

# The crabeater seal reference genome reveals hallmarks of persistently large effective population size and sustained population expansion in the World's most abundant pinniped

Beril Yıldız\*,<sup>1,2</sup>, Thomas Gelatt<sup>3</sup>, Luis A. Hückstädt<sup>4,5,6</sup>, Daniel P. Costa<sup>4,5</sup>, Michael Tift<sup>7</sup>, Jay Rotella<sup>8</sup>, Elizabeth Flesch<sup>8</sup>, Kaitlin Macdonald<sup>8</sup>, Nancy Chen<sup>9</sup>, Robert Garrott<sup>8</sup>, Mike E. Goebel<sup>10,11</sup>, Jaime Forcada<sup>12</sup>, Thorsten Wachtmeister<sup>13</sup>, David L. J. Vendrami<sup>1,2,14</sup> and Joseph I. Hoffman<sup>1,2,12,14</sup>

<sup>1</sup> Department of Evolutionary Population Genetics, Faculty of Biology, Bielefeld University, 33501 Bielefeld, Germany

<sup>2</sup> Center for Biotechnology (CeBiTec), Faculty of Biology, Bielefeld University, 33615 Bielefeld, Germany

<sup>3</sup> Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, USA

<sup>4</sup> Institute of Marine Science, University of California Santa Cruz, Santa Cruz, California, USA

<sup>5</sup> Centre for Ecology and Conservation, College of Life and Environmental Science, University of Exeter, Penryn, Cornwall, UK

<sup>6</sup> Sharjah Marine Science Research Centre, University of Khorfakkan, UAE

<sup>7</sup> Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, North Carolina, USA

<sup>8</sup> Department of Ecology, Montana State University, P.O. Box 173460, Bozeman, MT 59717, USA

<sup>9</sup> Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, USA

<sup>10</sup> Antarctic Ecosystem Research Division, Southwest Fisheries Science Center, NOAA Fisheries, La Jolla, California, USA

<sup>11</sup> Institute of Marine Sciences, University of California Santa Cruz, Santa Cruz, California, USA

<sup>12</sup> British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK

<sup>13</sup> Genomics & Transcriptomics Laboratory, Biologisch-Medizinisches Forschungszentrum, and West German Genome Center, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225, Düsseldorf, Germany

<sup>14</sup> Joint Institute for Individualisation in a Changing Environment (JICE), Bielefeld University and University of Münster, Bielefeld, Germany

\*Corresponding author: beril.yildiz@uni-bielefeld.de

## Abstract

Population genetic theory predicts that a species' demographic history shapes patterns of genome-wide variation. However, conservation genomic studies have disproportionately focused on small or declining species, where low genetic diversity and inbreeding are major concerns, while highly abundant species have attracted comparatively less attention. Here, we investigate the crabeater seal (*Lobodon carcinophaga*) which, despite being one of the most numerous large mammals on Earth, remains largely uncharacterised in terms of its genomic diversity and demographic history. We assembled a high-quality crabeater seal reference genome from a combination of Illumina and PacBio HiFi reads, generating a 2.44 Gb assembly spanning 138 scaffolds with high completeness. To evaluate genomic diversity in a comparative context, we whole-genome resequenced 20 crabeater seals alongside 20 individuals each of three Antarctic phocids spanning a population size gradient: the Weddell seal (*Leptonychotes weddellii*), leopard seal (*Hydrurga leptonyx*) and southern elephant seal (*Mirounga leonina*). Crabeater seals carried 61.5 million SNPs compared to 12–16 million in the other species and exhibited markedly higher nucleotide diversity and negligible genomic inbreeding. We observed an excess of rare alleles, with nearly half of all variants segregating at frequencies below 5%. Demographic reconstruction revealed persistently large effective population sizes over the past million years and sustained population expansion, paralleling inferred increases in Antarctic krill associated with sea-ice expansion during the late Pleistocene. This study provides a new genomic resource and sheds new light on the evolutionary dynamics of the world's most abundant pinniped.

**Keywords:** genome assembly; crabeater seal; pinniped; effective population size; genomic diversity; inbreeding; population expansion

## Introduction

Population size changes through time, whether expansion during periods of ecological opportunity or decline under environmental constraint, fundamentally shape patterns of genetic variation within species. Such changes alter the balance among mutation, genetic drift and natural selection, leaving distinct and enduring signatures across the genome (Tajima, 1989). For example, demographic expansions typically generate an excess of rare variants as newly arising mutations accumulate faster than they are removed by drift, whereas bottlenecks reduce genetic diversity and shift the site frequency spectrum (SFS) towards intermediate frequency variants (Nei et al., 1975). Consequently, demographic changes structure standing genetic variation in predictable ways, influencing genome-wide diversity and, in turn, adaptive potential (Frankham et al., 1999). Understanding how demographic histories translate into genome-wide patterns of variation is therefore central to evolutionary and conservation genomics.

A key parameter linking demography to genomic diversity is the effective population size ( $N_e$ ) (Wright, 1931; Frankham, 1995; Charlesworth, 2009). Because nucleotide diversity ( $\pi$ ), defined as the average number of pairwise sequence differences among individuals, is expected to scale with the product of long-term  $N_e$  and the mutation rate,  $\mu$ , such that  $\pi = 4N_e\mu$  in diploid organisms (Nei et al., 1975), demographic changes directly influence levels of genetic diversity through their effects on  $N_e$ . Consequently, species that have experienced prolonged small  $N_e$  or bottlenecks tend to exhibit reduced  $\pi$ , whereas those with historically large and stable  $N_e$  maintain higher levels of genetic variation.

Small  $N_e$  is also reflected in patterns of genomic inbreeding. As population size declines, the probability that individuals inherit identical by descent (IBD) haplotypes from both parents increases, leading to the accumulation of runs of homozygosity (ROH) across the genome (Ceballos et al., 2018). Elevated homozygosity unmask the effects of deleterious alleles, causing inbreeding depression, while also exposing them to selection. The fraction of the genome in ROH ( $F_{ROH}$ ) quantifies autozygosity and captures the cumulative effects of both recent and historical population size changes (McQuillan et al., 2008). Accordingly, ROH-based metrics are widely used to quantify autozygosity and to evaluate the genomic consequences of past demographic changes in wild populations (Shafer & Kardos, 2025).

Because species with small or declining  $N_e$  are most susceptible to the loss of genetic diversity and inbreeding, empirical studies of inbreeding have often focused on small or declining populations. By contrast, species with historically large and stable populations are expected to maintain higher  $\pi$  and to exhibit reduced inbreeding (Frankham, 1996). However, the relationship between  $N_e$  and genomic diversity is not always straightforward, as genetic diversity varies far less across species than predicted from differences in population size (“Lewontin’s paradox”, (Lewontin, 1974)). This arises because linked selection, variation in mutation rate and demographic changes can decouple population size from empirical patterns of genetic variation (Romiguier et al., 2014; Lynch et al., 2016; Charlesworth & Jensen, 2022). Consequently, the underrepresentation of species with large and stable populations in studies of wild, non-model organisms leaves a gap in our understanding of how demographic history shapes genomic variation across the full spectrum of  $N_e$ .

Pinnipeds provide an ideal system to study the relationship between  $N_e$  and genomic diversity (Stoffel et al., 2018; Peart et al., 2020). They vary enormously in population size, ranging from a few hundred Mediterranean monk seals (*Monachus monachus*) to several million crabeater seals (*Lobodon carcinophaga*) and encompass species with markedly different demographic histories. Some were driven to the brink of extinction by commercial sealing, such as the northern elephant seal (*Mirounga angustirostris*), whereas others inhabiting remote regions such as Antarctica experienced little to no human impacts (Mittermeier & Wilson, 2014). To investigate how these factors shape genomic variation across taxa, high-quality genome assemblies are needed to identify ROHs and accurately reconstruct historical changes in  $N_e$  (Brüniche-Olsen et al., 2018; Li & Durbin, 2011). However, despite recent advances and the growing availability of chromosomal reference genomes (Mohr et al., 2022; Hench et al., 2024; Nebenführ et al., 2025), pinniped species inhabiting remote, ice-covered environments remain underrepresented due to logistic challenges such as harsh weather, short breeding seasons and limited access for sampling.

