritance and rapid rate of evolution, mitochondrial DNA (mtDNA) is an ideal marker in the investigation of the demographic relationships among groups of animals such as harbor seals. The analysis of variation in this marker can be used to identify management units where the primary objectives are to preserve the species range as well as maintain healthy populations. In this report we present our findings from an extensive genetic study of population subdivision and dispersal patterns of harbor seals in Alaska using mtDNA.

Investigating population structure and dispersal in harbor seals in Alaska presented a number of challenges, including how to assess population subdivision in a continuously distributed species. To meet these challenges:

1. We conducted extensive sampling across the species’ range. In an earlier study we sampled individuals from the major centers of harbor seal abundance. While that study revealed much about the evolutionary history of the species and documented broad-scale patterns of population structure, it could not fully uncover the scale and form of population partitioning because of major gaps in sample coverage. In the present study, we analyzed a total of 881 seals sampled at 180 separate locations throughout the range. These samples were provided by Alaska Native hunters, scientists and from tissue archives.

2. We used information on the distribution, abundance and movement behavior of harbor seals in Alaska to define a set of 31 initial units (= groupings of sampling sites) for comparison in the analysis of population structure (Figure ES-2). These initial units were defined so as to be small enough in area to minimize the risk of missing structure and yet large enough to minimize the effects of low sample size. However, sample size in some
of these units was determined to be too low to be representative of the underlying genetic composition and so these units were excluded from further analysis (Figure ES-2).

3. We employed a variety of methods to analyze the genetic data for evidence of population structure, including the geographically-constrained clustering method, Boundary Rank, classic distance-based cluster and phylogeny reconstruction analyses and statistical hypothesis-testing.

4. We estimated rates of annual dispersal of seals among areas using a maximum-likelihood approach based on coalescent theory.

5. We compared our genetic results to recent findings from studies on harbor seal movement and foraging behavior, trends in abundance, and diet in order to better resolve the spatial pattern of population structure of this species in Alaska.

The analysis of mtDNA variation in 881 harbor seals sampled from 180 sites throughout Alaska revealed:

1. Substantial population subdivision across the State. The pattern of genetic differentiation was correlated with geographic distance indicating that female dispersal distances are a fraction of the species range and that when harbor seals do disperse, it is primarily to neighboring areas.

2. The form of genetic differentiation, in conjunction with the non-uniform abundance of seals across their range, indicates that Alaskan harbor seals are subdivided on spatial scales of 150-540km, depending on region, into a series of partially isolated sub-populations.

3. Using a variety of clustering methods, we identified 12 clusters of sampling sites that differed from each other at p<0.1 using a chi-square permutation test (Figure ES-3).

4. We estimated that dispersal among neighboring pairs of these 12 areas occurs at demographically low levels (≤4.25 females/year). Thus, despite their near-continuous distribution along the Alaska coast, harbor seals in Alaska consist of a series of discrete sub-populations that seldom exchange individuals.

5. These findings are consistent with other information relating to harbor seal dispersal patterns and population structure, including movement patterns, trends in abundance and foraging ecology.

**Conclusions**

1. These findings indicate that current stocks of harbor seals in Alaska are too broadly defined to meet the management objectives of the MMPA of maintaining population stocks as functioning elements of their ecosystem.

2. These findings also provide a framework for the identification of more meaningful management stocks and highlight the need for a re-appraisal of other information of relevance to stock structure including the interpretation of information on distribution,
movement patterns, trends in abundance and foraging ecology as well as the incorporation of traditional ecological knowledge.

3. The genetic study is still limited by sample coverage. Substantial gaps exist in areas of high conservation concern (see the non-circled areas in Figure ES-3), including the Aleutian Islands, the Alaska Peninsula, the northeastern Gulf of Alaska and parts of Southeast Alaska and the Kodiak Archipelago. Active collaboration with Alaska Native subsistence hunters and directed sampling is necessary if these important areas are to be sampled.

4. Although further sampling is needed to refine stock boundaries, the conclusion that there are multiple small units that need to be managed as separate stocks is not likely to change.
Figure ES-2. The 31 initial units defined using information on movement patterns and distribution of harbor seals in Alaska. For details on each stratum see Table 1. The 15 units that were excluded from the final analysis of population genetic structure due to low adjusted sample size and highlighted.
Figure ES-3. Graphic representation of the stage in the Boundary Rank analysis when all units were significantly differentiated from their neighbors at a chi-squared permutation p-value of < 0.1.
1. INTRODUCTION

Harbor seals (Phoca vitulina) inhabit the nearshore and continental shelf waters of Alaska from Dixon Entrance in the southeast, west throughout the Gulf of Alaska and the Aleutian Archipelago and as far as Kuskokwim Bay in the Bering Sea (Figure 1). This important marine predator occupies a diverse range of habitats, hauls out at thousands of discrete coastal sites (Hoover-Miller 1994; Rice 1998; Shaughnessy and Fay 1977; Withrow and Cesarone, 1998; Withrow et al., 1999, 2000, 2001, 2002; Boveng et al. 2003) and represents a significant marine resource to a range of users. Once considered abundant throughout their Alaska range, harbor seals have declined dramatically in some areas over the past few decades while in other areas their numbers have increased or remained stable over similar time periods (Pitcher, 1990; Frost et al., 1999; Matthews and Pendleton, 2000; Small et al., 2003). These declines and differences in trend highlight the need for the definition of management stocks at appropriate spatial scales in which abundance, population trend and other aspects of harbor seal biology can be monitored in relation to both direct and indirect human-caused mortality. Assessing the appropriate stock structure allows effective measures to be taken to fulfill the management objectives of preventing depletion, promoting recovery and maintaining the species range while ensuring harbor seals remain a sustainable resource for all users (Marine Mammal Protection Act (MMPA) 1972 as amended 1994; Alaska Native Harbor Seal Commission (ANHSC)-National Marine Fisheries Service (NMFS) Co-management Agreement, 1999).

The National Marine Fisheries Service (NMFS) currently recognizes three separate stocks of harbor seals in Alaska: the Southeast Alaska Stock, the Gulf of Alaska Stock and the Bering Sea Stock (Figure 1), identified primarily on the basis of regional differences in trends in abundance (Small and DeMaster, 1995; Hill et al., 1997; Angliss and Lodge, 2002). At the time of their designation, however, it was recognized that large gaps existed in our knowledge of dispersal and movement patterns and population structure, and it was recommended that more information on these aspects of seal biology be acquired to define more meaningful management units (Small and DeMaster, 1995; Marine Mammal Commission, 1995; DeMaster 1996). Over the past decade, a large body of research has been conducted that greatly improves our understanding of harbor seal population structure, including further trend studies and directed studies on movement patterns and population genetic structure.

The analysis of variation within genetic markers can reveal much about structure, dispersal patterns and gene flow in natural populations that can aid in the identification of management units, also known as stocks (Dizon et al., 1992; Mortiz, 1994; Avise, 1994, 1995; see Appendix 1). In this report we present our findings from an extensive molecular genetic study of population subdivision and dispersal patterns of harbor seals in Alaska. We integrate and compare our genetic results to recent findings from studies on harbor seal movement and foraging behavior, trends in abundance, and diet in order to better resolve the spatial pattern of population structure of this species in Alaska. Our studies indicate that current stocks are too broadly defined to meet the management objectives of the MMPA of maintaining population stocks as functioning elements of their ecosystem. The genetic findings provide a framework for the definition of more meaningful stocks.
Figure 1. The distribution of harbor seals, *Phoca vitulina*, in Alaska. The boundaries of the three currently recognized management stocks are indicated by dashed lines.
1.1 Trends in abundance, and management concerns

Over the past few decades, harbor seals have experienced dramatic declines in numbers in some parts of their Alaskan range, most notably in the Gulf of Alaska (Pitcher, 1990; Frost et al., 1999). For example, seal numbers on the Trinity Islands, at the southern end of the Kodiak Archipelago, decreased from nearly 17,000 animals in 1956 (Mathisen and Lopp, 1963) to fewer than 1,700 in 1996 (Withrow and Loughlin, 1997) an overall decline of 90%, while counts at terrestrial sites in Prince William Sound decreased by 63% between 1984 and 1997 (Frost et al., 1999). The causes for these declines are as yet unresolved but may involve several factors including ecosystem level changes and both direct and indirect human-caused mortality, all of which may differ on a regional scale.

Harbor seals are an important subsistence resource for many Alaska Native coastal communities, with an estimated annual take of 2,200-2,800 seals (Wolf, 2001; ANHSC, 2001). Management objectives are thus primarily concerned with maintaining harbor seals as functioning elements of their ecosystem while ensuring that this species remains a sustainable resource (MMPA as amended 1994; ANHSC-NMFS, 1999).

1.2 Defining stocks

The word “stock” is a management term defined in the context of a particular management regime, in this case by law in the MMPA. The Act specifies that endangered or depleted species and “population stocks should not be permitted to diminish beyond the point at which they cease to be a significant functioning element in the ecosystem of which they are a part, and consistent with this major objective, they should not be permitted to diminish below their optimum sustainable population...” It further states that “the primary objective of their management should be to maintain the health and stability of the marine ecosystem. Whenever consistent with this primary objective, it should be the goal to obtain an optimum and sustainable population keeping in mind the carrying capacity of the habitat." As guidance to interpreting these management objectives, the Act defines "optimum sustainable population" (OSP) as: "with respect to any population stock, the number of animals which will result in the maximum productivity of the population or the species, keeping in mind the carrying capacity of the habitat and the health of the ecosystem of which they form a constituent element." By regulation, NMFS has defined population stocks to be at OSP when they are between carrying capacity and the maximum net productivity level (MNPL) (Gerrodette and DeMaster 1990). Furthermore, the Act defines "population stock" as "a group of marine mammals of the same species or smaller taxa in a common spatial arrangement, that interbreed when mature." In a paper aimed at giving better quantitative guidance to implement these goals, Taylor (1997) emphasized the importance of a functional definition for population stock for use in calculating the number of animals that can be removed from stocks and that is consistent with both the population stock and ecosystem management goals of the Act. Because this functional definition provides the underlying logic for the harbor seal analyses that follow, we review material from that paper here.

The 1994 Amendments to the Act allow regulation of certain types of human-caused mortality through the calculation of Potential Biological Removals (PBRs). One
element of the equation used to calculate PBR is an estimate of abundance for the population stock. Use of the term population stock implies that both a biological (population) and a management (stock) meaning were intended. For brevity, we use "stock" instead of "population stock" and it carries the same double meaning: (1) groups that are delineated by a very low rate of genetic exchange, or (2) groups of animals that are essentially demographically separate and hence may experience differential risk and therefore need to be managed separately. A group of animals is considered "demographically separate" when the rate of animals coming into that population from its neighbors is so low that if this population were to decline it would not be prevented from becoming locally extinct by dispersal from its neighbors. These two meanings are often referred to by separate names, the former being called an evolutionarily significant unit (ESU) and the latter, a management unit (MU). The reason for defining populations as ESUs is to preserve the genetic diversity and evolutionary potential of the species (Waples 1991). Dizon et al. (1992) offer a phylogeographic approach that categorizes 'stocks' as to their probability of being an ESU, the unit defined as a 'species' under the Endangered Species Act. Note that the use of the term 'stock' as an ESU is not common and we do not use 'stock' to mean ESU in this paper. Instead we agree with Moritz (1994), who regards stocks to be synonymous with MUs and argues that they are the logical unit for short-term management. Perrin and Brownell (1994) contend that "stock" identity cannot be divorced from the management strategy adopted. There is no doubt that population units that are significant in an evolutionary sense qualify as population stocks under the MMPA. However, preserving only evolutionarily significant units could allow reduction and/or fragmentation of present ranges and thus violate the ecosystem goals of the MMPA.

