

Reduced representation sequencing detects only subtle regional structure in a heavily exploited and rapidly recolonizing marine mammal species

Nicolas Dussex^{1,2}  | Helen R. Taylor¹  | Willam R. Stovall¹  | Kim Rutherford¹ | Ken G. Dodds³  | Shannon M. Clarke³ | Neil J. Gemmell¹ 

¹Department of Anatomy, University of Otago, Dunedin, New Zealand

²Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

³Invermay Agricultural Centre, AgResearch, Puddle Alley, Mosgiel, New Zealand

Correspondence

Nicolas Dussex, Department of Anatomy, University of Otago, PO Box 913, Dunedin 9016, New Zealand.

Email: nicolas.dussex@gmail.com

Funding information

University of Otago; Allan Wilson Centre; National Institute of Health (NIH) Minority Health Research Grant; New Zealand Ministry for Business, Innovation, and Employment (MBIE) via its funding of the "Genomics for Production & Security in a Biological Economy" program

Abstract

Next-generation reduced representation sequencing (RRS) approaches show great potential for resolving the structure of wild populations. However, the population structure of species that have shown rapid demographic recovery following severe population bottlenecks may still prove difficult to resolve due to high gene flow between subpopulations. Here, we tested the effectiveness of the RRS method Genotyping-By-Sequencing (GBS) for describing the population structure of the New Zealand fur seal (NZFS, *Arctocephalus forsteri*), a species that was heavily exploited by the 19th century commercial sealing industry and has since rapidly recolonized most of its former range from a few isolated colonies. Using 26,026 neutral single nucleotide polymorphisms (SNPs), we assessed genetic variation within and between NZFS colonies. We identified low levels of population differentiation across the species range (<1% of variation explained by regional differences) suggesting a state of near panmixia. Nonetheless, we observed subtle population substructure between West Coast and Southern East Coast colonies and a weak, but significant ($p = 0.01$), isolation-by-distance pattern among the eight colonies studied. Furthermore, our demographic reconstructions supported severe bottlenecks with potential 10-fold and 250-fold declines in response to Polynesian and European hunting, respectively. Finally, we were able to assign individuals treated as unknowns to their regions of origin with high confidence (96%) using our SNP data. Our results indicate that while it may be difficult to detect population structure in species that have experienced rapid recovery, next-generation markers and methods are powerful tools for resolving fine-scale structure and informing conservation and management efforts.

KEYWORDS

commercial fisheries, conservation, genotyping-by-sequencing, pinnipeds, population genetics, single nucleotide polymorphisms

1 | INTRODUCTION

Population genetic analyses have become an integral part of conservation studies. These analyses allow the quantification of parameters relevant to endangered populations such as genetic diversity, effective population sizes, gene flow and the inference of past population histories (Beaumont, Zhang, & Balding 2002; Lopes & Boessenkool, 2010). Identifying management units and determining the degree of connectivity among demes are often the first steps in allocating conservation resources and developing management strategies (Moritz, 1999; Palsbøll, Bérubé, & Allendorf, 2007; Schwartz, Luikart, & Waples, 2007). For example, species with a relative lack of population structure that show little variation across their ranges benefit from management of all subpopulations as a single unit (Palsbøll et al., 2007). Conversely, species characterized by genetically-distinct, geographically isolated demes require management of distinct units to maximize maintenance of the species' evolutionary potential (Lowe & Allendorf, 2010; Palsbøll et al., 2007).

While describing population structure is crucial for species management, considering the underlying causes of this structure and discriminating between events of evolutionary significance and those with anthropogenic causes is also critical (e.g., rock wren; Weston & Robertson, 2015). For instance, if genetically distinct populations have diverged gradually as the result of glacial isolation or post-glacial colonization (e.g., kea; Dussex, Wegmann, & Robertson, 2014), it is often considered best to work toward maintaining that existing population structure. Conversely, for species that have undergone severe reductions in population size and genetic population fragmentation (e.g., mohua; Tracy, Wallis, Efford, & Jamieson, 2011), restoration of populations to better reflect the population structure prior to human disturbance may be more beneficial (Moritz, 1994; Palsbøll et al., 2007).

Recent advances in next-generation sequencing (NGS) allow the characterization of large numbers of molecular markers, increasing the resolution of comparisons of individuals within and between populations (Helyar et al., 2011). Single nucleotide polymorphisms (SNPs) have become the markers of choice for an increasing number of modern population genetic studies (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Helyar et al., 2011), despite initial speculation on their relative lack of genetic diversity compared to individual microsatellites (Rosenberg, Li, Ward, & Pritchard, 2003). While 100 SNPs may only be as informative as 10–20 microsatellites (Kalinowski, 2002), tens of thousands of SNPs can be discovered via NGS, even for non-model species (Ellegren, 2014; Seeb et al., 2011). The degree of definition afforded by such large marker datasets therefore statistically outperforms studies in which a few dozen microsatellites were used to describe population structure (Helyar et al., 2011; Liu, Chen, Wang, Oh, & Zhao, 2005).

A growing number of studies have also explored the specific applications of reduced representation sequencing (RRS) methods in evolutionary biology (e.g., RADseq; Andrews et al., 2016). One of these methods, referred to as Genotyping-by-sequencing (GBS; Elshire et al., 2011) has been used in a range of non-model species

as a means of inferring genotypes and describing population genomic structure (Johnson et al., 2015; Skovrind et al., 2016). Species for which traditional markers (i.e., microsatellite, single mtDNA sequences) have shown limited resolution in the description of population structure (e.g. Dussex et al., 2016; Robertson & Gemmill, 2005) could benefit from RRS methods, due to the higher definition and overall performance of the resulting marker sets. For instance, species characterized by strong dispersal potential and high gene flow are interesting candidate species for re-evaluations of population structure analyses using RRS.

The New Zealand fur seal (NZFS, *Arctocephalus forsteri*) has undergone a rapid recolonization following near extinction. Prior to the arrival of humans, NZFS were distributed throughout the east and west coasts of New Zealand's North and South Islands (Smith, 1989, 2005). Pre-historic Polynesian settlers hunted NZFS as a source of meat, which resulted in a rapid population decline (Emami-Khoyi et al., 2017; Smith, 1989, 2005) and the near elimination of the mainland fur seal population (Lalas & Bradshaw, 2001). By the time of European arrival in New Zealand in the late 18th Century, the NZFS mainland breeding range was confined to the south-western South Island (Smith, 1989). Sealing by European settlers for the acquisition of furs had a further devastating impact on NZFS and resulted in a depletion of the NZFS population (Lalas & Bradshaw, 2001). Overall, an estimated 1.5 million skins were harvested between 1792 and 1849 around Australia's southern coast, New Zealand, and at the adjacent subantarctic islands (Ling, 1999), suggesting that the number of NZFS before sealing may have been as high as 1–2 million (Baird, 2011). However, because records of the number of fur seals killed during that time are incomplete, inaccurate, or deliberately misleading to protect commercial interest of the day, it is unclear how many seals may have remained in New Zealand in the 1830s, when sealing stopped due to low profitability (Lalas & Bradshaw, 2001).

Since implementation of regulations in the late 19th century and the official cessation of commercial sealing in 1946 (Sorensen, 1969), the NZFS has rapidly recolonized much of its former range, and has re-established breeding colonies in the South Island of New Zealand (Baird, 2011; Bradshaw, Lalas, & Thompson, 2000; Crawley & Wilson, 1976). However, only one breeding colony, founded in the early 1990s, has been re-established in the North Island (Dix, 1993). Many of these colonies are showing steady population growth (i.e., 20%–25% annual growth) (Lalas & Bradshaw, 2001; Boren, Muller, & Gemmill, 2006; Figure 1). The current NZFS population is estimated at 50,000–100,000 animals although there is great uncertainty around this number (Baird, 2011).

