Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal

A. RUS HOELZEL,* R. C. FLEISCHER,* †C. CAMPAGNA, ‡B. J. LE BOEUF§ & G. ALVORD¶

*Department of Biological Sciences, Durham University, South Road, Durham DH1 3LE, UK
†Department of Zoological Research, National Zoological Park, Smithsonian Institution, Washington, DC 20 008, USA
‡Centro Nacional Patagónico, CONICET, 9120 Puerto Madryn, Chubut, Argentina
§Department of Biology, University of California, Santa Cruz, CA 95 064, USA
¶Data Management Services Inc., National Cancer Institute, FCRDC, Frederick, MD 21 702, USA

Keywords:
bottleneck;
fluctuating asymmetry;
genetic diversity;
marine mammal.

Abstract

The northern elephant seal (NES) suffered a severe population bottleneck towards the end of the nineteenth century. Theoretical expectations for the impact of population bottlenecks include the loss of genetic diversity and a loss of fitness (e.g. through a disruption of developmental stability); however, there are few direct demonstrations in natural populations. Here, we report on the comparison of archive samples collected prior to and following the NES population bottleneck. Measures of genetic diversity show a loss of variation consistent with expectations and suggest a strong disruption in the pattern of allele frequencies following the bottleneck. Measures of bilateral characters show an increase in fluctuating asymmetry.

Introduction

The impact of population bottlenecks on the diversity and fitness of natural populations is an important question in the context of both evolutionary process and conservation biology. The theoretical expectation with respect to genetic diversity can be calculated based on the stochastic loss of alleles, and the probability of recombining alleles that are identical by descent (e.g. Nei et al., 1975). Experimental studies have tended to support theoretical expectations, although they also illustrate the large variance in outcome among different bottleneck events, and show that the impact on allelic diversity is more predictable than the impact on heterozygosity (e.g. Leberg, 1992). However, most studies assessing the impact of known and putative bottleneck events in natural populations are based on comparative analyses, either between species or between conspecific populations. Direct comparisons of the same population before and after a bottleneck event in one of a number of extant populations of a species (Britten & Brussard, 1996; Bouzat et al., 1998; Groombridge et al. 2000; Whitehouse & Harley 2001). However, it is often difficult to control for the possible influence of immigration, which can have a substantial effect on the rate of recovery of genetic variation in post-bottleneck populations (e.g. Ratnaswamy et al., 1999; Keller et al. 2001).

In this study, we compare pre- and post-bottleneck populations in a species [the northern elephant seal (NES), Mirounga angustirostris] where only one population was left after the bottleneck, so that any post-bottleneck diversity must have survived the bottleneck event. In addition to the direct genetic comparison, we also analyse skulls from the prebottleneck sample for measures of bilateral symmetry, and compare them with measurements from post-bottleneck skulls and from a congeneric species, the southern elephant seal (SES) (Mirounga leonina). In this way, we test the prediction that a severe bottleneck can impact on developmental stability, as indicated by a disruption of fluctuating asymmetry (FA; Waddington, 1957; Leamy, 1984; Bryant et al., 1986; Hutchison & Cheverud, 1995). It has long been considered that FA could be used as an indicator of fitness related to developmental stability (Beardmore, 1960; Soulé, 1967; Parsons, 1992). Among

Correspondence: A. Rus Hoelzel, Department of Biological Sciences, Durham University, South Road, Durham, DH1 3LE, UK.
Tel.: +44-(0)191 3742417; e-mail: a.r.hoelzel@durham.ac.uk
the three types of asymmetry (directional, fluctuating and antisymmetry) only FA has been suggested to result from poorly coadapted gene complexes, and therefore to be useful as a measure of developmental stability. This builds on an earlier analysis of NES skulls (Hoelzel, 1999) that showed greater variability of morphometric characters in post than pre bottleneck skulls, consistent with earlier experiments (primarily on house flies) and predictions related to the disruption of non–additive genetic interactions following bottleneck events (see Bryant et al., 1986).