As an example, consider the schematic distributions in Figure 2. Distribution "a" is the pristine distribution where size of the oval represents abundance and height represents geographic distance. Constrictions in this schematic represent limited movement such that this distribution could be described as a series of populations or subpopulations connected by dispersal (the aggregate is often called a metapopulation). If we reduced abundance by 50%, we could obtain any of the other distributions: b, range contraction; c, range fragmentation; or d, range maintenance. Although all may meet the population goal of maintaining harbor seals within OSP, i.e., about 50% of historical abundance (K), b and c likely do not meet the ecosystem goal. However, because there are no stock definition "rules" for calculating PBRs, any of these alternatives could occur depending on the distribution of human-caused mortality.

The Act's definition of population stock gives little guidance. Unfortunately, for most species managers have found it impossible to use the criterion "interbreed when mature." If we interpret "interbreed when mature" to represent the degree of genetic interchange, then nature presents us with a continuum. Some geographically separate groups of animals may exchange members at the rate of one per generation and others at the rate of one percent per year. If we restrict our definition of population stock to only those virtually closed populations exchanging individuals at the rate of only a few individuals per generation then we will only have stock boundaries encircling large geographic ranges. Calculating the PBR based on abundance estimates for these units may allow depletion of areas with large human-caused mortalities, i.e., result in distributions b and c.
Figure 2. Distribution of pristine populations (a), versus potential distributions after 50% of the total abundance is removed (b, c, d). The area of the oval represents abundance, height represents geographic distance.
In the case of Alaska harbor seals, reduction of the range is also undesirable because it could violate the goal of sustained use by Alaska Natives. For example, the loss of the large population in scenario c may mean that hunters that traditionally hunted this population may have to travel further and spend more time to obtain the same number of seals.

Irrespective of difficulties defining stocks, NMFS must nevertheless draw lines on a map to represent stock boundaries, and they must do it for the 48 marine mammal species that occur in U.S. waters. Available data for making such stock boundary decisions ranges from very crude distributional data to very detailed data on movement, morphology, genetics and distribution. Most of the time, however, to meet the Act's management objectives, the implementing agency must make the best possible decisions in the face of considerable uncertainty.

There are two types of errors that can be made in making these decisions: (1) incorrect lumping of stocks, which could result in depleting or even eliminating one or more stock, or (2) incorrect splitting, which may unnecessarily restrict human activities. Call the first the "under-protection error," and the second the "over-protection error." To calculate the probabilities of making these errors, management objectives must be defined quantitatively. For purposes of illustration, we assume that population growth is logistic and thus MNPL occurs at 50% of carrying capacity (K). If maintenance of populations above 50% of K were the only objective, any of the distributions in Figure 2 would be an acceptable management outcome. However, because the Act emphasizes ecosystem integrity, a more comprehensive management target is required, i.e., one that considers range. Only Figure 2d would be an acceptable outcome if management objectives are to both maintain stocks above MNPL and maintain an unfragmented and undiminished geographic distribution.

Participants in a workshop to provide guidelines for implementing the MMPA recognized that maintaining the range would serve to meet the ecosystem goals (Wade and Angliss, 1997). To attain this objective, the definition of stocks should be based on demographic isolation where dispersal rates are "insufficient to deliver enough recruits from an unexploited population (source) to an adjacent exploited population (sink) so that the latter remains a functioning element of its ecosystem. Insufficient dispersal between populations where one bears the brunt of exploitation coupled with their inappropriate pooling for management could easily result in failure to meet MMPA objectives ". The report goes on to state that while careful consideration must be given to the definition process, "there is no intent to define stocks that are clearly too small to represent demographically isolated biological populations " (for more details on stock definition guidelines, see Appendix 2).

Harbor seals do not have obvious gaps in their distribution that allow easy stock definition (see below) and hence the first attempt at defining stocks used information on trends in abundance. The interpretation of data in this paper strives to meet the definition guidelines above in choosing a scale that will allow management to meet the objectives of maintaining the range while avoiding management units that are so small that demographic independence is not plausible. Fortunately, management objectives that seek to maintain the range of seals are completely compatible with maintaining locally
sustainable harvests, as both seek to identify the scale at which seal population dynamics are essentially independent.

1.3 Evidence of population structure in Alaskan harbor seals

Here we summarize data relevant to defining stock structure for harbor seals, much of which has been gathered during the 8 years since the provisional stocks were recognized. Tagging studies have shown that despite occasional long-distance movements (Bonner and Withams 1974; Lowry et al. 2001; Swain et al., 1996; Thompson et al. 1994), harbor seals in both the Atlantic and Pacific Oceans rarely forage more than 50 km (31 miles) away from haul-out sites (Thompson 1993; Lowry et al. 2001; Small and Ver Hoef, 2001). This, combined with the high rate of recorded returns to haul-out sites (Lowry et al. 2001; Swain and Small, 1997) has helped characterize harbor seals as a relatively sedentary species, exhibiting long-term fidelity to particular areas. Their distribution in Alaska, although effectively continuous, is non-uniform, with several centers of high density separated by areas of low density that often coincide with changes in habitat.

Studies on trends in abundance have revealed differing trajectories between areas within both the Gulf of Alaska and Southeast Alaska (Frost et al., 1999; Small et al., 2003), providing strong circumstantial evidence of population structuring within these two stocks. Information on trends in the Bering Sea stock is more fragmentary, yet available data suggest differences in trend among some areas over the past decade (Withrow and Loughlin, 1996; Jemison et al., 2001; Small et al. 2003). Regional differences in habitat and prey (Pitcher, 1980; Hoover-Miller 1994; Iverson et al., 1997), and morphology and reproductive physiology (Bigg 1969; Burns and Gol'tsev 1984; Burns et al. 1984; Kelly 1981; Shaughnessy and Fay 1977; Tempte et al. 1991) further suggest that harbor seals are structured on a finer-scale than reflected by the current three stocks. Finally, a recent molecular genetic study revealed strong genetic differentiation between five centers of abundance from throughout the State, again indicating that the current stocks fail to capture the true population structure of the species (Westlake and O’Corry-Crowe, 2002).

1.4 Difficulty of resolving population structure in a continuously distributed species

Despite the growing body of evidence indicating fine-scale population structure in Alaskan harbor seals, actually describing that structure has proven difficult. Most available methods of examining genetic structure require the researcher to divide their samples into what they think might be populations (or stocks) before statistical tests are performed. In discretely distributed species, distributional gaps can be used to guide stratification, as it can generally be safely assumed that movement rates are lower across such areas of low density. This assumption can be tested either directly (e.g., mark-recapture, telemetry) or indirectly (e.g., genetic markers) (Slatkin, 1985; Chepko-Sade and Halpin, 1987). However, in continuously distributed species there are few, if any, obvious distributional breaks, making choice of putative populations much more difficult. Furthermore, the marine environment presents few physical impediments to movement (at least that are obvious to humans), which, combined with the high mobility of many marine organisms, could facilitate long-distance movements and gene flow...
Consequently, population subdivision in marine species tends to be subtle, shaped more by behavioral and ecological factors than by obvious physical limitations to dispersal.

The traditional approach to resolving population structure has been to collect samples at a series of geographically separated sampling sites from across the distribution and look for statistically significant genetic differences among them (e.g., Lamont et al., 1996; Goodman, 1998; Stanley et al., 1996). While useful at revealing general patterns of population structure, this approach cannot fully uncover the scale and form of population partitioning in a continuously distributed species, nor can it be used to identify the precise locations of population boundaries, if they exist (Martien and Taylor, in press). Westlake and O'Corry-Crowe (2002) used such an approach when examining large-scale population structure of Pacific harbor seals. Their findings revealed much about the evolutionary history of this species, documented macro-geographic patterns of population subdivision, and showed that population structure adhered to a general isolation-by-distance model. However, while they found, for instance, that there was strong genetic differentiation between Prince William Sound and the Kodiak Archipelago, their analysis did not provide any guidance on where the stock boundary separating those two areas should be placed or whether the area between Prince William Sound and Kodiak constituted its own stock. Thus, the utility of such an approach is limited when defining stocks, since the exact placement of stock boundaries has important conservation and socio-economic consequences.

1.5 Population genetic structure and the definition of stocks of Alaskan harbor seals

We used patterns of variation within the mitochondrial genome (mtDNA) to investigate fine-scale population structure of Alaskan harbor seals. We employed a number of methods to analyze the genetic data, including the geographically-constrained cluster analysis, Boundary Rank (BR) (Martien and Taylor, in review). This method constrains the cluster sequence so that the resulting hypotheses are consistent with what is known about the behavior of the study species. Unlike other currently used methods, BR has been extensively tested for performance using MMPA management goals. It was also specifically designed to suggest boundary placement for continuously distributed species. We compared the results of the BR analysis to those obtained using classical distance-based clustering and phylogeny reconstruction analyses and a traditional hypothesis testing approach in order to examine how consistent these approaches were with the new method. Based on estimates of dispersal rates between the hypothesized stocks suggested by BR, we conclude that there are at least 12 genetically distinct units within the state of Alaska between which dispersal is low enough to warrant separate management. There are, however, still many areas of the state where we have little or no genetic data and further investigation in these areas will likely reveal more structure. Nevertheless, these 12 units provide a framework for the definition of more biologically meaningful management stocks of harbor seals in Alaska.
2. MATERIALS AND METHODS

2.1 LABORATORY METHODS

2.1.1 Sample collection

Tissue samples were collected from 881 harbor seals at 180 locations throughout Alaska (Figure 3, Table 1). All age classes and both sexes were represented, and samples were collected in all seasons. Because many samples were collected opportunistically for purposes other than the genetic analysis of population structure, the number of samples collected in an area is not proportional to the abundance of the area (Table 2).

2.1.2 DNA extraction

Several tissue types were used. Skin biopsies were taken from seals during tagging and rehabilitation operations; shed hair was collected at molting sites; and skin, muscle, and liver samples were collected from subsistence hunts, strandings, and scientific collections. Most tissues were preserved in 20% dimethyl sulfoxide (DMSO) saturated in NaCl (Amos and Hoelzel 1991). Molted hair samples and some tissues were stored at -80°C, and several skin samples were fixed in formalin then stored in ethanol. Total DNA was isolated from most samples using standard SDS lysis-proteinase K digestion or CTAB methods followed by phenol:chloroform extraction and ethanol precipitation protocols (Sambrook et al. 1989; Winnepenninckx et al. 1993). More recently, the tissue lysis and digestion steps were automated by using the FastDNA® kit and the FastPrep® instrument (BIO 101, Carlsbad, California). DNA from molted hair was successfully extracted using the lysis buffer of Tikel et al. (1996) followed by several phenol:chloroform extractions (DeAngelis, 2000). The salting-out protocol of Miller et al. (1988) was used to extract DNA from 20-year old formalin-preserved tissue. Concentration and quality of the purified DNA from all samples were estimated by spectrophotometry and visualized on 1% agarose gels stained with ethidium bromide (EtBr).