Population declines and fragmentation often lead to population differentiation via the effect of genetic drift when demographic units within a larger panmictic population become fixed for different alleles (Charlesworth, 2009; Frankham, 2005). Alternatively, genetic differentiation can arise via founder effects following the recolonization of a species' former range (Ramachandran et al., 2005; Templeton, 1980). However, previous studies on pinniped species with histories of intense exploitation have shown a lack of (e.g., *Arctocephalus gazella*, Hoffman, Grant, Forcada, & Phillips, 2011)

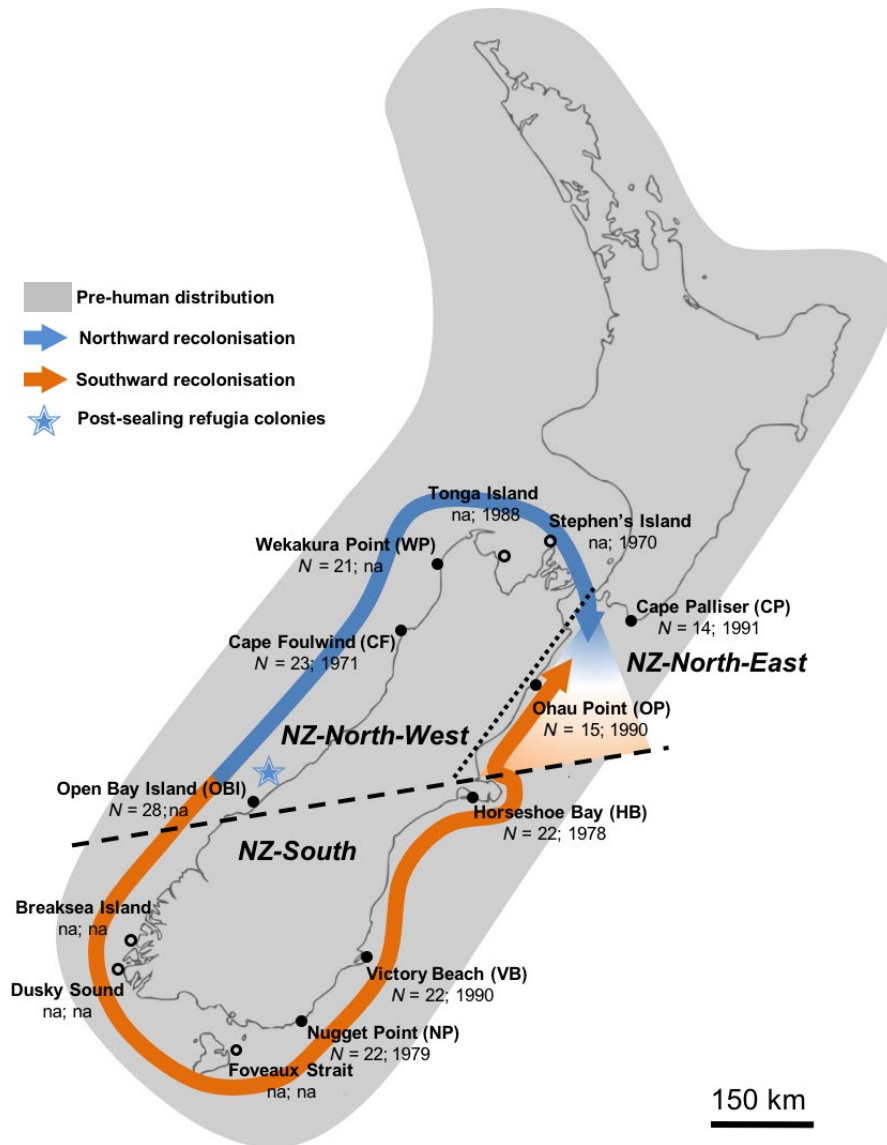


FIGURE 1 Extant New Zealand fur seal (*A. forsteri*) breeding colonies with full and empty dots depicting colonies sampled and not sampled in this study, respectively, with the number of samples analysed in this study and the year of first breeding from King (2005). Cape Palliser (CP) is currently the only known breeding colony in the North Island. The dashed line represents the genetic clusters based on Dussex et al. (2016) and Salis et al. (2016) microsatellite and mtDNA results, and the dotted line represents the approximate limit of the admixture zone between the two main clusters. Arrows depict the post-sealing recolonization scenario from hypothetical south-western South Island refugia depicted with a star inferred by ABC analyses (Dussex et al., 2016; this study)

or only moderate fine-scale structure (e.g., *Arctocephalus tropicalis*, Wynen et al., 2000; *Halichoerus grypus*, Fietz et al., 2016). Research based on modern and ancient DNA suggests that, while there was a shift in the genetic composition of NZFS in southern New Zealand resulting from the demographic decline, this decline may not have been long enough to induce a significant loss of genetic diversity as evidenced by comparable levels of genetic diversity in modern and ancient populations (Salis et al., 2016). This result is in stark contrast with the severe demographic decline and near range-wide extirpation in the 1800s. Moreover, in spite of potential founder effects resulting from recolonization from a few “refugia” colonies, studies based on microsatellite data have suggested that any population differentiation would have been transient, owing to the high vagility of the species combined with a rapid recolonization time (Figure 1, Dussex et al., 2016).

Studies aimed at assigning NZFS individuals to their colonies or broader regions of origin using microsatellite markers and classical assignment approaches have also had limited success (Dussex et al.,

2016; Robertson & Gemmill, 2005). Genetic distinction of potential management units has thus been difficult. A prevailing view from the literature published on NZFS genetics to date is that more numerous and/or variable genetic markers could significantly aid in resolving structure and improving assignment probabilities (Dussex et al., 2016; Robertson & Gemmill, 2005)—using an RRS approach provides the opportunity to empirically test this prediction.

Today, NZFS are the most commonly caught marine mammal in New Zealand commercial trawl net fisheries (Thompson & Abraham, 2010). Between 2003 and 2013 there were an estimated 501–1,273 fur seals caught per fishing season (Ministry for Primary Industries 2014; Thompson & Abraham, 2010) and males accounted for about 45% of reported bycaught individuals (Baird, 2011). Bycatch incidents seem to occur in clusters throughout New Zealand waters, mostly off the northern East (~40%) and West Coasts of the South Island and around offshore islands (Thompson & Abraham, 2010). While it is possible that bycatch may affect some colonies more than others, there is a high degree of uncertainty on the localized impact

of commercial fisheries on the species. West Coast colonies in particular face the challenges of lower habitat accessibility (Smith, 1989, 2005), climatic shifts that affect food availability (Bradshaw, Lalas, et al., 2000) and potentially higher conflict with fisheries (Ministry for Primary Industries, 2014). Moreover, the possibility of sexual segregation and niche divergence as seen in the country's other native pinniped, the New Zealand sea lion (*Phocarctos hookeri*) (Leung, Chilvers, Nakagawa, Moore, & Robertson, 2012), could result in sex-biased mortality which would then negatively impact population dynamics. This means that in spite of an overall trend in steady growth, these effects may negatively affect colony viability. The ability to assign bycaught individuals to their colonies of origin could thus help identify the colonies most at risk of decline.

In this study, we use SNP markers discovered via GBS to examine the population structure in NZFS and compare the markers' performance with the traditional markers used in previous studies on the same species (Dussex et al., 2016; Lento, Haddon, Chambers, & Baker, 1997; Robertson & Gemmell, 2005). We also assess the power of GBS-generated SNPs to assign potential bycatch to their colony by estimating the proportion of pups that can be correctly assigned to their known colony and/or regions of origin using a GBS SNP dataset. We hypothesize that in spite of the high individual vagility and rapid recolonization potential of the species, GBS-generated SNPs will allow the detection of fine-scale population differentiation among colonies. Moreover, we hypothesize that the SNPs presented here will outperform previously used microsatellite markers and produce more confident assignment probabilities.

2 | MATERIALS AND METHODS

2.1 | Sampling

Between 1998 and 2003, we obtained 169 NZFS skin samples from eight mainland NZFS breeding colonies (Figure 1). Only pups were sampled because they are a direct representation of each breeding colony, and to avoid bias caused by seasonal and yearly variation in adult dispersal. For each pup sampled, a small piece of skin was taken from the tip of a digit on a hind flipper using piglet ear notch pliers as per an established protocol (Majluf & Goebel, 1992) and stored in 70% ethanol.

2.2 | DNA extraction, library preparation, and sequencing

DNA was extracted from fur seal skin samples using a modified version of a lithium chloride DNA extraction protocol (Gemmell & Akayima, 1996). DNA quality and concentration were evaluated using both a Nanodrop 2000c spectrophotometer (Thermo Fischer Scientific) and a Fluorometric Quantitation Qubit system (Thermo Fischer Scientific). For library preparation, we used GBS as a sequencing preparation protocol (Elshire et al., 2011). We used two 96-well plates for library construction. Prior to

adapter ligation, an automated picogreen assay (Thermo Fisher Scientific) was performed to normalize the DNA concentration in all samples to 20 ng/μl. The restriction enzyme *Pst*1 was used due to its empirical success for sequencing of vertebrates and its capacity to produce libraries with high depth of coverage (De Donato, Peters, Mitchell, Hussain, & Imumorin, 2013). 17 μl H₂O, 2 μl CutSmart buffer, 0.5 μl NEB buffer, 0.5 μl *Pst*1 were added to wells containing DNA and adapters. A solution containing ligase buffer, ATP, and T4 ligase was then added to each well to aid in adhering adapters to sticky ends of the now-cut fragments. This combination of DNA, enzyme, adapters, and ligation solution was incubated at 22°C for 1 hr, and the temperature was then increased to 65°C for 30 min to inactivate T4 ligase (Elshire et al., 2011). 5 μl from each well were pooled together into one solution containing all fragments, each of which was identifiable by its respective barcode adapter. Fragments were amplified in three subsequent polymerase chain reactions and fragment length (in base pairs) was examined on a bioanalyser (Agilent 2100 bioanalyser, Santa Clara, CA, USA). The amplified library was purified using a Pippin Prep (SAGE Science, Beverly, MA, USA) to select the DNA sequencing library in the size range of 150–500 bp. Finally, single-end sequencing (1 × 100) was performed on an Illumina HiSeq2500 using v4 chemistry, which yielded approximately 25 Gb of raw sequence data per lane.