The NES was heavily exploited during the nineteenth century and reduced to a bottleneck population size estimated to be only 10–30 individuals (Bartholomew & Hubbs, 1960; Hoelzel et al., 1993). The hunt began around 1810 and became commercially unviable by the 1860s as a result of overexploitation. However, NES continued to be taken intermittently, including over 400 between 1880 and 1884 (Townsend, 1885). The population has now recovered to more than 100 000 seals (Stewart et al., 1994). This provides the basis for a direct assessment of the impact of a severe bottleneck on genetic diversity and a measure of developmental stability in a natural population.

Molecular genetic variation in this species is relatively low at mtDNA (Hoelzel et al., 1993; Weber et al. 2000), allozyme (Bonnell & Selander, 1972; Hoelzel et al., 1993) MHC (Hoelzel et al., 1999b) microsatellite and minisatellite DNA loci (Hoelzel, 1994; Hoelzel et al., 1999a), consistent with predictions, given the severity of the bottleneck (Hoelzel et al., 1993; Halley & Hoelzel, 1996; Hoelzel, 1999). As for a number of species (for marine mammals, see review in Hoelzel et al., 2002), interspecific comparisons suggest the loss of variation caused by a bottleneck event, however, a direct comparison of the same population before and after such an event is the only way to fully quantify and assess the impact. Weber et al. (2000) compared mtDNA diversity between modern NES samples and five prebottleneck samples, and found novel alleles in the prebottleneck sample (although the sample size was too small to provide an assessment of genetic diversity or statistical comparisons of pre- and post-bottleneck diversity levels). We expand on these data adding a further 16 prebottleneck samples and four new loci (microsatellite DNA loci).

A key element of this comparison is the assumption that the pre- and post-bottleneck samples are from the same source population. The historical data strongly indicate that this is the case. Around the time of the bottleneck (1860–1890), NES was known only from Guadalupe Island, San Benito Island, Cedros Island and on the mainland at Bahia San Cristobal (see Fig. 1). These sites are all within a radius of 200 km. There is no evidence from historical or archaeological data to indicate that NES was ever found further south than Baja California, although archaeological data suggest that they were in California 15 000 years ago (Walker & Craig, 1979). By the 1880s they were only known to be at Guadalupe Island and Bahia San Cristobal (on the Baja peninsula), and as far as anyone at the time could determine, all had been killed by 1884 (Scammon, 1870; Townsend, 1885).

The re-growth of the population began on Guadalupe with less than 10 seen there in 1892 and 1904, but 40 in 1907 and 125 in 1911 (Townsend, 1912). By 1918 a small number of elephant seals were seen on the relatively nearby Island San Benito, where they began breeding in the 1930s. The breeding population on Guadalupe Island continued to grow to 264 seals in 1922 (Anthony, 1924), and 469 in 1929 (Huey, 1930). San Miguel Island was colonized in 1925 and the population began breeding there in the early 1950s. Los Coronados and Santa Barbara islands were colonized in 1948, and breeding began in the 1950s (Bartholomew & Boolootian, 1960; Odell, 1974). Año Nuevo Island was colonized in 1955, and the, population, began, breeding, in, 1961 (Radford, et, al., 1965).

By 1960, the total population was estimated to be 15 000 with 91% on Guadalupe Island, 8% at San

Fig. 1 Map illustrating current range of the northern elephant seal, and the historical locations of the post-bottleneck population as it expanded out from the refuge on Guadalupe. Dates indicate first records of new colonies at specific sites.
Benito Island, and 1% on other islands further north (Bartholomew & Hubbs, 1960). The expansion continued onto the Farallones in the 1970s (Le Boeuf et al., 1974). The clear impression is of a remnant population on Guadalupe Island that then expanded over the last 110 years into the current species range, with sites near Guadalupe being colonized first (Fig. 1). All of the prebottleneck samples used in our study came from Guadalupe or sites relatively nearby (within approximately 200 km). Our post-bottleneck sample includes animals from Año Nuevo and the Channel Islands (Hoelzel et al., 1993; Weber et al. 2000). Data on post-bottleneck patterns of dispersal demonstrate that there is frequent movement between modern rookeries (about 30% dispersal per annum based on tag recoveries), and that females breed at newly adopted rookeries (see Le Boeuf et al. 1974; Le Boeuf, 1981). A comparison between samples from 109 modern animals from southern California (Channel Islands; Weber et al., 2000) and 40 modern animals from Año Nuevo in central California (Hoelzel et al., 1993) showed almost identical mtDNA allele frequencies for the two haplotypes (same most frequent allele at 71.6 and 72.5%, respectively), suggesting no genetic structure over this geographical range.