2.1.3 Mitochondrial DNA amplification and sequencing

A 588 base-pair (bp) region of the mitochondrial genome was then amplified by the polymerase chain reaction (PCR; Saiki et al., 1988) with custom designed primers (Kocher et al., 1989, Rosel et al., 1994; Westlake and O’Corry-Crowe, 2002). The amplified PCR products were purified by membrane-based filtration using Microcon® (Millipore, Bedford, Massachusetts) or QIAquick® (Qiagen, Valencia, California) columns. A total of 435 bp of the mtDNA control region and adjacent proline tRNA gene were sequenced by the direct dideoxy sequencing method of Sanger et al. (1977) using 4-dye fluorescent technology of Applied Biosystems (ABI, 1992). Excess dye-labeled terminators were removed from sequencing reactions using Centri-Sep™ spin columns (empBiotech GmbH, Berlin, Germany) or by ethanol precipitation. Sequences were electrophoresed on an ABI 373A or ABI 377 automated sequencer, and were edited and aligned with the SeqEd™ multiple-sequence editor program (ABI, 1992).
Figure 3. The distribution of 180 sampling sites of harbor seals in Alaska. 50 km radius buffer zones are indicated around each site.
Table 1. Summary of the number of sampling sites, the number of samples \((n)\), the proportion of individuals with unique haplotypes, haplotypic diversity \((H)\) and the adjusted sample size \((n_\text{a})\) within the 31 initial harbor seal units in Alaska. Units with \(n_\text{a} \leq 4\) are highlighted.

<table>
<thead>
<tr>
<th>site no.</th>
<th>Unit name</th>
<th>no. of sampling sites</th>
<th>sample size (n)</th>
<th>proportion unique</th>
<th>haplotypic diversity (H)</th>
<th>adjusted sample size (n_\text{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ketchikan</td>
<td>8</td>
<td>42</td>
<td>0.781</td>
<td>0.981</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Grand Island</td>
<td>8</td>
<td>34</td>
<td>0.609</td>
<td>0.977</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Red Bay</td>
<td>1</td>
<td>8</td>
<td>1.000</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Wrangell</td>
<td>1</td>
<td>10</td>
<td>0.714</td>
<td>0.911</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Frederick S(^{nd})</td>
<td>6</td>
<td>44</td>
<td>0.667</td>
<td>0.948</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>Angoon</td>
<td>3</td>
<td>9</td>
<td>1.000</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Vixen-Sitka</td>
<td>13</td>
<td>50</td>
<td>0.520</td>
<td>0.959</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>Glacier Bay</td>
<td>2</td>
<td>27</td>
<td>0.824</td>
<td>0.923</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>Yakutat Bay</td>
<td>3</td>
<td>21</td>
<td>0.882</td>
<td>0.967</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Icy Bay</td>
<td>1</td>
<td>20</td>
<td>0.944</td>
<td>0.984</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>Copper River Delta</td>
<td>3</td>
<td>22</td>
<td>0.643</td>
<td>0.939</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>East Prince William S(^{rd})</td>
<td>15</td>
<td>79</td>
<td>0.625</td>
<td>0.970</td>
<td>39</td>
</tr>
<tr>
<td>13</td>
<td>North Prince William S(^{rd})</td>
<td>10</td>
<td>30</td>
<td>0.727</td>
<td>0.975</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>South Prince William S(^{rd})</td>
<td>15</td>
<td>87</td>
<td>0.682</td>
<td>0.964</td>
<td>43</td>
</tr>
<tr>
<td>15</td>
<td>East Kenai</td>
<td>13</td>
<td>47</td>
<td>0.593</td>
<td>0.964</td>
<td>20</td>
</tr>
<tr>
<td>16</td>
<td>West Kenai</td>
<td>10</td>
<td>23</td>
<td>0.895</td>
<td>0.957</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>Kamishak</td>
<td>2</td>
<td>20</td>
<td>0.800</td>
<td>0.958</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>North Kodiak</td>
<td>12</td>
<td>27</td>
<td>0.960</td>
<td>0.922</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>West Kodiak</td>
<td>4</td>
<td>37</td>
<td>0.636</td>
<td>0.958</td>
<td>15</td>
</tr>
<tr>
<td>20</td>
<td>East Kodiak</td>
<td>10</td>
<td>29</td>
<td>0.650</td>
<td>0.973</td>
<td>9</td>
</tr>
<tr>
<td>21</td>
<td>South Kodiak</td>
<td>5</td>
<td>51</td>
<td>0.759</td>
<td>0.954</td>
<td>22</td>
</tr>
<tr>
<td>22</td>
<td>Puale Bay</td>
<td>1</td>
<td>8</td>
<td>1.000</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>Shumagin</td>
<td>1</td>
<td>5</td>
<td>0.750</td>
<td>0.900</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>Sanak</td>
<td>2</td>
<td>11</td>
<td>0.900</td>
<td>0.982</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>Akutan-UnAlaska</td>
<td>3</td>
<td>8</td>
<td>1.000</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>Atka</td>
<td>2</td>
<td>6</td>
<td>0.500</td>
<td>0.867</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>Nelson-Prt. Moller</td>
<td>3</td>
<td>8</td>
<td>0.714</td>
<td>0.972</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>Prt. Heiden</td>
<td>1</td>
<td>8</td>
<td>0.667</td>
<td>0.929</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>Ugashik-Egegik</td>
<td>3</td>
<td>35</td>
<td>0.647</td>
<td>0.901</td>
<td>18</td>
</tr>
<tr>
<td>30</td>
<td>Togiak-Nanvak</td>
<td>3</td>
<td>32</td>
<td>0.750</td>
<td>0.942</td>
<td>12</td>
</tr>
<tr>
<td>31</td>
<td>Pribilof Islands</td>
<td>2</td>
<td>16</td>
<td>0.800</td>
<td>0.892</td>
<td>6</td>
</tr>
</tbody>
</table>

\*A further 27 samples from 14 sites were analyzed but were not placed in any of the initial starting strata.

\^East Kodiak, despite having an \(n_\text{a}=9\) was excluded from the analysis of population subdivision because of uncertainty in defining the limits of this unit.
Table 2. Sample size and the number of sampling locations in relation to harbor seal abundance. The Alaska range is divided into 12 large regions. Minimum abundances are based on raw counts of seals hauled out during aerial and land-based surveys.

<table>
<thead>
<tr>
<th>Region*</th>
<th>Minimum† abundance</th>
<th>No. of sampling sites</th>
<th>Sample size</th>
<th>% total abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southeast Alaska - south</td>
<td>26,106</td>
<td>18</td>
<td>94</td>
<td>0.36</td>
</tr>
<tr>
<td>Southeast Alaska - north</td>
<td>18,557</td>
<td>25</td>
<td>132</td>
<td>0.71</td>
</tr>
<tr>
<td>Gulf of Alaska - east</td>
<td>4,681</td>
<td>4</td>
<td>41</td>
<td>0.88</td>
</tr>
<tr>
<td>Copper River Delta</td>
<td>3,053</td>
<td>4</td>
<td>27</td>
<td>0.88</td>
</tr>
<tr>
<td>Prince William Sound</td>
<td>2,952</td>
<td>40</td>
<td>196</td>
<td>6.64</td>
</tr>
<tr>
<td>Kenai - Cook Inlet</td>
<td>2,796</td>
<td>27</td>
<td>90</td>
<td>3.22</td>
</tr>
<tr>
<td>Kodiak</td>
<td>4,387</td>
<td>39</td>
<td>162</td>
<td>3.70</td>
</tr>
<tr>
<td>Alaska Peninsula - south</td>
<td>3,167</td>
<td>4</td>
<td>24</td>
<td>0.76</td>
</tr>
<tr>
<td>Aleutian Islands</td>
<td>3,508</td>
<td>5</td>
<td>14</td>
<td>0.40</td>
</tr>
<tr>
<td>Bristol Bay - south</td>
<td>2,570</td>
<td>5</td>
<td>17</td>
<td>0.66</td>
</tr>
<tr>
<td>Bristol Bay - north</td>
<td>6,169</td>
<td>7</td>
<td>68</td>
<td>1.1</td>
</tr>
<tr>
<td>Pribilof Islands</td>
<td>202</td>
<td>2</td>
<td>16</td>
<td>7.92</td>
</tr>
</tbody>
</table>

| Total                             | 180                | 881                   |             |                   |

* these regions do not represent stock delineations but are defined based primarily on aerial survey design, and thus should not be confused with units identified by genetic analyses.
Further details of the laboratory methods used to extract DNA from the various tissue types and to analyze sequence variation within the mitochondrial DNA control region are given in Westlake and O’Corry-Crowe (2002).

2.2 DATA ANALYSIS

2.2.1 Analysis of genetic diversity

The amount and nature of variation within the sequenced region were assessed by determining the number of variable sites and the number of unique haplotypes using MEGA 2.1 (Kumar et al. 2001) and MacClade 3.02 (Maddison and Maddison 1992) software, respectively. Haplotypic (H; Nei and Tajima 1981) and nucleotide (π; Nei, 1987) diversity estimates of genetic variation within geographic strata were calculated using ARLEQUIN 1.1 software (Schneider et al. 1997).

2.2.2 Analysis of genetic differentiation

We used two approaches to examine genetic differentiation. The first was a parameter estimation approach in which we defined geographic units using a clustering analysis and then estimated dispersal between those units using the program Migrate (Beerli and Felsenstein, 2001). The primary method used for clustering was Boundary Rank (BR) (Martien and Taylor, in review), a geographically constrained clustering method that uses the estimated genetic relationships between sampling sites to define geographically contiguous units. Because BR is a new method that is likely to be unfamiliar to most readers, we compared the results of the BR clustering to those of two traditional distance-based clustering and phylogeny reconstruction analyses, Unweighted Pair Group Method using Arithmetic averages (UPGMA; Sokal and Michener, 1958; Sneath and Sokal, 1973) and neighbor joining (N-J; Saitou and Nei, 1987). Comparing these commonly used methods to the results of the BR method can suggest how robust the results are and highlight the differences in BR that are pertinent to defining harbor seal stocks. The second approach we used was a standard hypothesis testing analysis in which population structure was assessed by looking for statistically significant genetic differentiation.