The crabeater seal (*Lobodon carcinophaga*) is the world’s most numerous pinniped and one of the

most abundant large mammals globally, with population size estimates varying from fewer than one million to 75 million individuals (Bengtson, 2009; Southwell et al., 2008), although the true population size is probably in the order of four million (Hückstädt, 2025). Occupying pack ice habitats across the Southern Ocean, it was largely protected from the commercial sealing of the 18<sup>th</sup> and 19<sup>th</sup> centuries. As a major consumer of Antarctic krill (*Euphausia superba*), it plays a key role in the krill-based food web of the Southern Ocean (Southwell et al., 2008). Despite its ecological importance, however, its genomic diversity and demographic history remain largely unexplored. Although this species has been included in a handful of phylogenetic and genetic studies based on mitochondrial DNA and small numbers of nuclear markers (Curtis et al., 2009, 2011; Davis et al., 2008), little is known about patterns of genome-wide diversity. Consequently, this species provides an opportunity to test whether the signatures of elevated  $\pi$  and reduced inbreeding expected under large  $N_e$  are confirmed at the genomic scale.

Here, we generated a high-quality reference genome for the crabeater seal and resequenced the genomes of 20 individuals. For comparison, we also sequenced 20 genomes each from three other Antarctic phocids: the leopard seal (*Hydrurga leptonyx*), southern elephant seal (*Mirounga leonina*) and Weddell seal (*Leptonychotes weddellii*). These species span a gradient of global population sizes, from ~220,000–440,000 leopard seals (Rogers, 2009) through ~750,000 southern elephant seals (Hindell et al., 2016) to ~202,000 Weddell seals (LaRue et al., 2021). Only the southern elephant seal experienced significant hunting during the 19<sup>th</sup> century hunting, yet it did not collapse and has since recovered (Kim et al., 2020).

Using whole-genome resequencing data from all four species, we quantified  $\pi$  and  $F_{ROH}$  and reconstructed historical changes in  $N_e$ . We hypothesized that the crabeater seal would exhibit (i) higher genome-wide diversity; (ii) minimal inbreeding; and (iii) historically large and stable  $N_e$  relative to the other Antarctic phocids.

## Materials and methods

**Study sites and sample collection.** Sample collection and sequencing procedures for each species are described below.

**Crabeater seal.** Whole blood samples were collected from eight females, six males and one individual of unknown sex, along with a skin sample from a male. Blood samples were obtained from the extradural intervertebral vein in the lower lumbar region using an 18-gauge syringe during an icebreaker cruise in 1994 around the Amundsen and Bellingshausen Seas. The skin sample was collected in 1987 in Barne Glacier, McMurdo Sound, Antarctica (Gelatt & Siniff, 1994). These samples were archived in liquid nitrogen-cooled cryovats that maintain vapor-phase nitrogen at  $-170^{\circ}\text{C}$  in the University of Alaska Museum (Arctos). Data associated with these samples are publicly available online via the Arctos database (<https://arctos.database.museum/>). Sampling was conducted in accordance with the Antarctic Conservation Act of 1978 (FR Doc No: 94-662) and under the National Marine Fisheries Service (NMFS) permit 976. An additional four whole blood samples (three females and one male) were collected in 2022 and 2023 from the Gerlache Strait in the western Antarctic Peninsula and stored at  $-80^{\circ}\text{C}$ . This sampling was conducted in accordance with the UC Santa Cruz Institutional Animal Care and Use Committee (Costd1808) under NMFS permit 25770. Sampling was granted to LAH, DPC and MST by National Science Foundation (NSF) grant (OPP 2042032).

**Leopard seal.** Skin tissue samples were collected from eight males, seven females and six individuals of unknown sex at Bird Island, South Georgia, between 2003 and 2016. Samples were collected from the webbing of the hindflipper using piglet ear notching pliers, then stored in dimethyl sulfoxide (DMSO) saturated with salt at  $-20^{\circ}\text{C}$ . Sampling was conducted in accordance with the Department for Environment, Food and Rural Affairs (DEFRA) (TARP/2016/054, AHZ/2024A/2005/1) and the Office of the Commissioner of South Georgia and the South Sandwich Islands (SGSSI) (SCI/2014/032).

**Southern elephant seal.** Skin tissue samples were collected from one adult male and 19 pups from a breeding colony at Half Moon Beach, Cape Shirreff in the South Shetland Islands between 2008 and 2016 (see Nichols et al., 2022 for details). Adults were sampled from the flanks using a 2 mm sterile disposable Miltex biopsy punch (Thermo Fisher Scientific). Pup skin samples were collected from the rear flipper using a tag hole punch or a 2 mm sterile disposable Miltex biopsy punch. The samples were immediately transferred to 95% ethanol and stored at  $-20^{\circ}\text{C}$ . Sampling was conducted in accordance with the MMPA of 1972 (16472-01 and 774-1847-04) under NMFS permits 2012-005 and 2008-008.

**Weddell seal.** Skin tissue samples were collected in 2006 from nine male and 11 female pups born that year in Erebus Bay, Ross Island, Antarctica, as part of a long-term demographic monitoring program established in 1969 (Powell et al., 2023; Proffitt et al., 2007; Rotella et al., 2009). The work that provided and curated the samples was supported by a series of grants from the NSF, Office of Polar Programs (Grant Nos. 1640481, 2147553 and 2147554) and prior NSF Grants to R.A. Garrott, J. J. Rotella, D. B. Siniff, and J. Ward Testa. Sampling was conducted in accordance with the Marine Mammal Protection Act of 1972 (MMPA; 16 U.S.C. 1361 et seq.) under NMFS permit 26375.

**Crabeater seal genome assembly.** A blood sample from a female juvenile crabeater seal captured in 2023 in the Le Maire Strait, western Antarctica, was used for the genome assembly. High-molecular-weight DNA was extracted using a QIAGEN Genomic-tip 20G kit. Illumina libraries were generated using the Illumina DNA PCR-Free Prep, Tagmentation (Illumina Inc.) following the manufacturer's instructions. PacBio libraries were constructed following the PacBio HiFi (CCS) standard protocol and sequenced on a PacBio Revio platform by BMKgene. HiFi sequencing generated a total of 106.6 Gb of HiFi data with an average sequencing depth of 51x and an average read length of 12.3 Kb. The raw data were processed for adapter and low-quality data removal using smrtlink (v9.0) CCC program with parameters “`-minLength 50 -maxLength 50000 -minPasses 1 -minPredictedAccuracy 0.99 -j 6`”. The filtered data contained 8,668,775 reads.

The genome assembly was performed using a hybrid approach in Hifiasm (v.024; Cheng et al., 2021) combining Illumina PE150 reads (BGISEQ-500) and PacBio HiFi reads (Pacbio Revio). Scaffolds shorter than 1,000 bp were excluded. The quality of the assembly was assessed by evaluating conserved gene content with CEGMA (v2.5) using the default parameters, and BUSCO (v5.8.2;

options: -m genome -c 24 -e 1e-3 --augustus) using the carnivora dataset from OrthoDB v10 as reference. In addition, the Illumina and PacBio reads were aligned to the assembly to determine the mapping rate, genome coverage and sequencing depth. Illumina reads were aligned with BWA (v0.7.10) using default settings and PacBio reads with minimap2 (v2.14-r883; options: -l 20G --MD -ax asm20 t). The assembly statistics were visualised as a snail plot using blobtoolkit (Challis et al., 2020).

**Whole genome resequencing.** We generated whole-genome resequencing data for 20 individuals from each species in this study, targeting a sequencing depth of 25x per individual. Whole genomic DNA was extracted using a standardised phenol-chloroform protocol (Sambrook 2001). The DNA extracts were first assessed photometrically using a NanoDrop One spectrophotometer (Thermo Fisher Scientific Inc.) to determine their purity. DNA integrity was then evaluated using an Agilent 5300 Fragment Analyzer System with the DNF-464 HS Large Fragment 50kb Kit (Agilent Technologies, Inc.) and the final DNA concentration was quantified fluorometrically using a Qubit dsDNA-BR assay (Thermo Fisher Scientific Inc.). Crabeater seal whole-genome sequencing libraries were prepared according to the PCR-free BGI-reseq protocol and 150bp PE sequenced on a BGISEQ-500 by BMKgene. All other samples were prepared using the Illumina DNA PCR-Free Prep Tagmentation kit (Illumina Inc.) according to the manufacturer's protocol, with >300 ng input DNA per sample. Libraries were normalized to 2 nM, pooled and 151bp PE sequenced on a NovaSeq 6000 system (Illumina Inc.) at the West German Genome Center.