2.3 | SNP calling and filtering

SNP calling was performed in Stacks v. 1.30 (Catchen et al., 2011). We used “*process_radtags*” to demultiplex the data and discard low quality data with a probability of sequencing error >0.10% (Phred score = 10). SNP loci were then called using the Stacks “*denovo_map.pl*” pipeline (i.e., without a reference genome), using a minimum stack depth (-m) of 5 and default parameters for mismatches between loci within samples (-M = 2) and when building the catalog (-n = 1). The “populations” program in Stacks was used to filter out monomorphic loci and sort individuals into separate files for each colony. 180,303 biallelic SNPs remained after this filtering. PLINK 1.9 (Purcell et al., 2007) was then used to exclude individuals with >90% missing data (-mind filter), SNPs with >20% missing call rates (-geno filter), and SNPs with minor allele frequency <5% (-maf filter). SNPs significantly deviating (exact test) from Hardy-Weinberg equilibrium (HWE) (-hardy filter) and pairs of SNPs in linkage disequilibrium (LD) (-ld filter) in at least six of our eight populations were removed. For downstream analyses, we used the whole set of SNPs generated in PLINK, but also generated a separate subset of 5,000 and 1,000 random SNPs for computationally intensive analyses (e.g., STRUCTURE, DIYABC). Finally, the resulting Ped files were converted into other file formats with the program PGDSPIDER 2.0 (Lischer & Excoffier, 2012), enabling their input into relevant population genetic analysis software.

2.4 | Genetic diversity and population structure

For our analyses of population structure, we evaluated genetic variation at the colony and genetic cluster levels. Three main clusters were defined based on previous NZFS genetic studies (Dussex et al., 2016; Salis et al., 2016) as well as the results presented here: NZ-North-West (OBI: Open Bay Island; CF: Cape Foulwind; WP: Wekakura Point), NZ-North-East (CP: Cape Palliser; OP: Ohau Point), previously identified as a zone of admixture (Dussex et al., 2016) and NZ-South (East Coast to south coast colonies; HB: Horseshoe Bay; VB: Victory Beach; NP: Nugget Point) (Figure 1).

Genetic diversity (H_o , H_e and F_{IS}) was estimated for each colony and region using Arlequin v. 3.5 (Excoffier & Lischer, 2010). Pairwise F_{ST} was calculated among the eight colonies in Arlequin v. 3.5 using Slatkin's linearized F_{ST} (Slatkin, 1995) and tested for Isolation by distance (IBD) using a Mantel test implemented in GenAIEx (Peakall & Smouse, 2006). We then tested for correlations between estimates of H_e and F_{ST} derived from microsatellite loci by Dussex et al. (2016) and the biallelic SNPs genotyped here. For H_e , we calculated the Pearson correlation coefficient ("cor.test" in R; R Development Core Team 2012) and for F_{ST} , we used a Mantel test for matrix correspondence for correlation ("mantel.rtest" in R; R Development Core Team 2012).

We used the Bayesian cluster methodology implemented in STRUCTURE 2.2 (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000) to make inferences regarding the number of population clusters in our sample of NZFS using a random subset of 5,000 SNPs generated with PLINK 1.9 (Purcell et al., 2007). We used an admixture model with correlated allele frequencies and performed 10 iterations (chain length 500,000 steps, burn-in = 200,000 steps) for each K (from 1 to 8), using the LOCPRIOR option to assist with clustering of faint-structure data (Hubisz, Falush, Stephens, & Pritchard, 2009). The number of distinct genetic clusters were inferred using the $\text{LnP}(K)$ value of STRUCTURE 2.2 and the ΔK (Evanno, Regnaut, & Goudet, 2005) method implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012). CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) was then used to estimate individuals' assignment coefficient (q) to each genetic cluster and to visualize the results. As a comparison, we also used the program ADMIXTURE (Alexander, Novembre, & Lange, 2009) to estimate the number of clusters for each K (from 1–8). This program estimates ancestry in a model-based manner where individuals are considered unrelated and uses a cross-validation procedure to determine the best number of possible genetic groups present in the dataset.

STRUCTURE and ADMIXTURE results can be biased in nonequilibrium populations and result in an upward bias in the estimation of K because the program relies on assumptions of HWE and linkage equilibrium within populations for each K tested (Kaeuffer, Réale, Coltman, & Pontier, 2007). The population structure was therefore also examined independently using a discriminant analysis of principal components (DAPC; Jombart, Devillard, & Balloux, 2010), in *adegenet* (Jombart & Ahmed, 2011) for R (R Development Core Team

2012). DAPC has been shown to perform as well as STRUCTURE and has the advantage of being unconstrained by the assumptions of HWE (Jombart et al., 2010). DAPC is similar to principal component analysis, but unlike PCA, which maximizes the total variation in the dataset, DAPC maximizes the variation among different groups and minimizes variation within groups (Jombart et al., 2010). This analysis was performed with prior information on individual colonies. Finally, analyses of molecular variance (AMOVA; Excoffier, Smouse, & Quattro, 1992) were performed in Arlequin v. 3.5 for (a) a two-cluster grouping: NZ-North (OBI, NP, WP, CP, OP) vs. NZ-South (HB, VB, NP), and (b) a three-cluster grouping: NZ-North-West (OBI, NP, WP) vs. NZ-North-East (CP, OP) vs. NZ-South (HB, VB, NP).

2.5 | Demographic history

Previous studies have established several aspects of NZFS population history using Approximate Bayesian Computation (ABC; Beaumont, Zhang, & Balding 2002) approaches with microsatellite (Dussex et al., 2016) and mitogenome data (Emami-Khoyi et al., 2017). Here, we did not attempt to reiterate the demographic model comparisons described in these studies, but rather to estimate demographic parameters based on SNP data for the most likely scenarios inferred in these studies. We thus designed a scenarios including a Polynesian-induced decline (~500–1,000 years ago; Emami-Khoyi et al., 2017), and a European-induced decline (~125–250 years ago; Dussex et al., 2016) followed by range-wide recolonization from a single refugium and subsequent population divergence with a NZ-North (OBI, CF, WK, CP, OP) and NZ-South cluster (HB,VB,NP) (Dussex et al., 2016; Salis et al., 2016) (Supporting Information Figure S1, Table 1). We also designed a null model of constant effective population size though time followed by the same population divergence. The generation time for NZFS is estimated at 4–6 years in females and at 5–9 years in males (Dickie & Dawson, 2003). Thus, we assumed an average 7 year generation time for NZFS.

Because DIYABC is a computationally intensive method, analyses were performed using a subset of 1,000 randomly selected SNPs from our complete dataset. We generated 10^6 simulations for each scenario. As summary statistics, for single sample statistics we used the mean gene diversity across polymorphic loci and mean gene diversity across all loci (Nei, 1987). For two sample statistics, we used mean of nonnull F_{ST} distances between the two samples across loci, the mean of F_{ST} distances between the two samples across all loci, mean of nonnull Nei's distances between the two samples across loci, and mean of Nei's distances between the two samples across all loci (Nei, 1972). We used the standard Hudson's (2002) algorithm for selection of minimum allele frequency (MAF).