**Methods**

**Genetics**

Mitochondrial DNA was amplified from 11 nineteenth century skulls (Smithsonian collection numbers 38 234, 21 887, 20 927, 38 208, 21 889, 38 285, 21 891, 14 929, 21 738, 21 890, 21 896), of which eight were collected between 1883 and 1884, three in 1892. DNA was also extracted from five bone samples from midden sites on San Miguel Island, dating from the fifteenth to seventeenth centuries (Walker et al., 2000). All modern samples were collected in the 1990s from both the southern and northern regions of the present-day range. DNA from modern samples was extracted from skin using standard protocols (see Hoelzel et al., 1993 for details of collection and extraction methods). Thirty post-bottleneck samples were sequenced for the HVR1 segment of the control region locus was either for 300 bp of the HRV1 region as described in Hoelzel et al. (1993), or for a 116-bp segment designed around the variable sites, using the primers: sense- GCCCTATGTATTATCGTGC; antisense- GGCTCGTTGGACTAGGTG in assay conditions: 2 mM MgCl2, 2 mM NTP, 2.4 mg mL⁻¹ BSA, 1 µM each primer, 1 U Taq Gold (hot start); cycled 45 times at 92 °C – 40 s, 50 °C – 1 min, 72 °C – 1 min, and product sequenced on ABI automated system. Microsatellite DNA loci NEMS11a, cloned from an NES genomic library (Hoelzel et al., 1999a) and Hg4.2, Hg8.9 and Hg8.10 cloned from a Halichoerus grypus library (Allen et al., 1995) were amplified from 100 post-bottleneck NES, and 4–14 nineteenth century NES samples. Locus NEMS11a used primers: sense – TACATTCAAGGCTCAA; antisense – TGGTTCCAGTTTACCA and all microsatellite loci used the same assay conditions and an annealing temperature of 50 °C (Hg primers as given in Allen et al., 1995). Amplifications were attempted from nineteenth century samples at least six times per locus.

Compliance with the expectations of the Hardy–Weinberg equilibrium and allele frequency differences were tested using Fisher exact tests, and the probability estimated using a Markov chain permutation analyses. Haplotypic diversities were compared using Welch’s approximate t-test, which corrects for unequal sample size and variance (Welch, 1938; Sokal & Rohlf, 1995). Average heterozygosity was compared using the Mann–Whitney U-test. Populations that experienced a recent reduction in effective population size will show a correlated reduction in allele number and gene diversity. However, allele number reduces faster than gene diversity, so in a recently bottlenecked population the gene diversity will be higher than expected at equilibrium. The program BOTTLENECK ( Cornuet & Luikart, 1996) was used to test for evidence of a recent bottleneck (using the microsatellite DNA data) on the basis of this theoretical expectation. The two-phase model was applied (with 70% stepwise), as this has been shown to be the most appropriate for microsatellite DNA data (DiRienzo et al., 1994). Demographic simulation analyses were after Hoelzel et al. (1993), and used the same age-specific reproduction and mortality data.