**Bioinformatics and variant calling.** Read quality was assessed using FastQC v0.11.9 (Andrews, 2010), checked for potential interspecific contamination using fastq\_screen v0.15.1 (Wingett & Andrews, 2018) and visualised using multiQC v1.12 (Ewels et al., 2016). Reads with phred quality scores below 20 (-q 20) were removed using fastp v0.23.4. The remaining reads were then mapped to the respective reference genome (LepWed1.0 for Weddell seals and ASM5032014v1 for leopard seals) of each species using bwa-mem v0.7.17-r1188 (Li, 2013), except for southern elephant seals, which were mapped to the northern elephant seal reference genome (mMirAng1.0.hap1) as described by Hoffman et al. (2024). This was justified by the higher quality and improved contiguity of the northern elephant seal assembly, as well as by the relatively recent divergence between the two species (0.4–4.0 million years ago, Arnason et al., 2006; Higdon et al., 2007). The resulting SAM files were converted to BAM format, sorted, filtered for duplicates and indexed with samtools 1.15.1 (Li et al., 2009). Variant calling was performed using bcftools mpileup v1.15.1 (Li, 2011) with a minimum mapping quality of 20 (min-MQ 20) and a minimum base quality of 20 (--min-BQ 20). The called variants were subsequently filtered using bcftools v1.15.1 (Li, 2011) and vcftools v0.1.16 (Danecek et al., 2011) to retain sites with a quality scores >100 and a mean (across individuals) sequencing depth of 5–40x. In addition, individual sites with sequencing depth below eight or above 80 were masked to exclude positions with unusually low or high coverage respectively. Finally, we applied a >33% missingness threshold across individuals.

**Sex scaffold removal.** For each species, we identified and removed putative sex scaffolds prior to further analysis. For the southern elephant seal, sex scaffolds were removed as described by Hoffman et al. (2024). For the leopard seal, the four annotated sex scaffolds were directly excluded. For the crabeater and Weddell seals, where sex scaffold information was unavailable, putative sex-linked scaffolds were identified by extracting chain alignments between each target genome and the California sea lion (*Zalophus californianus*; GCA\_900631625.1; Peart et al., 2021), a species with well characterised X and Y chromosomes. Alignments were obtained from a multi-species whole genome alignment of 13 pinniped species using Progressive Cactus v2.8.4 (Armstrong et al., 2020). Scaffolds with more than 20% of their sequence aligning to the California sea lion X or Y scaffolds were flagged as putative sex-linked regions. These candidates were then validated through visual inspection of ROH patterns, and the final set of sex-linked scaffolds was excluded from all downstream analyses.

**Genomic diversity, allele frequencies and inbreeding.** We calculated per-site allele frequencies and nucleotide diversity ( $\pi$ ) in 10 kb windows across all sites using vcftools v0.1.16 (Danecek et al., 2011). To quantify genomic inbreeding, we identified ROHs using bcftools v1.15.1 (Narasimhan et al., 2016), which applies a hidden Markov model to detect autozygous regions along the genome. Because ROH detection methods are commonly based on SNP data, their performance can be sensitive to the high density of variants in WGS datasets, which, despite quality filtering, is often accompanied by elevated genotyping error rates that can affect ROH identification. Specifically, false heterozygous calls can artificially break up ROHs, causing them to appear shorter or be missed

entirely depending on the minimum length threshold. To mitigate this bias, we reduced SNP density prior to ROH calling by thinning the WGS datasets using VCFtools v0.1.16 (Danecek et al., 2011), retaining approximately one SNP every 3kb. This approach reduces the artificial fragmentation of ROHs while preserving a sufficient density of markers for robust ROH detection. After filtering, the number of SNPs retained was 771,127 for the crabeater seal, 697,462 for the leopard seal, 745,293 for the southern elephant seal and 690,569 for the Weddell seal. A minimum ROH length of 100kb was used to detect the true autozygous segments, and  $F_{\text{ROH}}$  was calculated for each individual as the proportion of the genome contained within ROHs.

**Demographic reconstruction.** We inferred historical changes in  $N_e$  using SMC++ (v1.15.4; Terhorst et al., 2017). Autosomal VCF files were converted to SMC++ input format using `smc++ vcf2smc`, specifying the appropriate contig lengths. Analyses were performed assuming a mutation rate of  $\mu = 7 \times 10^{-9}$  mutations per site per generation (Peart et al., 2020) and a generation time ( $g$ ) of 14.9 years (crabeater seal; Hückstädt, 2025), 10.4 years (leopard seal; Hückstädt, 2015a), 9.5 years (southern elephant seal; Hofmeyr, 2015) and 10.8 years (Weddell seal; Hückstädt, 2015b), taken from IUCN estimates. Uncertainty in  $N_e$  estimates was assessed using a block bootstrapping approach. Each chromosome was divided into 5 Mb segments, which were randomly resampled and concatenated to generate pseudochromosomes representing full-genome-size datasets. This was implemented using the publicly available script `bootstrap_smcpp.py` (available at <https://github.com/popgenmethods/smcpp/issues/37>). We generated 20 bootstrap replicates and SMC++ inference was repeated independently for each replicate to evaluate variability in demographic trajectories.

## Results and discussion

### Genome assembly.

The final scaffolded and gap-closed genome assembly of the crabeater seal comprised 138 scaffolds with a total length of 2,444,645,479 bp ( $N_{50} = 99.1$  Mb,  $N_{90} = 32.4$  Mb; Figure 1). Gene completeness evaluated using BUSCO with the carnivora dataset identified 13,383 complete sequences, corresponding to 94.7% completeness, including 13,383 complete single-copy orthologs (92.3%), while 561 genes (3.9%) were missing. To further evaluate the completeness of the assembly, we aligned the Illumina and PacBio reads to the reference genome. The Illumina reads showed a mapping rate of 99.97% with a genome coverage of 99.98% and an average sequencing depth of 183x. The PacBio HiFi reads similarly achieved a mapping rate of 99.97%, with a genome coverage of 99.99% and an average sequencing depth of 44x.

These assembly statistics compare favourably with other recent chromosomal pinniped reference genomes (see Supplementary table 1). Until recently, achieving this level of contiguity required chromatin conformation capture approaches such as HiC to link scaffolds, which are both costly and dependent on very high-quality input material. By contrast, the present assembly demonstrates that highly contiguous genomes can now be generated using a combination of Illumina and HiFi data without the need for HiC. Although not strictly chromosomal, the relatively small number of scaffolds ( $n = 138$ ) represents a substantial improvement over earlier pinniped assemblies and showcases rapid advances in genome sequencing and assembly approaches.

### Genomic diversity and inbreeding.

To evaluate the crabeater seal's genomic diversity and place it into context, we computed the number of variant sites separately for each of the four species. Crabeater seals carried substantially more SNPs (61.5 million) than the other three species (~12–16 million). In line with this, they also exhibited the highest  $\pi$  ( $0.0031$ ,  $\pm 0.00126$  SD) compared to leopard ( $0.0010 \pm 0.00054$  SD), southern elephant ( $0.0016 \pm 0.00101$  SD) and Weddell seals ( $0.0009 \pm 0.00064$  SD; Figure 2A). Notably, these values are comparable to previous  $\pi$  estimates based on mitochondrial DNA or microsatellites (Bender et al., 2023; Curtis et al., 2009, 2011). Overall, these results suggest that crabeater seals have maintained high levels of standing genetic variation on a genome-wide scale, consistent with the expectation that large  $N_e$  promotes the accumulation and persistence of allelic diversity across generations.