The posterior probability of each scenario was estimated using both the direct and logistic regression approaches (Fagundes et al., 2007). The ten thousand datasets (1%) with the smallest Euclidean distances were then retained to build posterior parameter distribution. To check model performance, we first empirically evaluated the power of the model to discriminate among scenarios (confidence in scenario choice). The approach consists of simulating

TABLE 1 Prior and posterior distributions of parameters for a demographic scenario including a Polynesian- and European-induced decline followed by recolonization and population divergence that obtained the highest posterior probability (PP = 1) when comparing two demographic scenarios in New Zealand fur seal (*A. forsteri*). Timing of events is in generations, assuming a generation time of 7 years (Dickie & Dawson, 2003). Type I and II errors were of 0.04 and 0.05, respectively

Parameter	Prior distribution	Posterior mode	5%	95%
N_{e-NZ-S}	Uniform [10^2 – 10^4]	1,120	483	7,940
N_{e-NZ-N}	Uniform [10^2 – 10^4]	652	336	6,600
$N_{e-refugium}$	Uniform [10^2 – 2.5×10^3]	110	102	161
$N_{e-historical-NZ}$	Uniform [10^3 – 5×10^6]	572,000	228,000	4,750,000
$N_{e-pre-historical}$	Uniform [10^3 – 5×10^6]	4,900,000	356,000	4,810,000
t -Polynesian	Uniform [100–200]	106	105	194
t -Europeans	Uniform [25–50]	49.2	30.2	49
t -post-seal	Uniform [1–25]	20.3	5.89	24.2

Conditions: t -Europeans > t -postseal; $N_{e-pre-historical} > N_{e-NZ-N}$; $N_{e-pre-historical} > N_{e-NZ-S}$

pseudo-observed datasets with parameters drawn from the posterior parameter distribution of the considered scenario and positioning the summary statistics of the observed data in the summary statistic distribution of these pseudo-observed data. The scenario is then considered suitable if the observed data summary statistics are included in the confidence interval drawn from pseudo-observed data. We calculated statistical measures of performance and Type I and Type II error rates as a means of model checking (Cornuet, Ravigné, & Estoup, 2010; Excoffier, Estoup, & Cornuet, 2005).

2.6 | Outlier loci detection

To maximize assignment power to the colonies sampled, we attempted to identify colony-specific and genetic clusters-specific sets of outlier loci that showed high differentiation between colonies and genetic clusters. These analyses were performed for each of the eight colonies and for the three genetic clusters.

We identified outlier loci using two different methods. First, we ran Fdist2 (Beaumont & Nichols, 1996) with 50,000 simulations, a confidence interval of 0.99, and false discovery rate (FDR) of 0.05 in LOSITAN (Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008). To correct for multiple testing, we transformed the p -values obtained to q -values (i.e., the false discovery rate FDR, analog to the p -value; Benjamini & Hochberg, 1995). Secondly, we used a Bayesian simulation-based test (Beaumont & Balding, 2004) that has been further refined and implemented in the software Bayescan 2.0 (Foll & Gaggiotti, 2008). We based our analyses on 10-pilot runs each consisting of 5,000 iterations, followed by 100,000 iterations with a burn-in of 50,000 iterations. We used a FDR of 0.05 as the threshold for outlier locus detection in this test.

2.7 | Population assignments

To assess the power of GBS-generated SNPs to assign potential NZFS bycatch to their colony or genetic cluster, we treated pups as unknown in origin and attempted to assign them back to their home colonies and genetic clusters. We first performed assignments of pups to the eight colonies using all 26,026 SNPs, a random subset

of 5,000 SNPs and the 74 colony-specific outlier loci. Secondly, we performed assignments of pups to the three genetic clusters using all 26,026 SNPs, the same random subset of 5,000 SNPs and 18 genetic cluster-specific outlier loci (see Results). This made for a total of six analyses.

We used Genodive 2.0b23 (Meirmans & Van Tienderen, 2004) to assign pups to their colony or cluster. The program uses the leave-one-out (LOO) validation procedure in which an individual to be assigned is removed from its source population and treated as an individual of unknown population origin before calculation of the population allele frequency. The purpose of the LOO procedure is to remove a bias present when allele frequencies are calculated from the same individuals that are subsequently assigned. We used the home likelihood criterion (the likelihood that an individual comes from the population where it was sampled) because it is more appropriate when only part of all possible source populations have been sampled (Meirmans & Van Tienderen, 2004). We used a significance threshold of 0.05 and replaced zero frequencies by 0.005, as suggested by Meirmans and Van Tienderen (2004) and generated 10,000 permutations (i.e., number of datasets). Missing values were replaced by randomly picking alleles from the global allele pool. Assignment power was estimated by calculating the proportion of pups that were correctly assigned to their colony or cluster of origin using (a) the entire dataset of 26,026 SNPs, (b) the subset of 5,000 randomly selected SNPs, and (c) the loci identified as outliers in LOSITAN and Bayescan 2.0.

3 | RESULTS

3.1 | Data filtering

Out of a total of 169 genotyped individuals, we excluded two individuals for having more than 90% of loci missing. Out of a total 180,303 biallelic SNPs, we excluded 97,339 SNPs for being missing in at least 80% of individuals, and 56,937 SNPs with <5% minor allele frequency. We further excluded a total of 3,834 SNPs because they were either out of HWE, or in LD. After applying these filters, our dataset consisted of 167 individuals with data for 26,026 SNPs.

3.2 | Genetic diversity and population structure

Measures of heterozygosity were similar across colonies and regions (Table 2) and F_{IS} values were low and only significantly different from zero in one population (WP). Overall, F_{ST} values were low (0.003–0.022) and mostly significantly different from zero ($p < 0.05$) between each pair of colonies (27 out of 28; Table 3). NZ-North-West (OBI, NP, WP) colonies and NZ-South (HB, VB, NP) colonies had lower pairwise F_{ST} values when compared with each other than with other regions (Table 3). The NZ-North-East (CP, OP) colony Ohau Point (OP) had relatively low pairwise F_{ST} values when compared with every other colony in this study, especially the NZ-South colonies. Conversely, Cape Palliser (CP), had relatively high pairwise F_{ST} values when compared with every other colony except Ohau Point. The Mantel test revealed a weak, but significant IBD signal ($p = 0.01$, Supporting Information Figure S2). Neither expected heterozygosity estimates nor pairwise F_{ST} values were correlated between the microsatellite and SNP datasets (Pearson's $r = 0.25$, $p = 0.55$; Supporting Information Figure S3a; Mantel's $r_{xy} = 0.38$, $p = 0.11$; Supporting Information Figure S3b).

Based on LnP(K) and the ad-hoc statistic ΔK , Bayesian clustering analysis for 167 NZFS implemented in STRUCTURE identified two main clusters along a west to east genetic gradient (Figure 2, Supporting Information S4a, S5). NZ-North-West (OBI, CF, WK) and NZ-South (HB, VB, NP) colonies showed low admixture ($q \geq 0.83$ and 0.79, respectively) while NZ-North-East colonies (CP and OP) showed more pronounced admixture ($q = 0.59$ and 0.51). While ADMIXTURE and STRUCTURE results were consistent in describing a west to east genetic gradient, ADMIXTURE supported a single cluster (i.e. lowest cross-validation error for $K = 1$, Supporting Information Figure S4b, S5).

Discriminant analysis of principal components results were consistent with STRUCTURE and also supported a NZ-North-West and a NZ-South cluster (Figure 3). NZ-North-East (CP and OP) colonies shared principal component similarity with both of these clusters, though CP was more divergent than OP from other colonies. AMOVAs based on STRUCTURE and DAPC results for a NZ-North (OBI, NP, WP, CP, OP) vs. NZ-South (HB, VB, NP) grouping (Supporting Information Table S1a) and a NZ-North-West (OBI, NP, WP) vs. NZ-North-East (CP, OP) vs. NZ-South (HB, VB, NP) grouping (Supporting Information Table S1b) both indicated that 98% of

TABLE 2 Sample sizes, average expected heterozygosity (H_e), average observed heterozygosity (H_o) and F_{IS} values with an asterisk depicting values significantly different from zero ($p < 0.05$) after 10,000 permutations for each New Zealand fur seal (*A. forsteri*) cluster and breeding colony

Cluster/Colony	Code	N	H_o	H_e	F_{IS}
NZ-North-West		72	0.225	0.259	0.025*
Open Bay Islands	OBI	28	0.233	0.267	0.021
Cape Foulwind	CF	23	0.237	0.268	0.009
Wekakura Point	WP	21	0.230	0.268	0.039*
NZ-North-East		29	0.238	0.263	0.026
Cape Palliser	CP	14	0.247	0.272	0.010
Ohau Point	OP	15	0.250	0.271	0.021
NZ-South		66	0.230	0.261	0.027*
Horseshoe Bay	HB	22	0.227	0.263	0.033
Victory Beach	VB	22	0.246	0.270	0.012
Nugget Point	NP	22	0.246	0.269	0.029

variance was explained by variation among individuals. Only ~0.60% was explained by variation among regions.

3.3 | Demographic history

Demographic reconstructions using the Approximate Bayesian Computation (ABC) approach strongly supported a scenario of Polynesian- and European-induced declines followed by range-wide recolonization and population divergence with a posterior probability of 1. Our estimates supported severe bottlenecks with potential 10-fold and 250-fold declines in response to Polynesian and European hunting, respectively (Table 1). The observed data summary statistics for the preferred model were included in the confidence interval drawn from pseudo-observed data. Type I and II errors were of 0.04 and 0.05, respectively.