**Morphometrics**

The symmetry of bilateral characters was tested using four measurements from the mandibular region of 11 prebottleneck NES skulls, 43 post-bottleneck NES skulls and 32 SES skulls (Fig. 1). The absolute value of left minus right side was used as a measure of the magnitude of asymmetry (index 1 in Palmer & Strobeck, 1986). All skulls were measured by the same person (ARH) using a precision dial caliper. Only female skulls were used as a result of sexual dimorphism and the paucity of male skulls. Measurements indicated in Fig. 1 are as follows: mandibular height (MH); mandibular length (ML);
mandibular tooth row (MTR) and upper tooth row (UTR). Measurements were repeated, but for the level of resolution chosen (±1 mm for ML, ±0.1 mm for MH, UTR and MTR) no measurement error was detected. Tests for bias between right and left sides used the Wilcoxon statistic, and a one-sample t-test against a mean of zero. Soule (1967) points out that ‘asymmetry is fluctuating if the signed differences between paired structures are normally distributed with a mean of zero’. Distributions were tested for normality using the Kolmogorov–Smirnov test. The Mann–Whitney U statistic was used to test for the difference between the magnitude of asymmetry as measured, for each skull, by the absolute value of left minus right sides. We also used linear regressions of left vs. right sides to assess asymmetry, comparing residual means squared among regressions using an F-statistic. Data were compared with and without log transformation, and showed the same pattern of statistical significance either way (Fig. 2).

**Results**

**Genetic variation**

There were only two haplotypes among the 185 postbottleneck sequences for 300 bp from the 5’ mtDNA control region (115 from Weber et al. 2000; 40 from Hoelzel et al., 1993; 30 from this study), and seven haplotypes among 22 samples sequenced for 116 bp [including two novel haplotypes from six samples (five of which are prebottleneck) from Weber et al. 2000 sequenced for 179 bp incorporating the 116 bp sequence] collected in the nineteenth century or before (Table 1). All haplotypes from nineteenth and twentieth century
samples could be defined by the nucleotides at three positions (196, 214 and 224 from Hoelzel et al., 1993). Two further variable sites were found only in the midden samples (positions 225 and 235 from Hoelzel et al., 1993).

All four skulls from 1892 (including one from Weber et al., 2000 and one analysed in both studies) had the haplotype GTA (one of two modern sample haplotypes). The AAG haplotype shows a significantly different frequency in the post-bottleneck (lower 95% confidence estimate \(= 68.5\%\)) compared with the prebottleneck sample (upper 95% confidence estimate \(= 13.3\%\)).

Haplotypic diversity (Nei & Tajima, 1981) was significantly greater in the prebottleneck sample (from 1884 and before; \(h = 0.804 \pm SE 0.069\)) than in the post-bottleneck sample (\(h = 0.41 \pm 0.028; t' = 5.28, P < 0.001\)). This difference is still significant when the 1892 samples are counted among the prebottleneck samples (\(t' = 5.69, P < 0.001\)). An exact test comparing the haplotype frequencies shown in Table 1 was also highly significant (\(P < 0.0001\)).

In an earlier study, the size of the putative bottleneck had been back-calculated using a monte carlo simulation model (500 reiterations with random seeds; Hoelzel et al., 1993), based on life history data collected over 25 years for NES, and an estimate of prebottleneck genetic variation based on the congeneric SES. We repeated this analysis using all the same parameters as in Hoelzel et al. (1993), but using the estimate of prebottleneck mtDNA diversity from the data presented in Table 1. The model bottleneck event is imposed instantaneously, and the prebottleneck parameters are based on a long-term study of NES (see Le Boeuf & Reiter, 1988). A census in 1960 provided a population estimate of 15 000 seals at that time, and the modern level of genetic diversity at this locus is two haplotypes (Table 1). Therefore, the simulation is run for 76 years, and the number of simulations

### Table 1

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Control region GTA</th>
<th>GAA</th>
<th>GAG</th>
<th>AAG</th>
<th>AAA</th>
<th>GAAGC</th>
<th>GAGGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modern</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>135</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nineteenth century</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pre-eighteenth century</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Table 2

For putative bottlenecks of 5–20 seals, results are presented for the number of simulation runs that resulted in 15 000 seals and two haplotypes at the end of the run (outcome match), the number of runs that did not go to extinction (out of 500), and the mean final population size and haplotype number at the end of the run. Details for the most likely event are shown in bold.