Patterns of genomic inbreeding differed markedly among species (Figure 2B). Crabeater seals exhibited consistently extremely low inbreeding (mean  $F_{ROH} = 0.002 \pm 0.0007$  SD). By contrast, the other three species showed  $F_{ROH}$  values that were one order of magnitude higher (Weddell seals:  $0.017 \pm 0.0021$ , leopard seals:  $0.011 \pm 0.0032$ , southern elephant seals:  $0.019 \pm 0.0032$ ). Across all species, long ROHs (>5Mb), which are indicative of common ancestry within the last ~10 generations, contributed relatively little to  $F_{ROH}$  (Supplementary figure 1). Instead, inbreeding was mostly attributable to short- and intermediate-length ROHs. These patterns suggest that genomic inbreeding in Antarctic phocids is shaped primarily by historical demography rather than recent consanguineous mating. Hence, the very low  $F_{ROH}$  of crabeater seals is consistent with a historically large  $N_e$ .

### Allele frequency spectra and demographic histories.

Because the SFS reflects past changes in population size, with an excess of rare alleles indicating population expansion and a flatter spectrum suggesting past bottlenecks, we inspected the SFS of each species. All four Antarctic phocids showed an excess of rare variants, largely reflected in negative Tajima's D values and left-skewed allele frequency distributions (Figure 2C, Supplementary table 2). This pattern was most pronounced in the crabeater seal, where approximately half of all variants occurred at low frequency (MAF < 0.05: 48.96%). This enrichment of rare alleles in the crabeater seal is consistent with a non-equilibrium demographic history, most plausibly reflecting long-term population expansion, and it clearly distinguishes this species from the other three Antarctic phocids.

To investigate population size changes over time, we reconstructed historical  $N_e$  trajectories using SMC++, which implements a composite-likelihood framework based on the sequentially Markovian coalescent while leveraging information from multiple unphased genomes. The results are shown in Figure 3, while uncertainty in the estimates, assessed using a block bootstrapping approach, is shown in Supplementary figure 2. Crabeater seals exhibited consistently higher  $N_e$  over the past ~1 million years and showed clear evidence of sustained population expansion (Figure 3), consistent with the

excess of rare alleles observed in the SFS. Their  $N_e$  increased to a peak of around 1.6 million (95% bootstrap interval 0.59-0.66 million). By contrast,  $N_e$  was orders of magnitude lower for the other species, remaining below ~50,000 for leopard seals, ~60,000 for southern elephant seals and ~95,000 for Weddell seals, and exhibiting smaller temporal fluctuations (Figure 3, Supplementary figure 2).

The demographic expansion of crabeater seals is likely driven by their tight trophic coupling with Antarctic krill. Demographic reconstructions from krill genomes reveal a similar pattern of population expansion over the past ~100,000 years, which is associated with cooler conditions during the late Pleistocene that increased the extent of sea ice habitat on which krill depend (Shao 2023). Crabeater seals also exhibit considerable behavioural plasticity in response to variation in krill availability, which may have helped to sustain their population growth by buffering against shorter-term fluctuations in prey availability (Hückstädt et al., 2012). This apparent demographic coupling between predator and prey aligns with patterns observed in another krill-dependent species, the Antarctic fur seal (*Arctocephalus gazella*) at South Georgia, where increased krill availability following the near extirpation of the great whales from the Southern Ocean was thought to have contributed to population recovery from commercial sealing (Hoffman et al. 2022), while more recent population declines since 2009 have been linked to reductions in krill (Forcada & Hoffman, 2014; Forcada et al., 2023). By contrast, the other three Antarctic phocids did not show such pronounced population expansions, likely because they exploit a broader range of prey species (Casaux et al., 2005; Bender et al., 2023; Woodman et al., 2024) whose populations may have been less sensitive to the specific environmental changes that favoured krill specialists.

## Conclusions

We assembled a high-quality crabeater seal genome and performed a comparative analysis of genomic diversity among Antarctic phocids. Consistent with population genetic theory, we observed high genome-wide diversity, low inbreeding and evidence for a large  $N_e$  and sustained population growth in crabeater seals. While most conservation genetic studies of wild populations have focused on small or declining populations, our results contribute to understanding the genomic consequences of very large population sizes. By investigating a thriving polar pinniped, this study contributes to a more complete understanding of the ecological and evolutionary processes that shape genetic diversity across the full spectrum of population sizes.

## Data availability

The genome assembly and the whole genome sequencing data from the four pinniped species will be available at NCBI upon acceptance. The code, genome assembly and the whole genome sequencing data generated in this study will be available upon acceptance.

## Acknowledgments

The authors thank the researchers, staff, and volunteers who contributed to the collection of the samples and field data used for this study. We are also grateful to Mallory Gulbranson for helping us with retrieving the crabeater seal samples from archived freezer collections and coordinating the shipment. We thank Ira Wachendorf and Ann-Christin Polikeit for carrying out DNA extractions and quality checks, as well as Rebecca Carter for providing the seal artwork.

## Funding

This work was funded by the German Research Foundation (DFG) Sequencing Costs in Projects scheme (project number 497640428 awarded to J.I.H.) and the DFG Schwerpunktprogramm (priority programme) 1158, 'Antarctic Research with Comparative Investigations in Arctic Ice Areas' (project number 424119118 awarded to J.I.H.). This work was also supported by the DFG Research Infrastructure West German Genome Center (project number 388941457) as part of the Next Generation Sequencing Competence Network.

## Author contributions

**crabeater seal:** TG carried out the fieldwork, collected and stored the 16 crabeater samples. The genome assembly sample and the remaining tissues used for whole genome resequencing of four crabeater seal samples were collected, stored, and provided by LAH, DPC and MT.

**Weddell seal:** JR, RG, and KM carried out the fieldwork, collected and stored the Weddell seal samples. EF and NC selected the samples for sequencing. EF and KM sourced the samples from archived freezer collections and coordinated their shipment.

**Southern elephant seal:** MEG, FG and SS contributed the Southern elephant seal samples used in this study.

**Leopard seal:** JF carried out the fieldwork and collected the leopard seal samples.

DLJV performed quality control on the samples prior to sequencing and shipment to the sequencing centre. TW was responsible for sequencing the samples at the West German Genome Center in Düsseldorf, except for the crabeater seals. JIH acquired funding and supervised the PhD student (BY). BY and JIH coordinated the study and drafted the manuscript. BY developed the methodology, performed the analyses and created the figures. All co-authors commented upon and approved the final manuscript.

### Conflicts of interest

The authors declare no conflicts of interest.