3.4 | Outlier loci detection

Using the Fdist2 method implemented in LOSITAN, we did not find any outlier loci under putative selection. The Bayescan method

	OBI	CF	WP	CP	OP	HB	VB	NP
OBI	0							
CF	0.005	0						
WP	0.005	0.006	0					
CP	0.022	0.022	0.021	0				
OP	0.003	0.010	0.009	0.020	0			
HB	0.016	0.017	0.017	0.023	0.010	0		
VB	0.013	0.014	0.014	0.020	0.011	0.006	0	
NP	0.015	0.016	0.014	0.019	0.011	0.004	0.005	

Note. Bold values denote significant values following standard Bonferroni correction ($p < 0.05/n$ tests).

TABLE 3 Pairwise values of Slatkin's linearized F_{ST} between New Zealand fur seal (*A. forsteri*) breeding colonies



FIGURE 2 Individual clustering assignment in New Zealand fur seal (*A. forsteri*) ($n = 167$) from eight breeding colonies for the most likely number of clusters ($K = 2$) using 5,000 random SNPs in STRUCTURE

identified 74 colony-specific and 18 genetic cluster-specific outlier SNP loci putatively under diversifying selection when analysing the six colonies and East and West Coast clusters, respectively

(Supporting Information Table S2). For both datasets, all other loci were considered putatively neutral.

3.5 | Population assignments

We attempted to assign pups treated as unknowns to colonies and identified clusters using the LOO test. Using either all 26,026 SNPs or the 5,000 randomly selected SNPs, we assigned 167 pups to their colonies of origin with, on average, 44.2% and 42.1% accuracy, respectively, though low assignment probabilities to NZ-North-East colonies CP (14%–21%) and OP (0%–7%) drove these averages down (Table 4a). We observed the highest assignment probabilities at the colony level for OBI with 96% and 93% correct assignments for the 26,026 SNPs and 5,000 SNPs, respectively. Using outlier loci increased our colony-level assignment probabilities by around 10%. Average assignment probabilities to colonies using 74 colony-specific outlier loci were higher on average (53.3%) when considering all colonies and even higher (69%) when excluding CP and OP, which both showed the lowest assignment probabilities (Table 4a).

Individual assignment probabilities of pups to the three genetic clusters (NZ-North-West, NZ-North-East and NZ-South) were on average higher (64%–69%; Table 4b) than for assignments at the colony level. As expected, and in line with assignments probabilities to colonies, assignment probabilities were higher for the NZ-North-West (80%–97%) and NZ-South (77%–95%) clusters than for the highly-admixed NZ-North-East cluster (13%–34%; Table 4b) across all SNP datasets. However, the assignment probability to the admixed colonies was highest (34%) when using the 18 outlier loci (Table 4b).

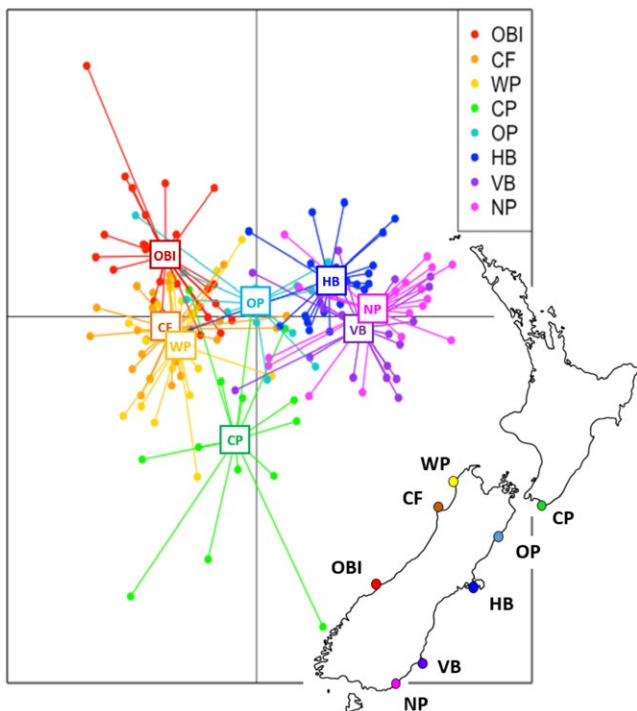


FIGURE 3 DAPC for 167 New Zealand fur seal (*A. forsteri*) from eight breeding colonies based on 26,026 SNP loci

TABLE 4 Proportions of New Zealand fur seal (*A. forsteri*) pups correctly assigned to (a) their colonies or (b) genetic clusters inferred from Bayesian clustering analyses. Assignments were performed independently using all 26,026 SNPs, 5,000 randomly selected SNPs, 74 outlier loci identified at the colony level, and 18 outlier loci identified at the genetic cluster level

Colony	Code	N	26K SNPs	5K SNPs	74 Outlier SNPs
(a)					
Open Bay Islands	OBI	28	0.96	0.93	0.68
Cape Foulwind	CF	23	0.30	0.26	0.61
Wekakura Point	WP	21	0.43	0.43	0.67
Cape Palliser	CP	14	0.21	0.14	0.00
Ohau Point	OP	15	0.00	0.07	0.07
Horseshoe Bay	HB	22	0.50	0.45	0.73
Victory Beach	VB	22	0.41	0.50	0.68
Nugget Point	NP	22	0.73	0.59	0.82
Average		167	0.442	0.421	0.533
Average (CP,OP excluded)		138	0.555	0.527	0.698
Cluster		N	26K SNPs	5K SNPs	18 Outlier SNPs
(b)					
NZ-North-West		72	0.97	0.96	0.80
NZ-North-East		29	0.17	0.13	0.34
NZ-South		66	0.94	0.95	0.77
Average		167	0.69	0.68	0.64

4 | DISCUSSION

4.1 | Subtle population structure and near panmixia

Our panel of 26,026 SNPs illustrates the potential of RRS methods to identify previously undetected sources of variation in non-model species. Overall, our data strongly support a pattern of near panmixia in NZFS, but pairwise F_{ST} values among range-wide colonies supported a weak but significant west to east genetic gradient. Bayesian clustering and DAPC results support the existence of two main genetic clusters (NZ-North-West and NZ-South), and further suggest that NZ-North-East colonies (CP and OP) are sites of genetic admixture between these two main clusters. This observed structure was further supported by constructing a PCA from a matrix of estimated genomic relationships (Patterson, Price, & Reich, 2006), which in turn were calculated following the bioinformatics and statistical methods in Dodds et al. (2015). This approach allows the PCA to be formed when there are missing genotype calls and without the use of stringent filtering as the method directly accounts for the read depth and missing observations and thus gives further support to our results.

The weak IBD pattern identified in the Mantel test is consistent with previous observations of highly migratory pinniped species undergoing post-bottleneck recolonizations (e.g., *Arctocephalus gazella*, Hoffman et al., 2011), as well as of other marine mammals with comparable vagility (e.g., dusky dolphin, Harlin, Markowitz, Baker, Würsig, & Honeycutt, 2003; common dolphin, Stockin, Amaral, Latimer, Lambert, & Natoli, 2014). However, it is crucial to

remain cautious when interpreting such a weak IBD pattern, especially when the underlying F_{ST} values are generated with tens of thousands of SNPs. As the significance of F_{ST} values is likely a result of the size and power of the dataset (Helyar et al., 2011; Morin, Martien, & Taylor, 2009), we stress that various lines of evidence in the form of additional analyses should be considered when interpreting such data, especially when the pattern observed is not clear-cut and where there is ample room for interpretation. Here, we achieve this goal using STRUCTURE, ADMIXTURE, DAPC, AMOVA and genetic assignments and conclude that NZFS breeding colonies are not markedly genetically different from each other.

4.2 | Performance of SNPs loci

H_e and H_o values did not differ substantially among colonies (0.225–0.250), which is consistent with previous results based on microsatellite data (Dussex et al., 2016; Robertson & Gemmill, 2005) with H_e and H_o between (0.67–0.79). Higher estimates of H_e derived from microsatellite markers may be a consequence of ascertainment bias caused by selecting the most polymorphic markers (Haas & Payseur, 2011; Queirós et al., 2015). However, there was no significant correlation between H_e or F_{ST} estimates derived from microsatellite and the SNPs generated here. The lack of correlation between H_e estimates could be due to low sample size (e.g., Fischer et al., 2017) while for pairwise F_{ST} , it could also be due to an ascertainment bias in the selection of microsatellite loci (Haas & Payseur, 2011; Ryyänen, Tonteri, Vasemägi, & Primmer, 2007).