<table>
<thead>
<tr>
<th>Bottleneck size</th>
<th>Outcome match</th>
<th>Surviving runs</th>
<th>End population size (SD)</th>
<th>End Haplotype number (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3</td>
<td>170</td>
<td>5367 (5597)</td>
<td>1.33 (0.48)</td>
</tr>
<tr>
<td>[2]</td>
<td>[128]</td>
<td>[5354 (5907)]</td>
<td>[1.41 (0.58)]</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>423</td>
<td>15 195 (12 067)</td>
<td>1.97 (0.72)</td>
</tr>
<tr>
<td>[8]</td>
<td>[408]</td>
<td>[14 963 (11 423)]</td>
<td>[2.44 (1.09)]</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>493</td>
<td>27 277 (16 281)</td>
<td>3.18 (1.08)</td>
</tr>
<tr>
<td>[4]</td>
<td>[485]</td>
<td>[27 867 (16 139)]</td>
<td>[3.88 (1.54)]</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>500</td>
<td>40 386 (19 493)</td>
<td>4.17 (1.14)</td>
</tr>
<tr>
<td>[2]</td>
<td>[496]</td>
<td>[41 040 (19 643)]</td>
<td>[5.34 (1.83)]</td>
<td></td>
</tr>
</tbody>
</table>

Values in brackets are from Hoelzel et al. (1993).
that reach 15 000 (±1000) and two haplotypes simulta-
neously are noted to provide a probability distribution for
different putative bottleneck sizes (Table 2). Although
the standard deviations are high for population size and
haplotype number estimates (due to demographic sto-
chasticity), the difference in the genetic and demographic
distributions permits some resolution (see Hoelzel
et al., 1996), although the power of the test is low
with only four loci. Both the sign test ($P = 0.04$) and
Wilcoxon test ($P = 0.03$) show a significant excess for all
loci, indicative of a recent bottleneck.

**Morphometric analyses**

Measures of $|L–R|$ are given for all comparisons in
Table 3. The distributions of $L–R$ for all NES measure-
ments were not significantly divergent from normal
distributions, and the means for NES and SES distribu-
tions were not significantly different from zero, indicat-
ing that the observed asymmetry is FA. Some SES
distributions were leptokurtic, but this does not imply
antisymmetry, and therefore does not alter interpreta-
tions regarding FA. One measurement (MTR) showed
significantly greater FA in post compared with prebot-
tleneck NES samples, and three of the four measures
(including MTR) showed significantly greater FA in post-
bottleneck NES compared with SES samples (Table 3).
There was no significant difference between the prebot-
tleneck NES and SES samples at MTR, or for any other
characters based on the non-parametric tests (see
Table 3). The comparison for MTR is illustrated in
Fig. 2 using linear regressions of right vs. left sides. The
robustness of the effect for MTR was tested by comparing
equal sized sample sets for pre- and post-bottleneck skulls
(every fourth sample was included for the post-bottle-
neck sample-set three times for three independent tests),
and the post-bottleneck increase in FA was still significant
($z = 2–3, P = 0.015–0.009$).

**Discussion**

These data show a strong increase in FA for one
morphometric measure (MTR) in post-bottleneck NES

Table 3 Fluctuating asymmetry of cranial measures comparing prebottleneck NES (bBN), post-bottleneck NES (aBN) and SES skulls.

| Measure | Range of $|L–R|$ | bBN vs. aBN $z$ | $F$ | SES vs. aBN $z$ | $F$ | SES vs. bBN $z$ | $F$ |
|---------|----------------|----------------|-----|----------------|-----|----------------|-----|
| UTR     | 0–0.03         | 0.79 1.75      | *** | 4.97 8.78      | 1.81 5.02 |
|         | 0.3–0.62       | 3.09 6.73      | *** | 5.23 8.93      | 0.50 1.32 |
| MTR     | 0–0.04         | 0.63 1.83      | *** | 2.25 1.47      | 2.08 1.11 |
|         | 0–0.5          | 0.49 1.10      |     |                |     |
| ML      | 0–0.5          | 0.21 1.53      |     |                |     |
| MH      | 0–0.4          | 0.28           |     |                |     |