### References

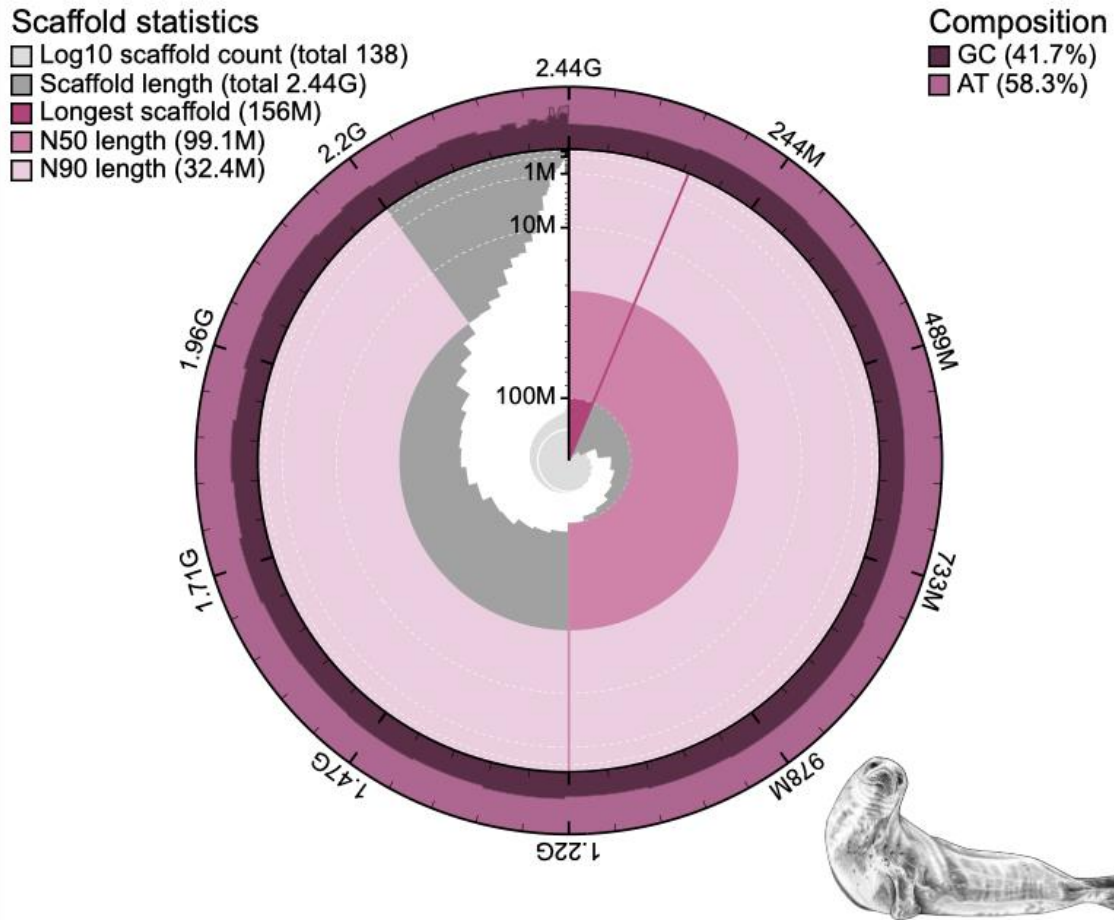
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [Computer software].
- Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I. T., Novak, A. M., Deran, A., Fang, Q., Xie, D., Feng, S., Stiller, J., Genereux, D., Johnson, J., Marinescu, V. D., Alföldi, J., Harris, R. S., Lindblad-Toh, K., Haussler, D., Karlsson, E., Jarvis, E. D., ... Paten, B. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature*, 587(7833), 246–251. <https://doi.org/10.1038/s41586-020-2871-y>
- Arnason, U., Gullberg, A., Janke, A., Kullberg, M., Lehman, N., Petrov, E. A., & Väinölä, R. (2006). Pinniped phylogeny and a new hypothesis for their origin and dispersal. *Molecular Phylogenetics and Evolution*, 41(2), 345–354. <https://doi.org/10.1016/j.ympev.2006.05.022>
- Bender, A. N., Krause, D. J., Goebel, M. E., Hoffman, J. I., Lewallen, E. A., & Bonin, C. A. (2023). Genetic diversity and demographic history of the leopard seal: A Southern Ocean top predator. *PLOS ONE*, 18(8), e0284640. <https://doi.org/10.1371/journal.pone.0284640>
- Bengtson, J. L. (2009). Crabeater seal *Lobodon carcinophaga*. In Perrin WF, Würsig B, Thewissen JGM (Eds) *Encyclopedia of marine mammals*. (Second Edition, pp. 290–292). (Original work published Academic Press, San Diego, USA)
- Brüniche-Olsen, A., Kellner, K. F., Anderson, C. J., & DeWoody, J. A. (2018). Runs of homozygosity have utility in mammalian conservation and evolutionary studies. *Conservation Genetics*, 19(6), 1295–1307. <https://doi.org/10.1007/s10592-018-1099-y>
- Casaux, R., Baroni, A., & Ramón, A. (2005). The diet of the Weddell Seal *Leptonychotes weddellii* at the Danco Coast, Antarctic Peninsula. *Polar Biology*, 29(4), 257–262. <https://doi.org/10.1007/s00300-005-0048-7>
- Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4), 220–234. <https://doi.org/10.1038/nrg.2017.109>
- Challis, R., Richards, E., Rajan, J., Cochrane, G., & Blaxter, M. (2020). BlobToolKit – Interactive Quality Assessment of Genome Assemblies.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205. <https://doi.org/10.1038/nrg2526>

- Charlesworth, B., & Jensen, J. D. (2022). How Can We Resolve Lewontin's Paradox? *Genome Biology and Evolution*, 14(7), evac096. <https://doi.org/10.1093/gbe/evac096>
- Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, 18(2), 170–175. <https://doi.org/10.1038/s41592-020-01056-5>
- Curtis, C., Stewart, B. S., & Karl, S. A. (2009). Pleistocene population expansions of Antarctic seals. *Molecular Ecology*, 18(10), 2112–2121. <https://doi.org/10.1111/j.1365-294X.2009.04166.x>
- Curtis, C., Stewart, B. S., & Karl, S. A. (2011). Genetically effective population sizes of Antarctic seals estimated from nuclear genes. *Conservation Genetics*, 12(6), 1435–1446. <https://doi.org/10.1007/s10592-011-0241-x>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Davis, C. S., Stirling, I., Strobeck, C., & Coltman, D. W. (2008). Population structure of ice-breeding seals. *Molecular Ecology*, 17(13), 3078–3094. <https://doi.org/10.1111/j.1365-294X.2008.03819.x>
- Ewels, P., Magnusson, M., Lundin, S., & Källér, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
- Forcada, J., & Hoffman, J. I. (2014). Climate change selects for heterozygosity in a declining fur seal population. *Nature*, 511(7510), 462–465. <https://doi.org/10.1038/nature13542>
- Forcada, J., Hoffman, J. I., Gimenez, O., Staniland, I. J., Bucktrout, P., & Wood, A. G. (2023). Ninety years of change, from commercial extinction to recovery, range expansion and decline for Antarctic fur seals at South Georgia. *Global Change Biology*, 29(24), 6867–6887. <https://doi.org/10.1111/gcb.16947>
- Frankham, R. (1995). *Conservation Genetics*. 29:305-327. <https://doi.org/10.1146/annurev.ge.29.120195.001513>
- Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, 10(6), 1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation*, 2(4), 255–260. <https://doi.org/10.1111/j.1469-1795.1999.tb00071.x>
- Gelatt, T. S., & Siniff, D. B. (1994). Line Transect Survey of Crabeater Seals in the Amundsen-Bellinghusen Seas, 1994. *Wildlife Society Bulletin (1973-2006)*, 27(2), 330–336.
- Hench, K., Vendrami, D. L. J., Forcada, J., & Hoffman, J. I. (2024). Refinement of the Antarctic fur seal (*Arctocephalus gazella*) reference genome increases continuity and completeness. *G3: Genes, Genomes, Genetics*, 14(11), jkae179. <https://doi.org/10.1093/g3journal/jkae179>
- Higdon, J. W., Bininda-Emonds, O. R., Beck, R. M., & Ferguson, S. H. (2007). Phylogeny and divergence of the pinnipeds (Carnivora: Mammalia) assessed using a multigene dataset. *BMC Evolutionary Biology*, 7(1), 216. <https://doi.org/10.1186/1471-2148-7-216>