The measure of genetic differentiation used by BR is \( \chi^2 \) per degree of freedom (\( \chi^2 / \text{dof} \)). Extensive performance testing showed that \( \chi^2 / \text{dof} \) was the measure that resulted in the best performance of the method (Martien and Taylor, in review). This result is consistent with Hudson et al.'s (1992) study that showed that \( \chi^2 \) is the most powerful statistic for detecting genetic differentiation in mtDNA markers. Consequently, we chose to use \( \chi^2 / \text{dof} \) as the distance measure for both the UPGMA and neighbor joining analyses and assessed statistical significance in the hypothesis-testing analysis using a \( \chi^2 \) permutation test. However, for completeness, we also estimated the more traditional index of genetic differentiation, Wright's (1951) \( F_{st} \), by the analysis of variance method of Weir and Cockerham (1984) using ARLEQUIN 1.1.
2.2.2.1 **Definition of initial units**

The three clustering methods that we used (UPGMA, N-J and BR) and the hypothesis-testing analysis all require that samples first be divided into initial units. For consistency, we used the same initial units in all four analyses. We used information on movement patterns and distribution to define a set of biologically meaningful initial units from the 180 distinct sampling locations (Figure 3). Treating each of the 180 sampling sites as an initial unit was not practical because we had only one or two samples from many of those sites. Furthermore, many of the sampling sites were so close to each other (for example, neighboring beaches in the same small bay) that treating each of them as separate units did not make biological sense.

Numerous studies have been conducted on harbor seal movement patterns in both the North Pacific and North Atlantic (e.g., Bonner and Witthames, 1974; Pitcher and McAllister, 1981; Harkonen and Harding, 2001; Lowry et al., 2001; Small and VerHoef, 2001; Thompson, 1989, 1993; Thompson et al., 1994). Though few of these investigations monitored animals for long enough periods to confirm dispersal or measure dispersal distances, many were sufficient to estimate minimum home range size and the minimum extent of annual movements for different age and sex classes. In all studies to date, the vast majority of seals stayed within 50km (31 miles) of the original capture area. To illustrate the geographic scale of these typical movements relative to the species range, the map in Figure 3 depicts 50km radius buffer zones around each sampling site. Based on this general pattern of limited movement, we created our initial strata by conservatively grouping sampling sites that were within 50km of each other. This sometimes resulted in grouping a string of sites along a coastline, the extremes of which were greater than 50km apart. In a number of cases, neighboring sites that were up to 100km apart were grouped together. This occurred when sample size was low and the next nearest sampling sites were much further (>100km) distant. Geographical features, such as channels and straits, and gaps in haul-out distribution (Figure 4) were used in some cases to refine the delineation of the initial units. These criteria resulted in the definition of 31 initial units (Figure 5). A number of sampling sites (n = 14 out of 180) were greater than 100km from the nearest neighboring site and had sample sizes too low (<4) to be considered initial units on their own. Data from these sites were excluded from further analysis.

While the rule of grouping sites within 50km of each other worked well in most areas, in Kodiak and Prince William Sound, the definition of initial units proved more difficult. In these areas, grouping all sites within 50km would have resulted in initial units that spanned several hundred kilometers (Figure 3). Therefore, we had to rely heavily on distributional gaps, habitat differentiation and subjective judgement when defining initial units in these areas. Figure 5 shows the primary set of initial units used in the analysis. However, we also performed a large number of sensitivity trials using alternate sets of initial units in these areas to examine the impact that the unit definitions had on our results.

Once the 31 initial units were delineated, we examined the sample size within each of them to determine whether it was sufficient for inclusion in the final analysis. The measures of genetic differentiation we used in all of our analyses ($\chi^2$/dof, $F_{ST}$) are
Figure 4. The abundance and distribution of harbor seals hauled out at terrestrial and glacial ice sites during fall molt surveys in Alaska. Data provided by J. Bengtson and D. Withrow, Polar Ecosystems Program, National Marine Mammal Laboratory, NMFS, Seattle, Washington. Data from the Pribilof Islands, Glacier Bay and Lake Iliamna are not present on this map.
Figure 5. The 31 initial units defined using information on movement patterns and distribution of harbor seals in Alaska. For details on each stratum see Table 1. The 15 units that were excluded from the final analysis of population genetic structure due to low adjusted sample size and highlighted.
based on the estimated haplotype frequencies in the two units being compared. How well a sample of a given size represents the actual haplotype frequencies in a population depends on the haplotypic diversity, \( H \). When sample size is low relative to diversity, estimates of genetic differentiation will be negatively biased (Appendix 3). Including such poorly sampled sites in the analyses could result in the definition of hypothesized units that are based more on sampling considerations than actual genetic relationships. Therefore, we determined whether the sample size within each of our initial 31 units was adequate relative to \( H \) using a statistic called the adjusted sample size (described in Appendix 3). The actual sample size \( (n) \) and adjusted sample size \( (n, ) \) for each of the 31 initial units is shown in Table 1. The adjusted sample size reflects the relationship between abundance and \( H \) as well as how well \( n \) represents the actual haplotype frequencies. For example, both Icy and Yakutat Bays have small \( n, \) because, although they have moderately large \( n \) (20 and 21, respectively), their large abundances \( (N_{\text{uncorrected}} = 2,754 \text{ and } 1,290 \text{ respectively}; \text{Withrow and Cesarone, 1998}) \) result in high \( H \) and mean that at most only 0.72% and 1.63% of these populations have been sampled. The Pribilof Islands, on the other hand, have a larger adjusted sample size \( (n, = 6) \) than either Icy Bay or Yakutat Bay despite having a smaller absolute sample size \( (n = 16) \). This is because the Pribilofs only have an estimated abundance of 202 \( (N_{\text{min}}; \text{Jemison, 1996}) \). Thus \( H \) is relatively low and as much as 8% of the population has been sampled.

In order to prevent our analyses from being biased by low sample size relative to haplotypic diversity, we excluded from the final analyses all initial units for which the adjusted sample size \( n, \leq 4 \). In addition, we excluded east Kodiak due to the uncertainty associated with the definition of this initial unit. This resulted in the exclusion of 15 units, including all 7 from the north and south sides of the Alaska Peninsula and the Aleutians, indicated in bold in Table 1 and shaded in Figure 5. The unshaded units in Figure 5 show the remaining 16 initial units used in the analyses.

### 2.2.2 Parameter estimation

**Boundary Rank** - Boundary Rank is a hierarchical clustering method that uses mtDNA data to generate sets of hypothesized units. The analysis begins by calculating \( \chi^2/\text{dof} \) between 'connected' pairs (see below) of geographically small initial units. The pair that exhibits the greatest similarity is coalesced into a single larger unit, thereby reducing the number of units by one. The process is repeated until all of the samples have been coalesced into a single unit. The result is a nested set of population structure hypotheses, each containing one fewer unit than the previous. These hypotheses can then be evaluated in light of the management goals the conservation units are intended to meet. Because they are based on the genetic relationships between the samples, these hypotheses are more likely to reflect the underlying population structure of the region than if the samples had been stratified on a purely subjective basis. Martien and Taylor (in review) used extensive simulation-based performance testing to demonstrate that the method performs well in a management context under a wide range of conditions.

Clustering methods have long been used in analyzing genetic data. The feature that distinguishes BR from previously published methods is that the clustering can be geographically constrained so that the geographic shapes of the clusters defined by BR conform to a specific model of population structure specified by the researcher. Thus,
BR allowed us to incorporate prior knowledge of the general form of harbor seal population structure into our analysis. Because it limits the number of comparisons made when deciding which units to cluster next, the geographic constraint renders the analysis more robust to sampling variance than other clustering methods. Similar geographic constraints have been used in the past in other studies aimed at understanding population structure. For instance, York et al. (1996) employed a geographically constrained clustering approach in their efforts to use trend data to model female population dynamics and metapopulation structure in Steller sea lions, *Eumetopias jubatus*.

The geographic constraint in BR is implemented through a connectivity matrix, which indicates which pairs of initial units have the potential to coalesce at the start of the analysis. At each step in the analysis, each unit can only be clustered with units to which it is ‘connected.’ Each time a pair of units is clustered, the connectivity matrix is updated so that the new larger unit is connected to everything the two original units were connected to. The connectivity matrix thus constrains the shapes of the units produced by Boundary Rank to reflect what is known about the behavior and movement patterns of the species in question.

The most common form of population structure observed in mammals is isolation-by-distance (Wright, 1943), in which individuals are more likely to move between geographically adjacent sites than they are between geographically distant sites. As mentioned earlier, tagging studies have shown that harbor seals exhibit high site fidelity, and that short movements between geographically adjacent sites are far more common than long-distance movements. A recent genetic study (Westlake and O’Corry-Crowe, 2002) showed that harbor seals exhibit a pattern of genetic isolation-by-distance in which sites that are close together are more closely related than sites that are far apart. Based on these convergent lines of evidence, we chose a connectivity matrix that reflects a stepping-stone model in which dispersal occurs only between adjacent sites. As such, connections were only allowed between initial units that were geographically adjacent (Figures 6A-C).

**UPGMA and Neighbor-Joining analyses** - We compared the results of the BR analysis to those of two more traditional genetic clustering methods, UPGMA and Neighbor-Joining (N-J). Both group units solely on the basis of estimated genetic distance, regardless of the geographic locations of the units. UPGMA links the two initial units with the smallest genetic distance first, followed by successively more distant units or groups of units. At each stage, the distance of a new group to all other units or groups is calculated as the mean distance of all the units in that group to the other units or groups. N-J was specifically developed to reconstruct evolutionary trees using distance data and seeks to minimize the overall length of the tree. We conducted both analyses using MEGA 2.1 (Kumar et al. 2001) software.

**Dispersal rate estimation** – As neighboring units are clustered together into fewer, larger units by the clustering algorithms, the level of genetic differentiation between units increases, reflecting the successively lower rates of dispersal that occur across the larger distances involved, as well as the larger sample sizes involved. A decision of what geographic scale to use when defining management units can thus be based on what rate of dispersal is necessary between units in order to meet a specific management objective.
Figure 6. The connectivity matrix used in the Boundary Rank analysis of population genetic structure.

A: Allowed connections among the 15 strata that were used in the analysis of population structure. Sampling sites are indicated in red, strata are outlined in blue, and the strata excluded based on limited samples are highlighted in jade.
Figure 6. The connectivity matrix used in the Boundary Rank analysis of population genetic structure.

B: Detail of the connectivity matrix for Southeast Alaska. Sampling sites are indicated in red, abundance is indicated in green, strata are outlined in blue, connections are indicated by dashed lines and the strata excluded based on limited samples are highlighted in jade.
Meeting the MMPA objective of preserving species as functioning elements of their ecosystem requires the definition of management units between which dispersal is low enough to render the units demographically independent, even if the dispersal rate is high enough to prevent evolutionary divergence.

In order to provide some guidance on how many stocks are necessary for this species to meet the management objectives of the MMPA, we focused on the clustering level at which all remaining units from the BR analysis had a $\chi^2$ permutation test $p$-value less than 0.1. We chose this starting point for a number of reasons. In standard hypothesis-testing the cut-off for considering a result to be statistically significant is typically arbitrarily set at $\alpha = 0.05$. In applied studies, the cut-off is often increased to 0.1 in order to introduce a small measure of precaution into the analysis by reducing the probability of making a type II error. We believe it is appropriate to use the precautionary value of 0.1 because the large abundance of harbor seals coupled with their continuous distribution make it very difficult to detect population structure using genetics.