$**P < 0.005, ***P < 0.0001.$

Data for the range of absolute values of left–right sides ($|L–R|$) is shown in centimetres without transformation. UTR = upper tooth row, MTR = mandibular tooth row, ML = length of the mandible, MH = height of the mandible. Mann–Whitney $U$-test ($z$) comparing $|L–R|$ and an $F$-statistic ($F$) comparing RMS are given.
compared with either prebottleneck NES or SES, and the direct loss of genotypes at all five loci investigated. The direct comparison of historical samples indicates that this is the result of a bottleneck that most probably occurred when over 400 seals were taken from an already depleted population between 1880 and 1884. Although the pattern of allele frequencies in the prebottleneck sample cannot be accurately assessed at all loci because of small sample sizes, the striking difference at some loci in the pre- and post-bottleneck samples reflects expectations about the stochastic disruption of allele frequencies caused by sampling effects following bottleneck events (cf. Luikart et al., 1998). The absence in the prebottleneck sample of a post-bottleneck allele at one of the five loci (NESM11a) most probably reflects this type of allele frequency distortion. An incomplete representation of alleles as a result of prebottleneck sample size is also a factor. For example, for a sample size of 14 (as for NESM11a) the allele would need to be present at 20% to give 95% confidence of detection (based on a binomial distribution).

The direct loss of alleles and the significant change in measures of diversity are the primary result from the molecular genetic data. However, tests against equilibrium expectations are also consistent with the bottleneck interpretation, although a larger number of loci should be included for the BOTTLENECK program for a more thorough test. Earlier work based on simulation models showed the consistency of molecular data for the modern population with the bottleneck interpretation (Hoelzel et al., 1993; Hoelzel, 1999). Using the prebottleneck mtDNA data from this study instead of the estimate based on SES (as presented in Hoelzel et al., 1993) reinforced the earlier interpretation. In Hoelzel et al. (1993) the most probable bottleneck according to the simulation was smaller that 20 individuals, and this was still true using the actual prebottleneck data.

Heritability of FA is typically low (e.g. Leamy, 1999), and we do not have the necessary data to assess the genetic component of the effect observed. Further, because of incomplete knowledge about the independence of the measured characters and the developmental processes involved, it is not clear that this result implies a generalized decrease in canalization, as only one character shows a strong and significant increase in FA. However, the pattern of post-bottleneck asymmetry seen in our results is consistent with results for some other populations thought to have undergone severe bottlenecks (Kieser & Groeneveld, 1991; Hutchison & Cheverud, 1995). Further, the magnitude of FA in MTR is comparable with that seen for some other inbred populations, such as FA based on cranial measures in the dama gazelle (Alados et al., 1995).

Swaddle et al. (1994) describe two potential problems (relevant to this study) with using museum material for studies of FA. First, human collections for museums may be biased towards symmetrical and otherwise ‘aesthetic’ samples. In this case, all skulls were from museums (so any bias should be similar for different samples) and most were collected opportunistically. Secondly, wear and damage asymmetry may not be discernible from FA. However, any wear would tend to diminish an apparent effect of increased post-bottleneck asymmetry, because the greater wear would be on the older (prebottleneck) skulls. Furthermore, either of these factors should affect the recent NES and SES skulls in a similar way. A greater difficulty is the small sample size for prebottleneck skulls. However, the post-bottleneck effect was strong enough that sub-samples of 11 skulls from the larger NES post-bottleneck sample still showed a significant increase in asymmetry for MTR compared with the prebottleneck NES sample. Sampling biases based on kinship are unlikely as samples were collected over a period of years.

These results show that the NES bottleneck had a quantifiable impact on genetic diversity and FA in the post-bottleneck population. An understanding of the details of such impacts is important for understanding the role of demographic fluctuations on the evolution and loss of diversity.

Acknowledgments

We thank Charles Potter, John Heyning, Philip Unitt, Maria Rutzmser, Robert DeLong, Phil Walker, Doug Kennett, Doug Jones, the Smithsonian Institution, the Los Angeles County Museum, the London Museum of Natural History, the San Diego Natural History Museum, and at the Harvard University Museum of Comparative Zoology for material and assistance.

References


Welch, B.L. 1938. The significance of the difference between two means when the population variances are unequal. *Biometrika* **29**: 350–361.


Received 8 February 2002; revised 15 March 2002; Accepted 22 March 2002