In spite of this very subtle population structure, assignment testing results indicate that large SNP marker sets allow NZFS pups to be assigned to their geographic regions of origin with high confidence, but less so to specific colonies. The proportion of pups correctly assigned to genetic clusters (64%–69%) and to breeding colonies (42%–53%) were on average higher for all SNP datasets used here than for microsatellite markers (59.3% and 22.7% correctly assigned to clusters and colonies, respectively) (Dussex et al., 2016; Robertson & Gemmell, 2005). Furthermore, a random subset of 5,000 SNPs was roughly as effective as the entire 26,026 SNP dataset in achieving these assignments, while using outlier loci in some instances increased the proportion of pups correctly assigned to their colony or genetic cluster of origin. As expected, the proportion of correct assignments was lowest for the NZ-North-East cluster and its colonies (CP and OP), which is in line with the identification of a zone admixture between the NZ-North-West and NZ-South clusters.

Our higher assignment probabilities can likely be attributed to the higher statistical power of our dataset (Helyar et al., 2011; Morin, Luikart, & Wayne, 2004; Morin et al., 2009), although we were still unable to achieve confident assignment to specific colonies. One exception is that of OBI, which also had one of the highest proportion of correct assignments with microsatellite data (68.8%; Dussex et al., 2016). This result is in line with the suggestion that OBI may represent a refugium from sealing (Dussex et al., 2016; Salis et al., 2016; Smith, 2005) and thus a colony with a higher proportion of private alleles, thus increasing the proportion of pups correctly assigned. While colony-level assignment for species may be unfeasible with the tools presently available, our results suggest that outlier loci assignment techniques could enable confident assignment in the future (e.g., Nielsen et al., 2012). However, identifying highly differentiated outlier loci will be crucial for accurate assignments as genetic structure is predicted to become rapidly eroded due to the high vagility of the species.

As a whole, our measures of intra and interspecific population structure and assignments are consistent with what might be expected from the severe historic bottleneck and rapid range-wide recolonization NZFS experienced, a model which has been confirmed through Bayesian reconstruction here and in previous studies (Dussex et al., 2016; Emami-Khoyi et al., 2016).

4.3 | Demographic recovery and transient population structure

The subtle population structuring seen here for NZFS is most likely due to the high vagility of the species and its recolonization history (Dussex et al., 2016; Emami-Khoyi et al., 2017). Demographic reconstructions strongly supported a scenario of Polynesian- and European-induced declines followed by range-wide recolonization and population divergence with a posterior probability of 1, which is consistent with archaeological records (Smith, 1989) and the recent exploitation history of the species (Lalas & Bradshaw, 2001). We estimated that a very small population survived European sealing with an N_e of ~110 (90%HPD 102–161), which is consistent

with Emami-Khoyi et al. (2017) who estimated a bottleneck N_e of 116. However, we estimated much larger pre-European and pre-Polynesian N_e of ~570,000 and ~4,900,000, which is one to two orders of magnitude higher than previous estimates of historical N_e (Dussex et al., 2016; Emami-Khoyi et al., 2017). Our contemporary N_e estimates seem more realistic with an N_e of 600–1,000 predicted following demographic recovery after cessation of sealing. The estimate for the post-European surviving population seems extremely low, yet not unlikely based on the history of near extermination of this and other fur seal species in the 1830s (Lalas & Bradshaw, 2001). For instance, only 5 NZFS were reported in the Bounty Islands in the 1830s (Taylor 1982). On the mainland, it has been estimated, colonies could have been reduced to <2% of their original size over 25–35 years (Lalas & Bradshaw, 2001), making hunting unprofitable and the possibility of few remaining breeders very likely. Moreover, preliminary runs comparing a model including a genetic bottleneck associated with Europeans and an alternative model that did not consider a genetic bottleneck, supported the former model, suggesting that a genetic bottleneck did actually occur. One explanation for the very low N_e in this surviving population compared to the larger N_e in modern populations may be due to unsampled populations that might have contributed genetic diversity to the modern populations and possibly range expansions of other species and, as yet, undetected hybridization (Lancaster et al., 2006). In fact, the main breeding locations in New Zealand in the 1970s comprised the Fiordland coastline at the south-western corner of South Island and the Chatham Islands, as well as islands in the Foveaux Strait, between Stewart Island and the South Island (Crawley, 1990; Wilson, 1981). It is thus possible that gene flow from unsampled colonies in Foveaux Strait, or subantarctic islands, may have rapidly inflated the N_e of modern populations.

While theory predicts that demographic decline can result in population structuring via the effects of genetic drift (Charlesworth, 2009; Frankham, 2005), NZFS as well as other otariids do not show marked population structuring (Dussex et al., 2016; Hoffman et al., 2011). This could be due to the fact that the demographic bottleneck in NZFS was too short in duration to induce a significant loss of genetic diversity and strong population differentiation via genetic drift and to a high dispersal rate contributing to range wide gene flow and homogenization of allelic distribution. NZFS most likely survived in refugia post-exploitation, and quickly recolonized after sealing became unprofitable (Dussex et al., 2016; Emami-Khoyi et al., 2017). Therefore, the lack of marked population structuring is most likely due to the high vagility of the species and its recolonization history. It is even possible that the subtle structure detected here is transient and that the whole species may represent a panmictic unit within a few generations because population size has increased dramatically in the years since official the cessation of seal hunting (Baird, 2011). Consequently, genetic assignments may become more difficult in the future and may require more traditional approaches, such as tagging and mark-recapture to assign bycatch to colonies (e.g., Bradshaw, Davis, Lalas, & Harcourt, 2000).

5 | CONCLUSION

To date, few studies have compared the performance of microsatellite and SNPs for highly vagile species. Overall, while the genetic structure described here was not substantially different from that examined with microsatellite loci, a larger number of markers allowed for better assignments of pups to their genetic cluster and colony of origin. This was especially true for colonies that were outside of the zone of admixture.

Our results, however, suggest that because of the ascertainment bias associated with the selection of microsatellite loci, overestimation of population structuring is possible when using such markers. SNP loci are thus not only less biased but also offer more statistical power. Nevertheless, due to the species' recolonization history and high vagility, describing finer-scale structuring remains challenging.

Recent advances in whole genome sequencing present new avenues for research in NZFS. For instance, hunting may have impacted fur seal functional diversity, either via the effects of drift or via artificial selection pressures imposed by hunting. While the effective population size may not have been significantly affected, modern functional genetic diversity may not be representative of the prehuman functional diversity as the shift of genetic composition in southern New Zealand since human arrival suggests (Salis et al., 2016). Comparison of pre- and post-hunting functional genetic diversity would thus provide insights in the potential effects of hunting as artificial selection pressure.

ACKNOWLEDGMENTS

We thank Sara Ferreira, Joanne Gillum, Tracey van Stijn, and Rachael Ashby for assistance with DNA extraction, GBS, and Illumina sequencing. The research was funded by the University of Otago, the Allan Wilson Centre, a National Institute of Health (NIH) Minority Health Research Grant, and by the New Zealand Ministry for Business, Innovation, and Employment (MBIE) via its funding of the "Genomics for Production & Security in a Biological Economy" program (Contract ID C10X1306). W.R. Stovall acknowledges support from the University of Otago Masters scholarship.

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

N. Dussex designed research, performed population structure analyses, and wrote main body of text. H.R. Taylor advised on population genetic theory and lab procedures, analysed sequencing output, reviewed, and contributed to writing the paper. W. R. Stovall designed research, performed lab work and bioinformatics processing of data, performed population structure analyses, and wrote main body of text. K. Rutherford performed bioinformatics processing of data,

analysed sequencing output, and oversaw data management and archiving. K. G. Dodds advised on visualization of results, and assisted with post-sequencing data management. S. M. Clarke assisted with formulation of project, performed GBS and Illumina sequencing, and reviewed and contributed to writing the paper. N.J. Gemmell designed research, performed and oversaw field sampling, and supervised project. All authors contributed substantially to the writing and refinement of the final text.

DATA ARCHIVING STATEMENT

Raw sequence data, processed data and scripts are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.63nm407>.