To ensure that the choice of $\alpha = 0.1$ would result in the definition of units between which dispersal is low enough to justify separate management under the MMPA, we estimated dispersal between pairs of units using the program Migrate (Beerli and Felsenstein, 2001). This program produces an estimate of the long-term average number of dispersers moving from population A to population B each generation. We constrained the program so that the rate of movement (though not necessarily the actual number of dispersers) was equal in both directions. To translate this into an annual number of seals moving, we had to divide by the generation time, $T$. Using the life-history data presented in Pitcher and Calkins (1979) and the generation time formula from Caswell (2000), we calculated that the generation time of a harbor seal population at carrying capacity, $T_c$, is approximately 11.5 years. To calculate generation time when abundance is close to zero ($T_o$), we assumed that $r_{max} = 0.12$, the default value for pinnipeds used in the Potential Biological Removal scheme (Wade, 1998; Taylor and Wade, 2000). This resulted in an estimate of $T_o = 4.5$. We averaged $T_c$ and $T_o$ to calculate an average harbor seal generation time of 8 years.

We estimated dispersal between the first three pairs of units clustered by BR that had a permutation $\chi^2$ $p$-value < 0.1. Because Boundary Rank clusters units on the basis of their genetic differentiation, units that cluster later in the analysis have a greater degree of genetic differentiation and are therefore likely to have a lower rate of dispersal. Thus, if the first pair of units that had a $p$-value < 0.1 exchanged dispersers at a low enough rate to warrant separate management, its likely that the same would be true of pairs of units between which differentiation was even stronger.

### 2.2.2.3 Hypothesis testing

In addition to the clustering and dispersal rate estimation approach described above, we also analyzed genetic structure among the initial units using a traditional hypothesis-testing approach. Subdivision among the same initial strata used in the clustering and phylogeny reconstruction analyses was assessed using a $\chi^2$ contingency test of independence (Roff and Bentzen, 1989). The statistical significance of the $\chi^2$
statistic under a hypothesis-testing framework, where the null hypothesis, \( H_0 \), was random mixing, was estimated by 10,000 randomizations of the original data using ChiSquare software (K. Martien, Southwest Fisheries Science Center).

3. RESULTS

3.1 Overall genetic diversity

A total of 435 bp of the mtDNA control region and adjacent proline tRNA gene were analyzed for sequence variation in 881 harbor seals sampled from across Alaska (Table 1). Eighty nine variable sites were identified, 85 with substitutions (78 transitions and 9 transversions) and 4 with indels. Altogether, 243 different haplotypes were identified, with more than half (145) represented by single individuals. Overall haplotypic diversity was high (\( H = 0.975 \)) due to the large number of rare haplotypes, and overall nucleotide diversity was moderate (\( \pi = 1.47\% \)), suggesting that most haplotypes were closely related phylogenetically. The majority of unique haplotypes have been submitted to GenBank (accession numbers AF522643 - AF522866). The remainder will be submitted shortly.

3.2 Population subdivision

3.2.1 Parameter estimation

**Boundary Rank** - The results of the BR analysis are represented as a dendrogram in Figure 7. The analysis began with the original set of units and proceeded until all 16 were coalesced into a single unit. The first units to cluster are the two Bristol Bay units, Togiak-Nanvak and Ugashik-Egegik, indicating a close relationship between the two. The analysis also indicates a close relationship between the three Prince William Sound units and between these three and the Copper River Delta. As the analysis proceeds, clustering groups that are further and further apart both genetically and geographically, the two units from the inner waters of southeastern Alaska, Frederick Sound and Ketchikan, group together, followed by western Kodiak coalescing with Kamishak Bay, and the eastern Kenai Peninsula with the Prince William Sound-Copper River Delta cluster. A number of the initial units do not group with neighboring units until very late in the clustering sequence, notably, southern Kodiak (Tugidak and Sitkinak islands), the two units from the outer coast of southeastern Alaska, Vixen-Sitka and Grand Island, and the Pribilof Islands. This reflects the high degree of genetic differentiation between these areas and geographically adjacent areas.

**UPGMA and Neighbor-Joining** - The results of the UPGMA (Figure 8) and N-J (Figure 9) analyses, both of which grouped units solely on the basis of genetic distance, are very similar to those of Boundary Rank. The two Bristol Bay strata and the three Prince William Sound plus the Copper River Delta strata all group early, suggesting close relationships among the strata in these two areas. Similarly, Frederick Sound and Ketchikan cluster together and eastern Kenai groups with the Prince William Sound-Copper River Delta group. However, these analyses differed from the geographically constrained BR analysis by grouping Glacier Bay with the geographically distant.
Figure 7. Dendrogram of the outcome of the BR analysis. The analysis began with 16 units, as shown by the legend at the bottom of the dendrogram, and proceeded until all 16 units were clustered into a single unit. At each step, the two most similar units were combined. The depth of the nodes in the tree indicate the step in the analysis at which the units connected by that node were coalesced.
Figure 8. Dendrogram of the outcome of the UPGMA cluster analysis based on $\chi^2$/dof of the 16 initial units of harbor seals in Alaska.
Figure 9. Neighbor-joining tree based on $\chi^2$/dof of the estimated relationships among the 16 initial units of harbor seals in Alaska.
Kamishak Bay, and linking western Kodiak with the even more distant Pribilof Islands. Analyses using $F_{st}$ as the estimate of genetic distance yielded similar results. $F_{st}$ among neighboring units tended to be low ($F_{st} \leq 0.02$) resulting in the early grouping of geographically proximate areas (data not shown).

**Dispersal rate estimation** - Once the samples had been coalesced into twelve units by BR (Figure 10), all remaining units were significantly genetically differentiated from their neighbors ($\chi^2$ permutation p-value < 0.1). The first three pairs of units to coalesce after this point were Ketchikan and Frederick Sound, West Kodiak and Kamishak Bay, and Prince William Sound/Copper River Delta and the Kenai Peninsula, in that order (Figure 7). We estimated the number of females moving per generation between these three pairs of units using the program Migrate (Beerli and Felsenstein, 2001) and converted into an annual movement rate by dividing by an average generation time of 8 years. The largest estimated movement rate was 4.25 females per year moving from Frederick Sound to Ketchikan (Table 3). All other movement rates were fewer than 2 females per year (Table 3).

**Effect of initial unit definition on estimates of population structure** - In the Kodiak Archipelago and, to a lesser degree, in Prince William Sound a relatively even distribution of haul-outs made the definition of initial units difficult. To investigate the potential influence of the configuration of the initial units on the outcome of the cluster analysis, we tested a number of alternative configurations of the starting strata in these two areas (data not shown). While the locations of boundaries between the initial units were altered for the sensitivity trials, we attempted to maintain the spatial scales and sample sizes involved. Re-defining the initial units in Prince William Sound did not affect the overall outcome, i.e. in all trials Prince William Sound strata clustered together early on. In the Kodiak Archipelago the initial cluster sequence of the eastern and northern strata varied depending on the configuration of the initial strata. However, in all sensitivity trials western Kodiak remained distinct from the Trinity Islands (Tugidak and Sitkinak), as shown in our primary analyses (Figures 7-9).

### 3.2.2 Hypothesis testing

Substantial levels of population subdivision were detected among most of the original 16 units and were found to be significant using an $\alpha = 0.05$ under a null hypothesis of random mixing (Table 4). In many cases, significant differentiation was found among neighboring strata that indicated population subdivision over spatial scales of 153-541 km. For example, Ketchikan was significantly differentiated from Grand Island ($\chi^2 = 56.45$, $p = 0.029$). The distance by sea between these two units is 153 km. Similarly, Vixen-Sitka was differentiated from Frederick Sound ($\chi^2 = 59.53$, $p = 0.002$), the two areas separated by 174 km of inner waterways. The Pribilof Islands were differentiated from neighboring strata in northern and eastern Bristol Bay some 541 km distant. By contrast, a number of adjacent units were not found to be significantly differentiated. None of the three Prince William Sound strata could be distinguished from each other ($p > 0.1$), nor were they differentiated from the Copper River Delta and the eastern Kenai Peninsula. Similarly, the two Bristol Bay units were not found to be significantly differentiated ($\chi^2 = 21.45$, $p = 0.924$). The primary haul-out sites of these two areas are 226-270 km apart.
Figure 10. Graphic representation of the stage in the Boundary Rank analysis when all units were significantly differentiated from their neighbors at a chi-squared permutation p-value of < 0.1.
Table 3. Number of female seals moving per annum between units identified by Boundary Rank. Estimates were made by dividing the estimates generated by the program Migrate (Beerli and Felsenstein, 2001) by an average harbor seal generation time of 8 years. 95% likelihood intervals are given in parentheses.

<table>
<thead>
<tr>
<th>Dispersal rate from:</th>
<th>Number of dispersers per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frederick Sound to Ketchikan</td>
<td>4.25 (2.2 - 6.2)</td>
</tr>
<tr>
<td>Ketchikan to Frederick Sound</td>
<td>0.8 (0.7 - 1.0)</td>
</tr>
<tr>
<td>West Kodiak to Kamishak Bay</td>
<td>0.9 (0.5 - 1.1)</td>
</tr>
<tr>
<td>Kamishak Bay to West Kodiak</td>
<td>1.6 (0.6 - 1.8)</td>
</tr>
<tr>
<td>Prince William Sound to Kenai Peninsula</td>
<td>1.1 (0.9 - 1.2)</td>
</tr>
<tr>
<td>Kenai Peninsula to Prince William Sound</td>
<td>1.1 (0.9 - 1.2)</td>
</tr>
<tr>
<td>Sample Size (n)</td>
<td>Ketchikan</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>87</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
</tr>
</tbody>
</table>
4. DISCUSSION

4.1 *Population genetic structure in Alaskan harbor seals*

The analysis of mtDNA variation in harbor seals from throughout Alaska revealed a pattern of genetic differentiation that was correlated with geographic distance. This general isolation-by-distance pattern of heterogeneity (Wright, 1943; Slatkin, 1993), which extends throughout the entire North Pacific ocean (Westlake and O’Corry-Crowe, 2002), indicates that dispersal distances are a fraction of the species range and that when harbor seals do disperse, it is primarily to neighboring areas. Furthermore, the form of genetic differentiation, in conjunction with the non-uniform abundance of seals across the distributional continuum (Figure 4), indicates that Alaskan harbor seals are subdivided into a series of partially isolated sub-populations. This stepping stone pattern of population subdivision (Kimura, 1953; Kimura and Weiss, 1964) may occur on a number of spatial and temporal scales. We recently reported substantial levels of genetic differentiation in North Pacific harbor seals over spatial scales on the order of 600-800km that are indicative of long-term isolation of populations (Westlake and O’Corry-Crowe, 2002). The current analysis found levels of genetic differentiation within Alaska indicative of very low dispersal rates over distances on the order of 150-540km, depending on the region.