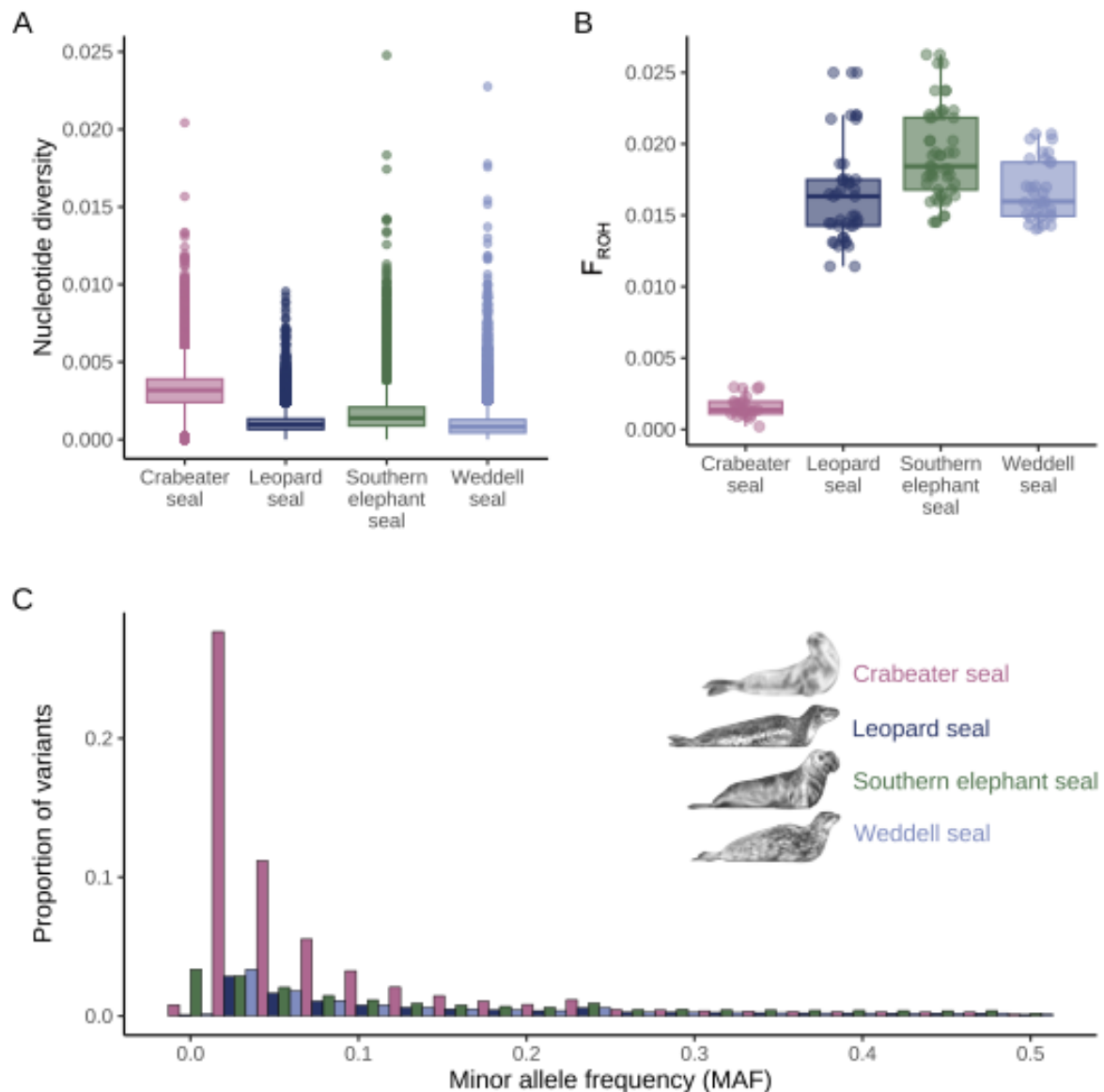
- Hindell, M. A., McMahon, C. R., Bester, M. N., Boehme, L., Costa, D., Fedak, M. A., Guinet, C., Herraiz-Borreguero, L., Harcourt, R. G., Huckstadt, L., Kovacs, K. M., Lydersen, C., McIntyre, T., Muelbert, M., Patterson, T., Roquet, F., Williams, G., & Charrassin, J. (2016). Circumpolar habitat use in the southern elephant seal: Implications for foraging success and population trajectories. *Ecosphere*, 7(5), e01213. <https://doi.org/10.1002/ecs2.1213>
- Hoffman, J. I., Vendrami, D. L. J., Hench, K., Chen, R. S., Stoffel, M. A., Kardos, M., Amos, W., Kalinowski, J., Rickert, D., Köhrer, K., Wachtmeister, T., Goebel, M. E., Bonin, C. A., Gulland, F. M. D., & Dasmahapatra, K. K. (2024). Genomic and fitness consequences of a near-extinction event in the northern elephant seal. *Nature Ecology & Evolution*, 8(12), 2309–2324. <https://doi.org/10.1038/s41559-024-02533-2>
- Hofmeyr, G. J. G. (2015). *Mirounga leonina*. The IUCN Red List of Threatened Species 2015. <https://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T13583A45227247.en>
- Hückstädt, L. A. (2015a). *Hydrurga leptonyx*. The IUCN Red List of Threatened Species 2015. <https://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T10340A45226422.en>
- Hückstädt, L. A. (2015b). *Leptonychotes weddellii*. The IUCN Red List of Threatened Species 2015. <https://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T11696A45226713.en>
- Hückstädt, L. A. (2025). *Lobodon carcinophaga*. The IUCN Red List of Threatened Species 2015. <https://dx.doi.org/10.2305/IUCN.UK.2025-2.RLTS.T12246A95969809.en>
- Hückstädt, L. A., Burns, J., Koch, P., McDonald, B., Crocker, D., & Costa, D. (2012). Diet of a specialist in a changing environment: The crabeater seal along the western Antarctic Peninsula. *Marine Ecology Progress Series*, 455, 287–301. <https://doi.org/10.3354/meps09601>
- Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, 189(1), 237–249. <https://doi.org/10.1534/genetics.111.130922>
- Kim, B.-M., Lee, Y. J., Kim, J.-H., Kim, J.-H., Kang, S., Jo, E., Lee, S. J., Lee, J. H., Chi, Y. M., & Park, H. (2020). The Genome Assembly and Annotation of the Southern Elephant Seal *Mirounga leonina*. *Genes*, 11(2), 160. <https://doi.org/10.3390/genes11020160>
- LaRue, M., Salas, L., Nur, N., Ainley, D., Stammerjohn, S., Pennycook, J., Dozier, M., Saints, J., Stamatiou, K., Barrington, L., & Rotella, J. (2021). Insights from the first global population estimate of Weddell seals in Antarctica. *Science Advances*, 7(39), eabh3674. <https://doi.org/10.1126/sciadv.abh3674>
- Lewontin, R. C. (1974). *The genetic basis of evolutionary change*. New York: Columbia University Press.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM (arXiv:1303.3997). arXiv. <https://doi.org/10.48550/arXiv.1303.3997>
- Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357), 493–496. <https://doi.org/10.1038/nature10231>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>

- Lynch, M., Ackerman, M. S., Gout, J.-F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. *Nature Reviews Genetics*, 17(11), 704–714. <https://doi.org/10.1038/nrg.2016.104>
- McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., ... Wilson, J. F. (2008). Runs of Homozygosity in European Populations. *American Journal of Human Genetics*, 83(3), 359–372. <https://doi.org/10.1016/j.ajhg.2008.08.007>
- Mittermeier, R. A., & Wilson, D. E. (2014). *Handbook of the Mammals of the World—Volume 4: Sea Mammals*. Lynx Edicions, Barcelona, Spain.
- Mohr, D. W., Gaughran, S. J., Paschall, J., Naguib, A., Pang, A. W. C., Dudchenko, O., Aiden, E. L., Church, D. M., & Scott, A. F. (2022). A Chromosome-Length Assembly of the Hawaiian Monk Seal (*Neomonachus schauinslandi*): A History of “Genetic Purging” and Genomic Stability. *Genes*, 13(7), 1270. <https://doi.org/10.3390/genes13071270>
- Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C., & Durbin, R. (2016). BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11), 1749–1751. <https://doi.org/10.1093/bioinformatics/btw044>
- Nebenführ, M., Hamadou, A. B., Aguilar, A., Borrell, A., Gkafas, G. A., & Janke, A. (2025). Mediterranean monk seal (*Monachus monachus*) and leopard seal (*Hydrurga leptonyx*) de novo genomes to study the demographic history and genetic diversity of southern seals. *BMC Biology*, 23(1), 102. <https://doi.org/10.1186/s12915-025-02207-w>
- Nei, M., Maruyama, T., & Chakraborty, R. (1975). *The Bottleneck Effect and Genetic Variability in Populations*.
- Nichols, H. J., Fuchs, B., Paijmans, A. J., Lewis, G., Bonin, C. A., Goebel, M. E., & Hoffman, J. I. (2022). Where are the beachmasters? Unexpectedly weak polygyny among southern elephant seals on a South Shetland Island. *Journal of Zoology*, 316(2), 104–117. <https://doi.org/10.1111/jzo.12936>
- Peart, C. R., Tusso, S., Pophaly, S. D., Botero-Castro, F., Wu, C.-C., Auriolles-Gamboa, D., Baird, A. B., Bickham, J. W., Forcada, J., Galimberti, F., Gemmell, N. J., Hoffman, J. I., Kovacs, K. M., Kunnsaranta, M., Lydersen, C., Nyman, T., De Oliveira, L. R., Orr, A. J., Sanvito, S., ... Wolf, J. B. W. (2020). Determinants of genetic variation across eco-evolutionary scales in pinnipeds. *Nature Ecology & Evolution*, 4(8), 1095–1104. <https://doi.org/10.1038/s41559-020-1215-5>
- Peart, C. R., Williams, C., Pophaly, S. D., Neely, B. A., Gulland, F. M. D., Adams, D. J., Ng, B. L., Cheng, W., Goebel, M. E., Fedrigo, O., Haase, B., Mountcastle, J., Fungtammasan, A., Formenti, G., Collins, J., Wood, J., Sims, Y., Torrance, J., Tracey, A., ... Wolf, J. B. W. (2021). Hi-C scaffolded short- and long-read genome assemblies of the California sea lion are broadly consistent for syntenic inference across 45 million years of evolution. *Molecular Ecology Resources*, 21(7), 2455–2470. <https://doi.org/10.1111/1755-0998.13443>
- Powell, J. H., Kalinowski, S. T., Taper, M. L., Rotella, J. J., Davis, C. S., & Garrott, R. A. (2023). Evidence of an Absence of Inbreeding Depression in a Wild Population of Weddell Seals (*Leptonychotes weddellii*). *Entropy*, 25(3), 403. <https://doi.org/10.3390/e25030403>