ORCID

Nicolas Dussex  <http://orcid.org/0000-0002-9179-8593>

Helen R. Taylor  <http://orcid.org/0000-0001-7951-0772>

William R. Stovall  <http://orcid.org/0000-0003-1512-5415>

Ken G. Dodds  <http://orcid.org/0000-0002-9347-6379>

Neil J. Gemmell  <http://orcid.org/0000-0003-0671-3637>

REFERENCES

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*, 81–92. Nature Publishing Group. <https://doi.org/10.1038/nrg.2015.28>
- Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a *Fst*-outlier method. *BMC Bioinformatics*, *9*, 323. <https://doi.org/10.1186/1471-2105-9-323>
- Baird, S. J. (2011). New Zealand fur seals – summary of current knowledge. New Zealand Aquatic Environment and Biodiversity Report. In New Zealand Aquatic Environment and Biodiversity Report No. 72 (P. 50). Wellington, New Zealand.
- Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*, 2025–2035. <https://doi.org/10.1002/9780470710777>
- Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, *13*, 969–980. <https://doi.org/10.1111/j.1365-294X.2004.02125.x>
- Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *263*, 1619–1626. <https://doi.org/10.1098/rspb.1996.0237>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289–300.
- Boren, L. J., Muller, C. G., & Gemmell, N. J. (2006). Colony growth and pup condition of the New Zealand fur seal (*Arctocephalus forsteri*) on the Kaikoura coastline compared with other east coast colonies. *Wildlife Research*, *33*, 497–505. <https://doi.org/10.1071/WR05092>
- Bradshaw, C., Davis, L., Lallas, C., & Harcourt, R. G. (2000). Geographic and temporal variation in the condition of pups of the New Zealand

- fur seal (*Arctocephalus forsteri*): Evidence for density dependence and differences in. *Journal of Zoology*, 252, 41–51 University of Otago Library. <https://doi.org/10.1111/j.1469-7998.2000.tb00818.x>
- Bradshaw, C. J. A., Lalas, C., & Thompson, C. M. (2000). Clustering of colonies in an expanding population of New Zealand fur seals (*Arctocephalus forsteri*). *Journal of Zoology*, 250, 105–112. <https://doi.org/10.1111/j.1469-7998.2000.tb00581.x>
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J. H., & De Koning, D.-J. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes|Genomes|Genetics*, 1, 171–182. <https://doi.org/10.1534/g3.111.000240>
- Charlesworth, B. (2009). Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10, 195–205. <https://doi.org/10.1038/nrg2526>
- Cornuet, J.-M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, 11, 401. <https://doi.org/10.1186/1471-2105-11-401>
- Crawley, M. C. (1990). Suborder Pinnipedia. In King, C. M. *The handbook of New Zealand mammals* (pp. 243–280). Oxford, UK: Oxford University Press.
- Crawley, M. C., & Wilson, G. J. (1976). The natural history and behaviour of the New Zealand fur seal (*Arctocephalus forsteri*). *Tuatara*, 22, 1–29.
- De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-Sequencing (GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLoS ONE*, 8, e62137. <https://doi.org/10.1371/journal.pone.0062137>
- Dickie, G. S., & Dawson, S. M. (2003). Age, growth, and reproduction in New Zealand fur seals. *Marine Mammal Science*, 19, 173–185. <https://doi.org/10.1111/j.1748-7692.2003.tb01100.x>
- Dix, B. (1993). A new record this century of a breeding colony in the north island for the new zealand fur seal *Arctocephalus forsteri*. *Journal of the Royal Society of New Zealand*, 23, 1–4. <https://doi.org/10.1080/03036758.1993.10721212>
- Dodds, K., Mcewan, J., Brauning, R., Anderson, R., Van Stijn, T., Kristjánsson, T., ... Clarke, S. M. (2015). *Calculation and uses of relatedness estimated from genotyping-by-sequencing (GBS) data Estimating relatedness from genotypes*. P. in New Zealand Molecular Mapping Workshop 10, Abstract 12.
- Dussex, N., Robertson, B. C., Salis, A. T., Kalinin, A., Best, H., & Gemmell, N. J. (2016). Low spatial genetic differentiation associated with rapid recolonization in the New Zealand fur seal *Arctocephalus forsteri*. *Journal of Heredity*, 107, 581–592. <https://doi.org/10.1093/jhered/esw056>
- Dussex, N., Wegmann, D., & Robertson, B. C. (2014). Postglacial expansion and not human influence best explains the population structure in the endangered kea (*Nestor notabilis*). *Molecular Ecology*, 23, 2193–2209. <https://doi.org/10.1111/mec.12729>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology and Evolution*, 29, 51–63. <https://doi.org/10.1016/j.tree.2013.09.008>
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6, e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Emami-Khoyi, A., Hartley, D., Paterson, A., Cruickshank, R., Boren, L., Ross, J., ... Else, T. (2016). Mitochondrial DNA structure and colony expansion dynamics of New Zealand fur seals (*Arctocephalus forsteri*) around Banks Peninsula. *New Zealand Journal of Zoology*, 4223, 1–14. <https://doi.org/10.1080/03014223.2016.1179649>
- Emami-Khoyi, A., Paterson, A. M., Hartley, D. A., Boren, L. J., Cruickshank, R. H., Ross, J. G., ... Else, T.-A. (2017). Mitogenomics data reveal effective population size, historical bottlenecks, and the effects of hunting on New Zealand fur seals (*Arctocephalus forsteri*). *Mitochondrial DNA Part A*, 1394, 1–13. <https://doi.org/10.1080/24701394.2017.1325478>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., Estoup, A., & Cornuet, J. M. (2005). Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, 169, 1727–1738. Genetics Society of America. <https://doi.org/10.1534/genetics.104.036236>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin ver. 3.5: A new series of programs to perform population genetics analysis under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Excoffier, L., Smouse, P., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131, 479–491.
- Fagundes, N. J., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F. M., Bonatto, S. L., & Excoffier, L. (2007). Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 17614–17619. <https://doi.org/10.1073/pnas.0708280104>
- Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164, 1567–1587.
- Fietz, K., Galatius, A., Teilmann, J., Dietz, R., Frie, A. K., Klimova, A., ... Olsen, M. T. (2016). Shift of grey seal subspecies boundaries in response to climate, culling and conservation. *Molecular Ecology*, 25, 4097–4112. <https://doi.org/10.1111/mec.13748>
- Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., ... Widmer, A. (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, 18, 1–15. <https://dx.doi.org/10.1186/s12864-016-3459-7>
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, 126, 131–140. <https://doi.org/10.1016/j.biocon.2005.05.002>
- Gemmell, N. J., & Akayima, S. (1996). An efficient method for the extraction of DNA from vertebrate tissues. *Trends in Genetics*, 12, 338–339. [https://doi.org/10.1016/S0168-9525\(96\)80005-9](https://doi.org/10.1016/S0168-9525(96)80005-9)
- Haas, R. J., & Payseur, B. A. (2011). Multi-locus inference of population structure: A comparison between single nucleotide polymorphisms and microsatellites. *Heredity*, 106, 158–171. <https://doi.org/10.1038/hdy.2010.21>
- Harlin, A. D., Markowitz, T. M., Baker, C. S., Würsig, B., & Honeycutt, R. L. (2003). Genetic structure, diversity, and historical demography of New Zealand's dusky dolphin (*Lagenorhynchus obscurus*). *Journal of Mammalogy*, 84, 702–717. [https://doi.org/10.1644/1545-1542\(2003\)084\[702:GSDAHD\]2.0.CO;2](https://doi.org/10.1644/1545-1542(2003)084[702:GSDAHD]2.0.CO;2)
- Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., ... Nielsen, E. E. (2011). Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. *Molecular Ecology Resources*, 11, 123–136. <https://doi.org/10.1111/j.1755-0998.2010.02943.x>