The BR analysis indicated close relationships among harbor seals in eastern and northern Bristol Bay and among seals throughout Prince William Sound and the Copper River Delta in the central Gulf of Alaska. The short genetic distances and early clusterings indicate that movements and dispersal may be extensive among these initial units. This relatively high degree of mixing is also suggested by the other cluster analyses and by our inability to reject the null hypothesis of random mixing (i.e., $p > 0.1$) under the hypothesis-testing approach.

More distant relationships, indicated by greater genetic distances and deeper nodes in the BR dendrogram, were observed among harbor seals in the central and southern inner waters of Southeast Alaska (Ketchikan and Frederick Sound, respectively), the Kenai Peninsula and Prince William Sound, and the Kodiak Archipelago and Kamishak Bay, indicating that dispersal among these areas is limited. This pattern was also evident in the UPGMA and N-J analyses and in most of these comparisons we were also able to reject the $H_0$ of random mixing ($p < 0.1$). Estimated dispersal rates between the Kenai Peninsula and Prince William Sound, Ketchikan and Frederick Sound, and Kodiak and Kamishak Bay did indicate that the level of dispersal between these areas is indeed low from a demographic perspective.

Finally, a number of the initial strata were found to be substantially differentiated from neighboring strata, not coalescing until very late in the BR clustering process. These areas include the Trinity Islands (Tugidak and Sitkinak) off Kodiak’s southern coast, the outer coast of central and southern Southeast Alaska (Vixen-Sitka and Grand Island, respectively), and the Pribilof Islands in the Bering Sea. Those areas are genetically so different as to indicate that dispersal between them and their nearest neighbors is extremely low and has likely been so for a period of time long enough for allele frequencies at neutral genetic markers such as mtDNA to diverge substantially. These
deep divisions are also present in the UPGMA cluster analyses and N-J tree, and in all of these cases the $H_0$ of random mixing was rejected with a high degree of statistical significance ($p < 0.001$).

4.2 Comparison of the various methods used to resolve population structure

As mentioned in the Introduction, resolving population structure in a continuously distributed species, such as Alaskan harbor seals, is difficult. The task is typically reduced to a more tractable one by comparing sets of samples collected at a number of geographically distant locations from across the continuum. While adequate perhaps, for describing broad-scale patterns of population structure, the informativeness of this approach is limited by the coarse a priori stratifications. In the current study we collected and analyzed samples from across much of the harbor seal’s Alaska range and used the geographically constrained cluster analysis, Boundary Rank (Martien and Taylor, in review) to allow the genetic data to reveal the underlying pattern of population subdivision by purposely starting with a series of small initial strata and successively grouping geographically adjacent strata based on genetic similarity. Unlike more traditional clustering methods this approach allows the researcher to incorporate prior knowledge about the general form of population structure into the analysis.

The UPGMA and N-J analyses, which grouped units based on genetic distance regardless of geographic location, gave similar results to the BR analysis. While most strata were quite distinct (i.e., long branches), these analyses indicated that a number were closely related (short branches), notably, the two Bristol Bay strata and the Prince William Sound and Copper River Delta strata. Significantly, these groupings involved geographically adjacent strata, further demonstrating the general isolation-by-distance pattern to population subdivision in this species.

Grouping areas solely on the basis of genetic distance, however, may link areas that are not close geographically. This is the case, for example, with Glacier and Kamishak Bays. Although these two strata were not among the first to group, both the UPGMA and N-J analyses eventually did link them together (Figures 8 and 9). A more extreme example involves western Kodiak and the Pribilof Islands. Although these two areas are quite distinct genetically, as indicated by the long branch lengths separating them in both the UPGMA and N-J trees and the highly significant differentiation recorded in the hypothesis test, they were eventually grouped together. In the case of UPGMA, as the analysis reached the last 4 groups these two units were found to have the shortest distance and were thus linked. Similarly, linking these two areas to each other rather than the two other remaining groups (i.e., Bristol Bay and Gulf of Alaska-Southeast Alaska) minimized the overall length of the N-J tree.

It is likely that these geographically disjunct groupings are simply an artifact of sampling bias due to low sample size. Inspection of the haplotypic composition of the Glacier and Kamishak Bay strata reveals that apart from the ubiquitous haplotype #6, these strata shared only 2 other haplotypes, #22 and #25. This, however, was enough to suggest a closer relationship of these two areas to each other than to any other unit or group. Similarly, although western Kodiak and the Pribilof Islands have few haplotypes in common, the frequencies of those that are shared are similar enough to suggest a
closer relationship to each other than either are to Bristol Bay or the rest of the Gulf of Alaska and Southeast Alaska. The geographic constraint in BR greatly reduced the number of pairwise comparisons made in determining which two units to coalesce at each step of the analysis. Consequently, it is less susceptible to being misled by poor sampling than the other clustering methods.

It is also possible that the geographically disjunct groupings suggested by UPGMA and N-J reflect some historical relationship between the distant strata that is still detectable today. In the absence of evidence to support such a historical relationship, we view this explanation as unlikely. Even if such a relationship did exist, while it might be of some academic interest, it would offer little direction to the manager, since the grouping of geographically distant strata may have little relevance to the contemporary relationships among areas and is not a practicable option from the perspective of defining management units. The BR analysis accommodates these two considerations when clustering strata by limiting the range of clustering options based on prior knowledge of the species' behavior and on management considerations (e.g., most logical to incorporate groups of contiguous sites into discrete management units). In the case of harbor seals in Alaska, restricting clustering to spatially adjacent units is based on the previously discovered isolation-by-distance pattern to population subdivision and provides a workable framework for the definition of biologically meaningful and geographically sensible management stocks (see below).

Statistical hypothesis-testing has been the traditional approach to analyzing subdivision in natural populations where the sound rejection of the null hypothesis of random mixing is taken as strong evidence of limited dispersal between areas (Nei, 1987; Hudson et al., 1992; Excoffier et al., 1992). The rejection of such a null hypothesis may reflect large differences in the haplotype frequencies among areas, which in itself can be taken as evidence of limited dispersal, and indeed the highly significant differences (low p-values) recorded among most of the strata in Southeast Alaska (Table 4) reflects widely differing haplotype frequency profiles among these areas (data not shown). This approach, however, suffers from a number of limitations that have caused many researchers to question its usefulness in an applied context (Johnson, 1999; Paetkau, 1999; Taylor and Dizon, 1999; Anderson et al., 2000; Martien and Taylor, in press). To begin with, the null hypothesis is rarely, if ever, true (Johnson, 1999; Anderson, 2000). Furthermore, the probability of detecting differentiation when it is present, called statistical power, is often low, making it very likely that hypothesis-testing will result in the definition of too few units in a management context (Martien and Taylor, in press). Another undesirable feature is that hypothesis-testing is not designed to place boundaries. A low p-value, therefore, indicates that structure is present but provides no statistical validity to any particular boundary. There is also no valid process to combine areas. For example, a scientist contemplating Table 4 would have no statistical guidance concerning combining areas in the Prince William Sound region.
4.3 The influence of sample size and the configuration of initial strata on estimates of population structure

Sample size can greatly influence the precision of estimates of heterogeneity and the power of statistical tests to detect genetic differentiation (Hudson et al., 1992; Taylor et al., 1997, Martien and Taylor, in press) and, therefore, potentially also influence the grouping sequence of units in distance-based cluster and phylogeny re-construction analyses such as BR, UPGMA and N-J. To minimize the effects of poor sample size, areas that had low adjusted sample size \( n \leq 4 \) were excluded from the analysis of population differentiation in Alaska harbor seals. Current sampling efforts are focused on increasing sample sizes in these areas.

The analysis of genetic structure may also be influenced by the size and configuration of the initial strata that are to be compared (Martien and Taylor, in press). If defined incorrectly, the underlying structure may be concealed and inappropriate inferences may be made about the relationships among areas, with ramifications for the use of genetic data in the identification of management units. To reduce the likelihood of making such errors we defined a set of initial strata that were small enough in area to minimize the risk of missing structure and yet large enough to minimize the effects of low sample size, and used cluster analysis to allow the genetic data to reveal the true pattern of subdivision. In a continuously distributed species such as Alaskan harbor seals, the appropriate size and shape of these initial stratifications may not be immediately obvious. Even when a set of criteria based on known patterns of behavior of the species is used to guide the process, configuring some strata may be partly subjective. The concern, therefore, is about the effects the initial stratification may have on the outcome of the analysis, especially if there is a subjective element in the definition process.

To illustrate, apart from the discrete grouping of seals in the Trinity Islands (Tugidak and Sitkinak) and in the narrow bays on the west coast of Kodiak Island, there are few obvious delineations of haulout sites in the Kodiak Archipelago that facilitate the definition of initial strata based on our criteria (Figure 6C). Combined with limitations of poor sample size, we thus decided to limit our analysis of population subdivision in the Archipelago to these two discrete areas. Even in areas where the distribution of seals in relation to typical movement distances (i.e. ≤50km) makes the delineation of starting strata fairly straightforward, such as in PWS, alternative configurations might be envisioned (Figure 6C). Sensitivity trials revealed that although it is clear that west and south Kodiak differ strongly, exact boundary placement within the archipelago remains problematic because of sampling considerations and difficulties associated with defining initial units in some areas. Thus, the current genetic data strongly support at least two stocks but data are inadequate to provide strong evidence for exact boundary placement. These trials also showed that the differentiation between PWS and neighboring areas is clear regardless of the configuration of the initial strata.
4.4 Comparison of genetic findings to other information of relevance to population structure

A number of aspects of harbor seal biology suggest that dispersal is limited and that this species is subdivided into discrete sub-populations. Geographic variation in morphology and reproductive physiology likely reflect limited dispersal and gene flow but may also be influenced by environmental factors. Tracking the movements of individual seals indicates strong site fidelity over relatively short distances, but current limits on the duration of the tracking period preclude the use of tagging in estimating dispersal. Regional differences in foraging ecology may reflect population subdivision, but could also occur in the face of extensive dispersal. Although differing trends in abundance can be consistent with different stocks, trend interpretation can be difficult without knowledge of population structure. The findings of the current study can aid in interpreting these and other aspects of harbor seal biology in the context of population structure and dispersal patterns. Here, we treat in some detail recent data on harbor seal movement patterns, trends in abundance and foraging ecology.

4.4.1 Movements

An increasing body of information on movement patterns of harbor seals in Alaska is becoming available. Although seals have not been tracked for long enough periods to confirm dispersal events, average and maximum distances traveled by telemetered seals allow us to make conservative estimates of likely dispersal distances, while individual movements often provide direct evidence of the relationships between areas.

In a satellite-tagging study of harbor seal movements in Prince William Sound (PWS), the majority of the 76 telemetered individuals remained within PWS (Lowry et al., 2001; Small et al., in prep.). A number of seals moved over 100km out of the Sound into the Gulf of Alaska, some moving 125km east to the Copper River Delta or 120 km south to Middleton Island. Many of these animals were recorded making return trips to PWS. Three juvenile seals made even longer movements of 300-500km away from their capture site, while a single pup made extensive movements southwest into the Gulf of Alaska, reaching Cape Douglas 350km away before returning to PWS. No seals, however, were recorded moving farther southwest than this pup or farther southeast than a single juvenile seal that visited Yakutat Bay. Although the majority of tagged seals were caught in south-central PWS, these findings suggest a close relationship among all seals within PWS, the Copper River Delta and Middleton Island and limited movement of these seals to other areas. Significantly perhaps, the three PWS strata were clustered together and with Copper River Delta early in the Boundary Rank and other cluster analyses of the genetic data. This entire region was significantly differentiated from all other areas.