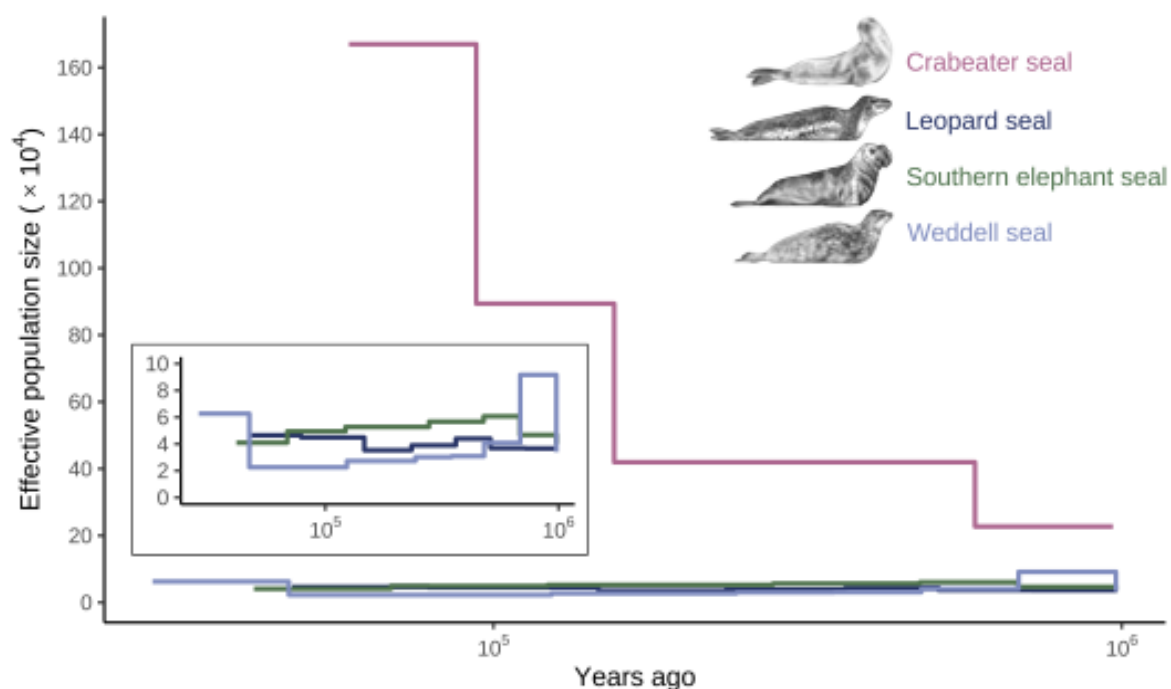
- Proffitt, K. M., Garrott, R. A., Rotella, J. J., & Wheatley, K. E. (2007). Environmental and senescent related variations in Weddell seal body mass: Implications for age-specific reproductive performance. *Oikos*, 116(10), 1683–1690. <https://doi.org/10.1111/j.0030-1299.2007.16139.x>
- Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A. A.-T., Weinert, L. A., Belkhir, K., Bierne, N., ... Galtier, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515(7526), 261–263. <https://doi.org/10.1038/nature13685>
- Rotella, J. J., Link, W. A., Nichols, J. D., Hadley, G. L., Garrott, R. A., & Proffitt, K. M. (2009). An evaluation of density-dependent and density-independent influences on population growth rates in Weddell seals. *Ecology*, 90(4), 975–984. <https://doi.org/10.1890/08-0971.1>
- Shafer, A. B. A., & Kardos, M. (2025). Runs of Homozygosity and Inferences in Wild Populations. *Molecular Ecology*, 34(3), e17641. <https://doi.org/10.1111/mec.17641>
- Southwell, C., Paxton, C. G. M., Borchers, D., Boveng, P., & De La Mare, W. (2008). Taking account of dependent species in management of the Southern Ocean krill fishery: Estimating crabeater seal abundance off east Antarctica. *Journal of Applied Ecology*, 45(2), 622–631. <https://doi.org/10.1111/j.1365-2664.2007.01399.x>
- Tajima, F. (1989). The effect of change in population size on DNA polymorphism. *Genetics*, 123(3), 597–601. <https://doi.org/10.1093/genetics/123.3.597>
- Terhorst, J., Kamm, J. A., & Song, Y. S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, 49(2), 303–309. <https://doi.org/10.1038/ng.3748>
- Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. 7(1338). <https://f1000research.com/articles/7-1338/v2>
- Woodman, S. M., Borrás-Chavez, R., Goebel, M. E., Torres, D., Aguayo, A., & Krause, D. J. (2024). CS-PHOC: Weekly census counts of Southern Ocean phocids at Cape Shirreff, Livingston Island. *Scientific Data*, 11(1), 895. <https://doi.org/10.1038/s41597-024-03744-9>
- Wright, S. (1931). Evolution in Mendelian Populations. 16(2):97–159. <https://doi.org/10.1093/genetics/16.2.97>
- Sambrook, J., & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.



**Figure 1.** Snail plot summarizing the crabeater seal genome assembly. The innermost circle shows scaffold statistics scaled to the length of the longest scaffold, with the colours from dark to light purple indicating the longest scaffold, N50 and N90 lengths respectively, as shown in the legend. The light grey distribution shows the cumulative scaffold count on a log scale, while the dark grey distribution shows the scaffold length in descending order of size. GC and AT composition is shown in the outer circle. The seal illustration was created by Rebecca Carter (<http://www.rebeccacarterart.co.uk>) and is reproduced with her permission.



**Figure 2.** Nucleotide diversity, genome-wide inbreeding and allele frequency spectra in four Antarctic phocids: the crabeater seal, leopard seal, southern elephant seal and Weddell seal. (A) Nucleotide diversity ( $\pi$ ) estimated in 10 kb non-overlapping windows across the genome; (B) individual inbreeding coefficients  $F_{ROH}$  estimated from runs of homozygosity; (C) minor allele frequency (MAF) spectra showing the proportion of variants across frequency intervals. Raw data points are shown alongside boxplots; the thick centre lines represent median, the lower and upper hinges correspond to the first and third quartiles, respectively, and the whiskers represent largest and smallest value but no further than  $1.5\times$  the interquartile range. The bars in the barplots represent the proportion of variants falling within each frequency interval, with colours corresponding to species as indicated in the legend. The seal illustrations were created by Rebecca Carter (<http://www.rebeccacarterart.co.uk>) and are reproduced with her permission.



**Figure 3.** Historical effective population size ( $N_e$ ) trajectories inferred with SMC++. Coloured lines show the maximum-likelihood estimates of  $N_e$  based on whole-genome sequencing data from 20 individuals per species. Time is shown in years before present on a logarithmic scale. Analyses were restricted to autosomal chromosomes and assumed a mutation rate of  $\mu = 0.7 \times 10^{-8}$  mutations per site per generation (Peart et al., 2020) and species-specific generation times of 14.9 years (crabeater seal; Hückstädt, 2025), 10.4 years (leopard seal; Hückstädt, 2015a), 9.5 years (southern elephant seal; Hofmeyr, 2015) and 10.8 years (Weddell seal; Hückstädt, 2015b), taken from IUCN estimates. The inset shows the  $N_e$  estimates for leopard seal, southern elephant seal and Weddell seal on a rescaled y-axis (0-10  $\times 10^4$ ). The seal illustrations were created by Rebecca Carter (<http://www.rebeccacarterart.co.uk/>) and are reproduced with her permission.