- Hoffman, J. I., Grant, S. M., Forcada, J., & Phillips, C. D. (2011). Bayesian inference of a historical bottleneck in a heavily exploited marine mammal. *Molecular Ecology*, 20, 3989–4008. <https://doi.org/10.1111/j.1365-294X.2011.05248.x>
- Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, 18, 337–338. <https://doi.org/10.1093/bioinformatics/18.2.337>
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 1–15. <https://doi.org/10.1186/1471-2156-11-94>
- Johnson, J. L., Wittgenstein, H., Mitchell, S. E., Hyma, K. E., Temnykh, S. V., Kharlamova, A. V., Gulevich, R. G., Vladimirova, A. V., Fong, H. W. F., Acland, G. M., Trut, L. M., & Kukekova, A. V. (2015). Genotyping-by-sequencing (GBS) detects genetic structure and confirms behavioral QTL in tame and aggressive foxes (*Vulpes vulpes*). *PLoS One*, 10, e0127013. <https://doi.org/10.1371/journal.pone.0127013>
- Kaeuffer, R., Réale, D., Coltman, D. W., & Pontier, D. (2007). Detecting population structure using STRUCTURE software: Effect of background linkage disequilibrium. *Heredity*, 99, 374–380. <https://doi.org/10.1038/sj.hdy.6801010>
- Kalinowski, S. (2002). How many alleles per locus should be used to estimate genetic distances? *Heredity*, 88, 62–65. <https://doi.org/10.1038/sj.hdy.6800009>
- King, C. M. (2005). *The handbook of New Zealand mammals*. New York, NY: Oxford University Press.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Lalas, C., & Bradshaw, C. J. A. (2001). Folklore and chimerical numbers: Review of a millennium of interaction between fur seals and humans in the New Zealand region. *New Zealand Journal of Marine and Freshwater Research*, 35, 477–497. <https://doi.org/10.1080/00288330.2001.9517017>
- Lancaster, M. L., Gemmill, N. J., Negro, S., Goldsworthy, S., & Sunnucks, P. (2006). Ménage à trois on Macquarie Island: Hybridization among three species of fur seal (*Arctocephalus* spp.) following historical population extinction. *Molecular Ecology*, 15, 3681–3692. <https://doi.org/10.1111/j.1365-294X.2006.03041.x>
- Lento, G. M., Haddon, M., Chambers, G. K., & Baker, C. S. (1997). Genetic variation of southern hemisphere fur seals (*Arctocephalus* spp.): Investigation of population structure and species identity. *Journal of Heredity*, 88, 202–208. <https://doi.org/10.1093/oxfordjournals.jhered.a023089>
- Leung, E. S., Chilvers, B. L., Nakagawa, S., Moore, A. B., & Robertson, B. C. (2012). Sexual segregation in juvenile New Zealand sea lion foraging ranges: Implications for intraspecific competition, population dynamics and conservation. *PLoS ONE*, 7, e45389. <https://doi.org/10.1371/journal.pone.0045389>
- Ling, J. K. (1999). Exploitation of fur seals and sea lions from Australian, New Zealand and adjacent subantarctic islands during the eighteenth, nineteenth and twentieth centuries. *Australian Zoology*, 31, 323–350. <https://doi.org/10.7882/AZ.1999.036>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28, 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Liu, N., Chen, L., Wang, S., Oh, C., & Zhao, H. (2005). Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. *BMC Genetics*, 6, 1–5.
- Lopes, J. S., & Boessenkool, S. (2010). The use of approximate Bayesian computation in conservation genetics and its application in a case study on yellow-eyed penguins. *Conservation Genetics*, 11, 421–433. <https://doi.org/10.1007/s10592-009-0032-9>
- Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, 19, 3038–3051. <https://doi.org/10.1111/j.1365-294X.2010.04688.x>
- Majluf, P., & Goebel, M. (1992). The capture and handling of female South American fur seals and their pups. *Marine Mammal Science*, 8, 187–190. <https://doi.org/10.1111/j.1748-7692.1992.tb00382.x>
- Meirmans, P. G., & Van Tienderen, P. H. (2004). GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4, 792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>
- Ministry for Primary Industries (2014). *Aquatic environment and biodiversity annual review 2014*. Wellington, New Zealand: Ministry for Primary Industries.
- Morin, P. A., Luikart, G., & Wayne, R. K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, 19, 208–216. <https://doi.org/10.1016/j.tree.2004.01.009>
- Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, 9, 66–73. <https://doi.org/10.1111/j.1755-0998.2008.02392.x>
- Moritz, C. (1994). Defining “evolutionarily significant units”. *Trends in Ecology and Evolution*, 9, 373–375. [https://doi.org/10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4)
- Moritz, C. (1999). Conservation units and translocations: Strategies for conserving evolutionary processes. *Heredity*, 130, 217–228.
- Nei, M. (1972). Genetic distance between populations. *American Naturalist*, 106, 283–292. <https://doi.org/10.1086/282771>
- Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press. [https://doi.org/10.1016/0047-2484\(88\)90032-2](https://doi.org/10.1016/0047-2484(88)90032-2)
- Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., ... Carvalho, G. R. (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications*, 3, 851–856. Nature Publishing Group. <https://doi.org/10.1038/ncomms1845>
- Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, 22, 11–16. <https://doi.org/10.1016/j.tree.2006.09.003>
- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2, e190. <https://doi.org/10.1371/journal.pgen.0020190>
- Peakall, R., & Smouse, P. E. (2006). GenAlEx version 5.0: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575. <https://doi.org/10.1086/519795>
- Queirós, J., Godinho, R., Lopes, S., Gortazar, C., de la Fuente, J., & Alves, P. C. (2015). Effect of microsatellite selection on individual and population genetic inferences: An empirical study using cross-specific

- and species-specific amplifications. *Molecular Ecology Resources*, 15, 747–760. <https://doi.org/10.1111/1755-0998.12349>
- R Development Core Team (2012). *R: A language and environment for statistical computing*. Vienna Austria: R Found. Stat. Comput.
- Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences*, 102, 15942–15947. <https://doi.org/10.1073/pnas.0507611102>
- Robertson, B. C., & Gemmell, N. J. (2005). Microsatellite DNA markers for the study of population structure in the New Zealand fur seal *Arctocephalus forsteri*. *DOC Science Internal Series*, 196, 4–18.
- Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of genetic markers for inference of ancestry*. *American Journal of Human Genetics*, 73, 1402–1422. <https://doi.org/10.1086/380416>
- Ryyänänen, H. J., Tonteri, A., Vasemägi, A., & Primmer, C. R. (2007). A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *Journal of Heredity*, 98, 692–704. <https://doi.org/10.1093/jhered/esm093>
- Salis, A., Easton, L., Robertson, B., Gemmell, N., Smith, I., Weisler, M., ... Rawlence, N. J. (2016). Myth or relic: Does ancient DNA detect the enigmatic Upland Seal? *Molecular Phylogenetics and Evolution*, 97, 101–106. Elsevier Inc. <https://doi.org/10.1016/j.ympev.2015.12.012>
- Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, 22, 25–33 Elsevier. <https://doi.org/10.1016/j.tree.2006.08.009>
- Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., & Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources*, 11, 1–8. <https://doi.org/10.1111/j.1755-0998.2010.02979.x>
- Skovrind, M., Olsen, M. T., Vieira, F. G., Pacheco, G., Carl, H., Gilbert, M. T. P., & Møller, P. R. (2016). Genomic population structure of freshwater-resident and anadromous ide (*Leuciscus idus*) in north-western Europe. *Ecology and evolution*, 6, 1064–1074. <https://doi.org/10.1002/ece3.1909>
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457–462.
- Smith, I. W. G. (1989). Maori Impact on the marine megafauna: pre-European distributions of New Zealand sea mammals. In D. G. Sutton (Ed.), *Saying so doesn't make it so: Papers in Honour of B. Foss Leach*, 76–108. Dunedin, New Zealand Archaeological Association Monograph No. 17.
- Smith, I. (2005). Retreat and resilience: Fur seals and human settlement in New Zealand. In M. G. Gregory (Ed.), *Exploitation and cultural importance of sea mammals* (pp. 6–18). Durham, UK: Oxbow Books.
- Sorensen, J. (1969). *New Zealand fur seals, with special reference to the 1946 open season: NZ Fisheries Division*. pp. 80.
- Stockin, K. A., Amaral, A. R., Latimer, J., Lambert, D. M., & Natoli, A. (2014). Population genetic structure and taxonomy of the common dolphin (*Delphinus* sp.) at its southernmost range limit: New Zealand waters. *Marine Mammal Science*, 30, 44–63. <https://doi.org/10.1111/mms.12027>
- Templeton, A. R. (1980). The theory of speciation via the founder principle. *Genetics*, 94, 1011–1038.
- Taylor, R. H. (1982). New Zealand fur seals at the Bounty Islands. *New Zealand Journal of Marine and Freshwater Research*, 16, 1–9. <https://doi.org/10.1080/00288330.1982.9515942>
- Thompson, F. N., & Abraham, E. R. (2010). Estimation of fur seal (*Arctocephalus forsteri*) bycatch in New Zealand trawl fisheries, 2002 – 03 to 2008 – 09. In New Zealand Aquatic Environment and Biodiversity Report No. 61 (P. 37). Wellington, New Zealand.
- Tracy, L. N., Wallis, G. P., Efford, M. G., & Jamieson, I. G. (2011). Preserving genetic diversity in threatened species reintroductions: How many individuals should be released? *Animal Conservation*, 14, 439–446. <https://doi.org/10.1111/j.1469-1795.2011.00448.x>
- Weston, K. A., & Robertson, B. C. (2015). Population structure within an alpine archipelago: Strong signature of past climate change in the New Zealand rock wren (*Xenicus gilviventris*). *Molecular Ecology*, 24, 4778–4794. <https://doi.org/10.1111/mec.13349>
- Wilson, G. J. (1981). *Distribution and abundance of the New Zealand Fur Seal, Arctocephalus Forster*. Wellington, New Zealand: Ministry of Agriculture and Fisheries.
- Wynen, L., Goldsworthy, S. D., Guinet, C., Bester, M. N., Boyd, I. L., Gjertz, I., ... Slade, R. (2000). Postsealing genetic variation and population structure of two species of fur seal (*Arctocephalus gazella* and *A. tropicalis*). *Molecular Ecology*, 9, 299–314. <https://doi.org/10.1046/j.1365-294x.2000.00856.x>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Dussex N, Taylor HR, Stovall WR, et al. Reduced representation sequencing detects only subtle regional structure in a heavily exploited and rapidly recolonizing marine mammal species. *Ecol Evol*. 2018; 8:8736–8749. <https://doi.org/10.1002/ece3.4411>