In a study of seal movements in the Kodiak Archipelago, the majority of adult and sub-adult seals that were tracked remained within 30-50km of where they were tagged (Swain et al., 1996; Swain and Small, 1997; Small and Ver Hoef, 2001). Some made trips of 50-150km but usually returned. A small number of seals made longer
movements of 160-200km away from the capture site. In most of these cases, however, the seal was near the point of capture when transmissions ceased. By contrast, pups tagged on Tugidak Island made extensive movements out over the continental shelf and the submarine canyon in Shelikof Strait in their first year, some traveling over 400km from their capture site (Small et al., in prep.). Reviewing the movement data on all seals tagged in Kodiak between 1993 and 1999, however, revealed that few movements were recorded between northwestern Kodiak and the Trinity Islands (Tugidak and Sitkinak) or across Shelikof Strait to the Alaska Peninsula. No seals traveled north to the Kenai Peninsula or PWS. These findings are consistent with the genetic evidence of substantial differentiation within the Kodiak Archipelago as well as between Kodiak and Cook Inlet, Kenai and PWS.

In a companion study of harbor seal movements in Southeast Alaska, the scale and pattern of sub-adult and adult movements were similar to those of Kodiak seals (Swain et al., 1996; Swain and Small, 1997; Small and Ver Hoef, 2001). Seals were tagged in two sub-regions of central Southeast, and in only a few instances were seals recorded crossing Chatham Strait. Movements of seals tagged off the southeast coast of Admiralty Island were primarily restricted to the inner waters of Frederick Sound, Stephens Passage and Lynn Canal, whereas the movements of seals tagged off the south coast of Chichagof Island were limited to Hoonah Sound, Tenakee Inlet and the outer coast. No seals traveled as far south as Ketchikan or Prince of Wales Island or as far north as Glacier Bay. These findings are consistent with the genetic evidence of a break between the inner and outer coast at Chatham Strait (Frederick Sound v. Vixen-Sitka) and again at Prince of Whales Island (Ketchikan v. Grand Island) and with limited dispersal of seals between central Southeast Alaska (Frederick Sound and Vixen-Sitka) and southern Southeast (Ketchikan and Grand Island) or northern Southeast Alaska (Glacier Bay).

Preliminary findings from an ongoing study of harbor seal movements in Bristol Bay show seals making extensive movements along the coast and out over the continental shelf, some in excess of 300 km (G. Blundell, pers. comm). The scale of movements in this region may differ from that of other regions due to the presence of seasonal sea ice which forces seals in the northern half to move south each winter. This may also facilitate dispersal across long distances, a finding already indicated by the genetic analysis.

A number of tagging and telemetry studies in the north Atlantic have also found that although few physical barriers to movement exist and that individual seals may travel several hundred kilometers from haul-out sites, harbor seals typically range over much shorter distances there as well. Studies in the U.K. documented individual pups traveling over 100km, exceptionally >200km, away from their natal site (Bonner and Witham, 1974; Thompson et al., 1994). Most pups, however, remained in their natal area. Furthermore, movements of older seals between haul-out sites were on the order of 10-20km and the majority of movements from haul-out sites to at-sea locations were <50km (Thompson, 1989; 1993). In a study of long-term movement patterns of harbor seals in Sweden, individual seals were branded as pups and monitored for up to 14 years (Härkönen and Harding, 2001). None of the 163 branded seals were observed to haul out more than 32 km from where they were branded as pups.
In summary, while movement patterns of harbor seals differ both regionally and seasonally most likely in relation to the relative distance to foraging areas and suitable haul-out sites (Thompson, 1993), some generalizations can be made as to the spatial scale of movements in this species. Maximum distances traveled rarely exceed 200km and typically are much less. The majority of movements from haul-outs to at-sea locations are <50km, and if seals are monitored for long enough they are often observed to make the return trip. That the spatial scale and pattern of movements of harbor seals in four regions of Alaska are similar to the scale and configuration of population genetic structure found in those regions further confirms that seals are philopatric on scales of 150-540km, depending on the region.

4.4.2 Trends in abundance

Harbor seals in different regions of Alaska currently exhibit widely different trends in abundance (Frost et al., 1999; Matthews and Pendleton, 2000; Small et al., 2003). These differences occur on similar spatial scales to the genetic findings presented here and do not correspond to currently recognized stock structure (Angliss and Lodge, 2002).

In the Gulf of Alaska Stock, harbor seal numbers at terrestrial sites in Prince William Sound declined at a rate of -4.6%/year during the 1990s with an estimated total decline of 63% between 1984 and 1997 (Frost et al., 1999). By contrast, although parts of the Kodiak archipelago witnessed some of the most dramatic declines in recent decades (Pitcher, 1990), a trend route in east Kodiak has recorded a 6.6%/year increase since 1993 (Small et al., 2003).

Three separate trend routes within the Southeast Alaska Stock all exhibit different trends. Numbers in the Ketchikan area increased significantly at 7.4%/year between 1983 and 1998, while counts in the Sitka area were stable across a similar time period (Small et al., 2003). By contrast, seal numbers in Glacier Bay have been declining at a rate of -4.9% to -10.9%/year since 1992 (Mathews and Pendleton, 2000).

Finally, numbers in much of Bristol Bay, part of the Bering Sea Stock, appear to have stabilized in recent years after a period of possible decline (Small et al., 2003), while numbers in Nanvak Bay in northern Bristol Bay, also in the Bering Sea Stock, increased at a rate of 2.1%/year during the 1990s (Jemison et al., 2001). Interpretation of trends at this and other northerly haul-outs in the Bering Sea, is hampered by the co-occurrence of spotted seals (Phoca largha) at these sites, the juveniles, sub-adults and adults of which are difficult to distinguish from harbor seals.

It is possible that differences in trend may occur even with extensive dispersal among regions or may be driven, in part, by substantial redistribution of animals (Härkönen et al., 2002). Hoover-Miller et al. (2001) argue that declines of seals at terrestrial sites in Prince William Sound after the Exxon Valdez oil spill in 1989 can be explained by a redistribution of seals, while Jemison (1996) suggests that recent declines in harbor seals on Otter island in the Pribilof Islands may be due, in part, to displacement
by fur seals (*Callorhinus ursinus*). Small et al. (2003), however, cite the strong site fidelity observed in seal-tagging studies and conclude that dispersal among the different regions in Alaska is unlikely. More importantly, our genetic findings clearly indicate that these differences in trend are occurring in demographically independent sub-populations.

### 4.4.3 Foraging ecology

A variety of approaches have been and are currently being used to study the foraging ecology of harbor seals in Alaska. Scat and stomach content analysis document food habits, satellite and time-depth recorder telemetry of seals describe movement and diving behavior at sea and are used to make inferences about feeding, while fatty acid signatures of seals and their prey measure temporal and spatial foraging patterns. The genetic evidence can now facilitate the interpretation of the findings from these studies in terms of population structure and dispersal patterns.

Pitcher (1980) found significant differences in the stomach contents of seals in Prince William Sound and Kodiak. Iverson et al. (1997) recently found differences in fatty acid signatures between harbor seals in Prince William Sound, Kodiak and Southeast Alaska. The genetic findings indicate that these differences in diet and foraging ecology exist between sub-populations where dispersal, if it occurs, is at demographically low levels. The latter food-habits study also observed differences among sub-regions within PWS, and in some cases among haul-out sites within sub-regions (Iverson et al., 1999). No genetic differentiation has been found to date at these spatial scales, suggesting that dispersal is occurring at higher levels within PWS. Thus, seals from particular haul-outs or clusters of haul-outs tend to utilize traditional sub-areas for foraging within sub-populations. This is supported by the tagging studies which indicate that harbor seals predominantly forage within 50km of primary haul-out sites (Lowry et al., 2001; Frost et al., 2001).

### 4.5 Identification of stocks

The examination of patterns of variation within genetic markers is a long established method of analyzing subdivision and dispersal patterns in natural populations, and defining units of conservation (Dizon et al., 1992; Moritz, 1994). Variation within mtDNA is of particular utility as it predominantly reflects patterns of female dispersal over time and can thus be used to identify demographically independent populations and sub-populations which, depending on the management objectives, may require separate management (Avise, 1995; Taylor, 1997; Taylor and Dizon, 1999). Differences in haplotype composition among populations reflect the interplay between genetic drift, dispersal and mutation, and although rapid divergence is possible, differences between populations typically take a long time to accrue. For historically large populations like Alaska harbor seals it may take many generations for geographic differentiation to develop within mtDNA such that the genetic patterns we observe today may represent long-established patterns of female mediated site fidelity and dispersal.

The primary goals of the MMPA (Section 2.2) are to prevent marine mammals from diminishing below their optimal sustainable population and to maintain them as
functioning elements of their ecosystem. The latter goal has been interpreted to mean maintaining the species range. The inappropriate lumping together of two or more demographically distinct sub-populations of harbor seals into a single population stock risks local depletion and ultimately extirpation of sub-populations, thus failing to achieve the objectives of the MMPA as well as failing to preserve harbor seals as a sustainable resource for Alaska Native communities and other users, the stated goal of the Alaska Native Harbor Seal Commission, the primary Alaska Native Organization concerned with harbor seal management, and one of the guiding principles of the co-management agreement between the ANHSC and NMFS.

The substantial level of mtDNA differentiation found in Alaska harbor seals indicates that female dispersal occurs at demographically low rates (≤4.25 females/year) across spatial scales of 150-540 km, depending on the region. These findings indicate that the current stocks are too large and thus risk local depletion and failure to achieve the stated goals of management. The 12 demographically independent clusters identified in this study do not represent a complete picture of population subdivision and dispersal in Alaskan harbor seals as gaps remain in our sample coverage (see below). Nevertheless, they do reveal the scale, and in some cases the configuration, of population subdivision and show a high degree of concordance with other aspects of harbor seal biology pertinent to population subdivision. The genetic findings can thus serve as a framework for the definition of biologically meaningful management stocks of harbor seals in Alaska. They highlight the need for a re-appraisal of other information of relevance to stock identity, including variation in morphology and life history as well as an analysis of new information such as distribution, movement patterns, trends in abundance, foraging ecology and traditional ecological knowledge (TEK) in the context of improving stock identity. Final stock designation requires a synthesis of all these kinds of information.

4.6 Future directions

The genetic findings are consistent with other information relating to harbor seal dispersal patterns and population structure and indicate the spatial scales at which this species should be managed. The analyses however, are still limited by sample coverage. Substantial sample gaps still exist in areas of high conservation concern. In particular, the Aleutian Islands, the Alaska Peninsula, the northeastern Gulf of Alaska (Yakutat, Icy Bays) and parts of Southeast Alaska and the Kodiak Archipelago (see uncircled ares in Figure 10). Many of these gaps are areas with substantial subsistence harvest where active collaboration with the Alaska Native Harbor Seal Commission and subsistence hunters would be greatly beneficial. In other areas directed sampling is required. As sample coverage increases we expect the configuration of some existing strata to change somewhat, new strata to be added and the connectivity matrix modified accordingly.