### Supplementary material

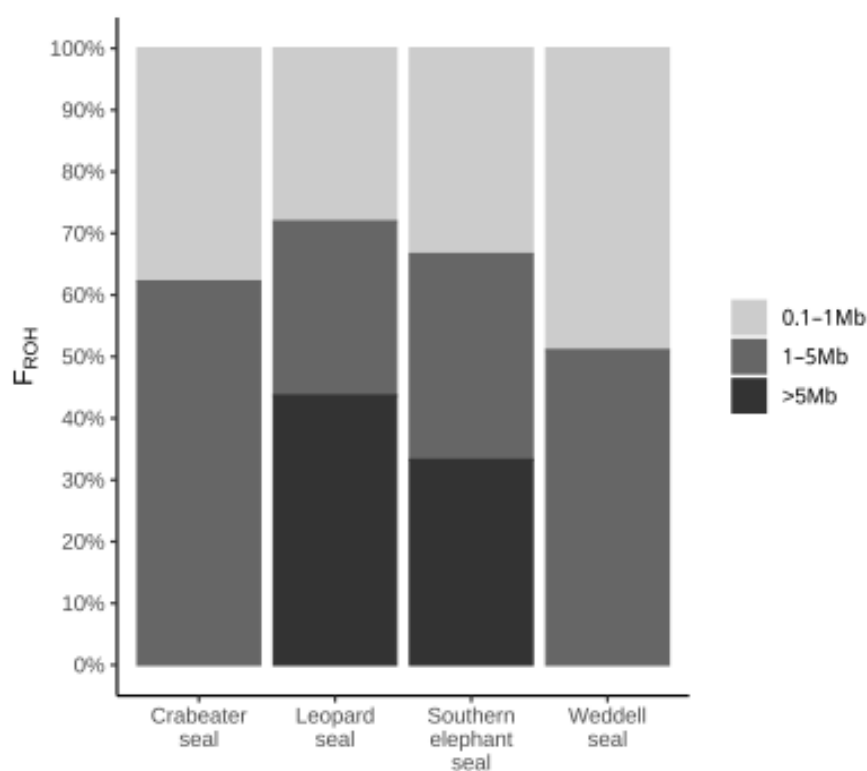
**Table S1.** Assembly statistics of the crabeater seal reference genome compared to recent genome assemblies of the leopard seal, northern elephant seal, Antarctic fur seal and California sea lion.

	Crabeater seal	Leopard seal	Northern elephant seal	Antarctic fur seal	California sea lion
<b>Assembly ID</b>	LobCar_v1	ASM5032014 v1	mMirAng1.0. hap1	arcGaz4_h1	mZalCal1.pri. v2
<b>Assembly level</b>	scaffold	scaffold	scaffold	scaffold	chromosome
<b>Genome size (bp)</b>	2,444,645,479	2,454,374,544	2,430,305,188	2,527,999,884	2,409,685,272

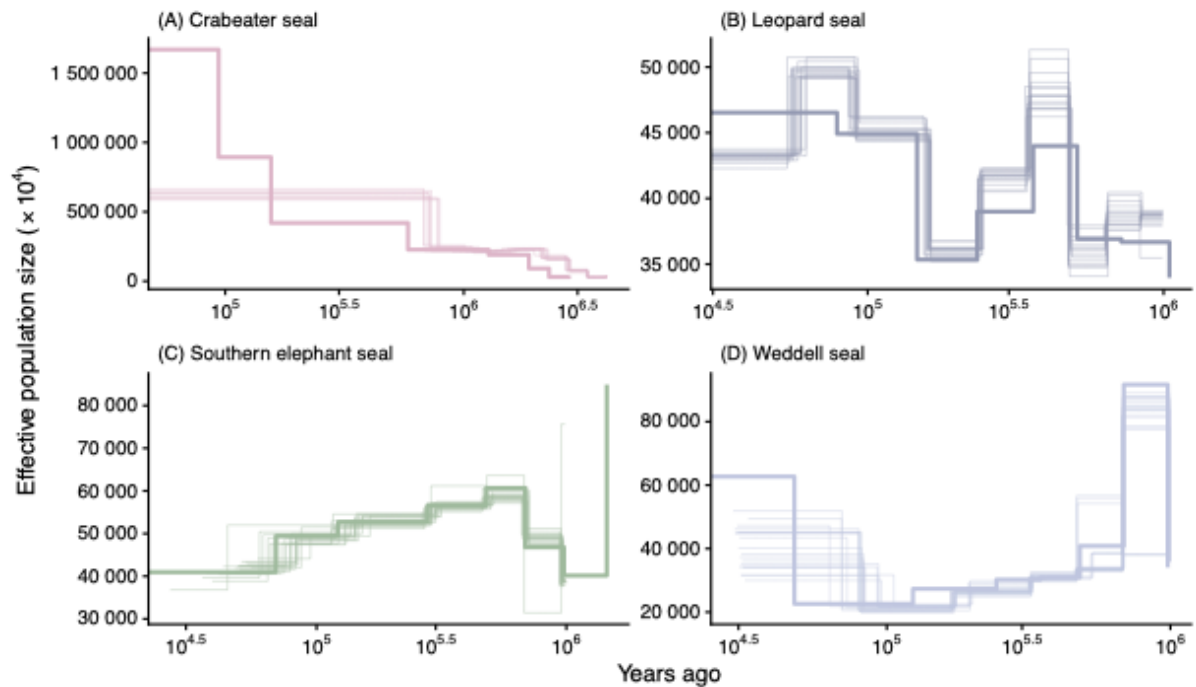
<b>Number of scaffolds</b>	138	203	497	541	47
<b>Scaffold N50 length</b>	99,142,401	99,454,515	154,172,358	141,635,559	147,124,152
<b>Scaffold N90 length</b>	32,405,651	31,951,825	93,938,456	64,524,230	94,255,706
<b>% BUSCO completeness</b>	94.7 (carnivora_odb 10)	98.2 (carnivora_odb 10)	96.1 (mammalia_odb10)	95.2 (carnivora_odb 10)	93.9 (mammalia_odb9)

**Table S2.** Summary statistics of the allele frequency spectrum for four Antarctic phocids: the crabeater seal, leopard seal, southern elephant seal and Weddell seal. Proportions of singletons (minor allele count = 1) and doubletons (minor allele count = 2), and median Tajima's D values (interquartile range) calculated in 10kb non-overlapping windows with at least five SNPs.

	<b>Crabeater seal</b>	<b>Leopard seal</b>	<b>Southern elephant seal</b>	<b>Weddell seal</b>
<b>% singletons</b>	44.1	20.8	10.3	24.3
<b>% doubletons</b>	17.3	11.5	7.3	11.5
<b>Tajima's D</b>	-1.76 (-1.92, -1.57)	-0.38 (-0.86, 0.1)	0.03 (-0.51, 0.56)	-0.51 (-0.1,-0.02)



**Figure S1.** Genome-wide inbreeding composition across ROH length categories for four Antarctic phocids: the crabeater seal, leopard seal, southern elephant seal and Weddell seal. Individual inbreeding coefficients ( $F_{ROH}$ ) estimated from runs of homozygosity (ROHs) are grouped into three length categories: 0.1–1 Mb, 1–5 Mb and >5 Mb. The stacked bars show the mean proportion of the total  $F_{ROH}$  attributable to each ROH length class per species. Short ROHs reflect more ancient shared ancestry, whereas long ROHs indicate more recent shared ancestry.



**Figure S2.** Uncertainty in historical effective population size ( $N_e$ ) trajectories inferred using SMC++. Bold coloured lines show the maximum-likelihood estimates of  $N_e$  based on whole-genome sequencing data from 20 individuals per species, while the thin lines show  $N_e$  trajectories from 20 block bootstrap replicates generated by resampling 5 Mb genomic segments. Time is shown in years before present on a logarithmic scale. Analyses were restricted to autosomal chromosomes and assumed a mutation rate of  $\mu = 0.7 \times 10^{-8}$  mutations per site per generation (Peart et al., 2020) and species-specific generation times of 14.9 years (crabeater seal; Hückstädt, 2025), 10.4 years (leopard seal; Hückstädt, 2015a), 9.5 years (southern elephant seal; Hofmeyr, 2015) and 10.8 years (Weddell seal; Hückstädt, 2015b), taken from IUCN estimates.