Although we have used the best genetic techniques and analytical methods available to estimate population structure, no method can fill in knowledge gaps left be inadequate sampling. Thus, although this research should lead to significant improvement in understanding harbor seal population structure in Alaska, appropriate management will still require bridging knowledge gaps with expert opinion awaiting further data.
ACKNOWLEDGMENTS

Funding for this research was provided by the National Marine Fisheries Service (NOAA Fisheries) and the Alaska Department of Fish and Game’s harbor seal grant from the National Oceanic and Atmospheric Administration. A special words of thanks to Robin L. Westlake who conducted much of the early laboratory aspect of this study. Invaluable laboratory assistance was provided by L. Hansen, K. Raschen, C. Reed, A.S. Costa, C. Lux, and D. Campbell. Thanks are due to all those, many anonymous, who went to the trouble of collecting and sending us samples, including many Alaska Native hunters, J. Cesarone, L. Dzinich, K. Frost, G. Jarrell, L. Jemison, M. Kookesh, J. Lewis, L. Lowry, B. Mahoney, E. Mathews, K. Raum-Suryan, M. Riedel, R. Small, U. Swain, V. Vanek, D. Withrow, and K. Wynne, the Alaska Department of Fish and Game, the Alaska Native Harbor Seal Commission, the University of Alaska Museum, and the National Marine Fisheries Service. We appreciate the advice and support we received from A. Dizon, L. Lowry, K. Brix, J. Bengtson, M. Payne, J. Burns and D. DeMaster. We would like to thank M. Adkison, A. Dizon, L. Lowry, P. Morin and W. Perrin for reviewing earlier drafts of this report, and R. LeDuc and E. Archer for helpful discussions on many aspects of this study. Finally, a special word of gratitude to Rich Cosgrove for his tireless work on the GIS analysis of data and the production of the maps in this report.


Kumar, S., K. Tamura, I.B. Jakobsen, and M. Nei (2001) MEGA2: Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, Arizona, USA.


Taylor, B.L. and A. E. Dizon. 1999. First policy then science: why a management unit based solely on genetic criteria can’t work. Molecular Ecology 8: S11-S16


Reports, Agreements and Unpublished Manuscripts


Use of genetic markers to detect population structure and estimate dispersal

For many reasons, the direct measurement of rates and modes of dispersal in harbor seals (i.e., using mark-recapture, telemetry) is difficult. Harbor seals are long-lived and so individual animals need to be monitored for several years in order to confirm dispersal events. Furthermore, large numbers of seals would have to be tracked in order to get accurate estimates of dispersal rates. Fortunately there are a number of indirect methods that use naturally occurring traits as ‘tags’ to estimate dispersal and describe breeding patterns. To be informative, these traits need to be heritable (that is, passed from parent to offspring), variable, and be under little or no selective pressure. Many sites along the DNA sequence of each individual possess these properties such that geographic patterns of variation at these genetic sites reflect patterns of dispersal and breeding behavior (see Box 1 below). Modern molecular genetic techniques enable us to screen many individuals for variation at these ‘genetic markers’ and then use the geographic patterns of variation to estimate levels of dispersal and interbreeding.

When resolving the demographic relationships among groups of seals, the dispersal patterns of females are of greatest interest. In many vertebrate species the degree to which females disperse among areas is the primary determinant of the rate at which a depleted area will be replenished by recruits from outside or whether a vacant area will be (re)colonized (Moritz, 1994; Avise, 1995). By contrast, the dispersal patterns of males, while having an important influence on the extent of genetic exchange, typically contribute little to the demographic trajectories of natural groupings. Thus, where female dispersal among two groups of seals is low, these groups represent demographically distinct entities, even in the face of high male dispersal (Avise, 1995), and, depending on the management objective, may need to be managed as separate ‘population stocks’ or management units.

Because of its maternal mode of inheritance (Hutchison et al., 1974; Ingman et al., 2000), the pattern of variation at one particular genetic marker, mitochondrial DNA (mtDNA), is influenced primarily by the dispersal patterns and breeding behavior of females. This, combined with its high level of variability (Brown et al., 1979; Denver et al., 2000), makes mtDNA an ideal marker to investigate the demographic relationships among natural groupings of animals, such as harbor seals.

Furthermore, this marker offers two advantages over other genetic markers when detecting genetic differentiation. In many species of mammals, both effective (interbreeding) and actual (emigration) dispersal is biased towards males (Greenwood, 1980), such that genetic differentiation at bi-parentally inherited nuclear markers (e.g., microsatellites) will be lower and more difficult to detect than at the maternally inherited mtDNA. Secondly, the effective population size, $N_e$, for mtDNA is typically much lower than that of nuclear markers (Birky et al., 1983), thus the rate of genetic drift is much greater, again resulting in increased power to detect genetic differentiation (Moritz et al., 1994).
The pattern of variation at selectively neutral markers within populations is influenced by three forces: gene flow, mutation and random genetic drift (Wright, 1969; Hartl and Clark, 1989). Genetic drift is the process where allele or haplotype frequencies change through time because of the random nature of inheritance. The rate at which the genetic composition of a population changes, and thus the rate at which two populations diverge, due to drift is inversely proportional to effective population size, $N_e$. The larger the population, the slower the rate of genetic drift. Mutation ($\mu$) is a change in the genetic code of a haplotype due primarily to errors in the replication process that produces another, often new, haplotype. The rate at which these changes occur is typically so low that mutation may have a limited influence on the genetic composition of populations over ecological time scales. Finally, gene flow, the amount of genetic exchange among populations, is determined by the rate of dispersal ($m$, $d$) between populations. The higher the rate of dispersal, the lower the level of differentiation.

Wright's (1943, 1951) idealized island model of population structure demonstrates the relationship between these forces and the level of genetic divergence, measured as $F_{st}$ at nuclear markers for populations at equilibrium. Takahata and Palumbi (1985) modified this model for uni-parentally inherited genetic markers such as mtDNA:

$$F_{st} = \frac{1}{2N_e(m+\mu)} + 1 = \frac{1}{2N_e(dT+\mu)} + 1$$

where $N_e$ = effective number of females in the population, $m$ = dispersal rate per generation, $\mu$ = mutation rate, $d$ = the annual dispersal rate and $T$ = generation time. In statistical terms $F_{st}$ is the effect size. The figure depicts the relationship between the effect size and dispersal and shows that the effect size for a given dispersal rate is greatly reduced as abundance increases (because drift is greatly reduced).
Marine Mammal Protection Act stock definition guidelines

Participants in a workshop to provide guidelines for implementing the MMPA recognized that maintaining the range would serve to meet the ecosystem goals (Wade and Angliss, 1997). The stock definition section of this report states:

Many types of information can be used to identify stocks of a species: distribution and movements, population trends, morphological differences, genetic differences, contaminants and natural isotope loads, parasite differences, and oceanographic habitat differences. Evidence of morphological or genetic differences in animals from different geographic regions indicates that these populations are reproductively isolated. Reproductive isolation is proof of demographic isolation, and thus separate management is appropriate when such differences are found. Failure to detect differences experimentally, however, does not mean the opposite. Dispersal rates, though sufficiently high to homogenize morphological or genetic differences detectable experimentally between putative populations, may still be insufficient to deliver enough recruits from an unexploited population (source) to an adjacent exploited population (sink) so that the latter remains a functioning element of its ecosystem. Insufficient dispersal between populations where one bears the brunt of exploitation coupled with their inappropriate pooling for management could easily result in failure to meet MMPA objectives. For example, it is common to have human-caused mortality restricted to a portion of a species' range. Such concentrated mortality (if of a large magnitude) could lead to population fragmentation, a reduction in range, or even the loss of undetected populations, and would only be mitigated by high immigration rates from adjacent areas.

Therefore, careful consideration needs to be given to how stocks are defined. In particular, where mortality is greater than a PBR calculated from the abundance just within the oceanographic region where the human-caused mortality occurs, serious consideration should be given to defining an appropriate management unit in this region. In the absence of adequate information on stock structure and fisheries mortality, a species' range within an ocean should be divided into stocks that represent defensible management units. Examples of such management units include distinct oceanographic regions, semi-isolated habitat areas, and areas of higher density of the species that are separated by relatively lower density areas. Such areas have often been found to represent true biological stocks where sufficient information is available. There is no intent to define stocks that are clearly too small to represent demographically isolated biological populations, but it is noted that for some species genetic and other biological information has confirmed the likely existence of stocks of relatively small spatial scale, such as within Puget Sound, WA, the Gulf of Maine, or Cook Inlet, AK.
Appendix 3

Sample size \((n)\), haplotypic diversity \((H)\), and adjusted sample size \((n_a)\)

The size of sample (i.e., number of individuals) needed to achieve adequate performance in any genetic analysis will depend on the haplotypic diversity \((H)\) of the study species. This is particularly true for frequency-based analyses, where the genetic differentiation between two units is estimated by comparing their haplotype frequencies. The higher \(H\), the larger the sample size \((n)\) that will be needed in order to accurately estimate haplotypic frequencies. When \(H\) is very high (i.e., close to 1) and sample size is low, many haplotypes may be represented in the sample by only a single individual. These haplotypes contribute no information to a frequency-based analysis of genetic differentiation. If there are a large number of unique haplotypes in a comparison of two units then the estimate of genetic distance between them will be negatively biased and the \(p\)-value for the comparison will be positively biased. For illustrative purposes, consider the extreme example where ten samples are collected from each of two sampling sites, and each sample is found to have a different haplotype. The level of genetic differentiation, \(F_{st}\), (i.e., the proportion of the total variance that is due to variance between the two samples, Wright, 1943, 1951) is zero and the associated \(p\)-value is one, despite the fact that the two samples have completely non-overlapping sets of haplotypes. Thus, inadequate sampling relative to diversity can result in a bias that causes poorly sampled sites to appear to be more genetically similar (less differentiated) than they actually are. Because the three clustering methods used in this study combine units on the basis of their estimated genetic similarity, a small sample size relative to \(H\) can result in two units coalescing early in the clustering process, not because they are very similar genetically, but simply because they both are inadequately sampled.

In order to gauge whether the estimate of genetic similarity between two initial units is likely to be biased due to inadequate sampling relative to \(H\), we developed a diagnostic statistic we call the adjusted sample size \((n_a)\). The adjusted sample size is calculated by subtracting the number of haplotypes detected at a sampling site from the number of samples at that site, \(n\). For instance, we sequenced 16 samples from the Pribilofs and detected 10 distinct haplotypes, resulting in an adjusted sample size of \(n_a = 16 - 10 = 6\). From Icy Bay, 20 samples yielded 18 haplotypes, so \(n_a = 20 - 18 = 2\). Thus, the adjusted sample size is the number of samples collected at a site that represented a haplotype that had already been observed at that site at least once. Because it does not contribute any information to frequency-based comparisons, a sample that represents a unique haplotype at the sampling site is not counted.