1

2 DR. AMY M VAN CISE (Orcid ID : 0000-0002-0613-4072)

3

4

6

7

8 # Familial social structure and socially-driven genetic differentiation in

9 # Hawaiian short-finned pilot whales

10

11 Amy M. Van Cise[1*], Karen. K. Martien[2], Sabre D. Mahaffy[3], Robin W. Baird[3], Daniel L.

12 Webster[3], James H. Fowler[4], Erin M. Oleson[5], Phillip A. Morin[2,1]

13

14 1. Scripps Institution of Oceanography, 8622 Kennel Way, La Jolla, CA 92037
15 2. Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 8901 La Jolla Shores Dr.,
16   La Jolla, CA, 92037, USA
17 3. Cascadia Research Collective, 218 1/2 4th Ave W, Olympia, WA 98501
18 4. University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92037
19 5. Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA, 1845 Wasp
20   Boulevard, Building 176, Honolulu, HI 96818

21

24 [*]Corresponding author: Amy Van Cise, avancise@gmail.com

25 8901 La Jolla Shores Drive

26 La Jolla, CA 92037

Abstract

29

30   Social structure can have a significant impact on divergence and evolution within species, especially in the marine environment, which has few environmental boundaries to dispersal. On the other hand, genetic structure can affect social structure in many species, through an individual preference toward associating with relatives. One social species, the short-finned pilot whale (*Globicephala macrorhynchus)*, has been shown to live in stable social groups for periods of at least a decade. Using mitochondrial control sequences from 242 individuals and SNPs from 106 individuals, we examine population structure among geographic and social groups of short-finned pilot whales in the Hawaiian Islands, and test for links between social and genetic structure. Our results show that there are at least two geographic populations in the Hawaiian Islands: a Main Hawaiian Islands (MHI) population and a Northwestern Hawaiian Islands/Pelagic population (F_{ST} and Φ_{ST} $P < 0.001$), as well as an eastern MHI community and a western MHI community (F_{ST} $P = 0.009$). We find genetically-driven social structure, or high relatedness among social units and clusters ($P < 0.001$), and a positive relationship between relatedness and association between individuals ($P < 0.0001$). Further, socially-organized clusters are genetically distinct, indicating that social structure drives genetic divergence within the population, likely through restricted mate selection (F_{ST} $P = 0.05$). This genetic divergence among social groups can make the species less resilient to anthropogenic or ecological disturbance. Conservation of this species therefore depends on understanding links among social structure, genetic structure, and ecological variability within the species.

49

Introduction

51   While the concept of culture has traditionally been reserved for human societies, more recently biologists have identified and described aspects of culture in non-human species, such as elephants, birds, primates, pinnipeds, and cetaceans (e.g., Mundinger, 1980; Lachlan & Slater, 1999; Rendell & Whitehead, 2001, 2003; McComb & Semple, 2005; Laland & Janik, 2006; Wittemyer et al., 2009; de la Torre & Snowdon, 2009; Kershenbaum et al., 2012; Riesch et al., 2012; Kessler et al., 2014). Theoretical studies have long suggested the existence of gene-culture

57 coevolution outside humans, and integrative studies of genomic and cultural traits are beginning

58 to provide evidence of gene-culture coevolution in social mammals, both in a narrow sense (i.e.,

59 direct links between genes and cultural phenotypic traits), and a broad sense (i.e., population-

60 level genetic differences among groups with different cultures or societies). Sociality has been

61 shown to increase inclusive fitness in cooperative species (e.g., Connor et al., 1992; Rendell and

62 Whitehead, 2001), and therefore be an evolutionarily advantageous trait. Socially-driven, fine-

63 scale genetic structure has been documented in primates and some other social mammals, such as

64 elephants, rock wallabies (*Petrogale penicillata*), prairie dogs (*Cynomys ludovicianus*), killer

65 whales (*Orcinus orca*) and sperm whales (*Physeter macrocephalus*) (e.g., Pope, 1992; Dobson et

66 al., 1998; Hazlitt et al., 2006; Wittemyer et al., 2009; Cantor et al., 2015; Foote et al., 2016).

67 These species all form socially-defined groups that are genetically distinct due to nonrandom

68 mating and dispersal patterns, and are often characterized by matrilineal societies with male-

69 biased dispersal. These types of societies, if stable over many generations, could lead to the co-

70 evolution of genes and culture.

71 Because cetaceans live in an environment with few boundaries to dispersal, social

72 structure may play an important role in driving population structure and evolution. Stable social

73 structures (i.e., hierarchical group associations that remain stable for decades to generations)

74 have been identified in four species of cetacean - sperm whales, killer whales, long-finned pilot

75 whales (*Globicephala melas*) and short-finned pilot whales (*G. macrorhynchus*) (e.g., Amos et

76 al., 1993; Baird & Whitehead 2000; Cantor et al., 2015; Mahaffy et al., 2015). Whitehead (1998)

77 suggests that the dearth of mitochondrial diversity in these four highly social cetaceans may be

78 driven by selection for maternally-inherited cultural traits. In killer whales and sperm whales, the

79 effects of social structure and cultural learning (e.g., foraging techniques, migration patterns,

80 predator avoidance, and vocal traditions) as drivers of genetic structure have been well

81 documented (e.g., Ford & Fisher 1982; Janik & Slater 1997; Weilgart & Whitehead 1997; Foote

82 et al., 2009, 2016; Filatova et al., 2012; Rendell et al., 2012; Riesch et al., 2012; Cantor et al.,

83 2015). However, little is understood of the social and genetic structure of pilot whales, or the

84 links between the two.

85 Just as social structure can affect genetic structure, genetic structure can have a driving

86 effect on social structure, if individuals choose to associate with close relatives rather than

87 disperse throughout their range, even though it may or may not provide an evolutionary

88   advantage (Beck et al., 2011). The positive feedback loop created by these two complementary
89   processes may stabilize social units or clusters, allowing co-evolutionary genetic and social
90   divergence to occur. While many aspects of this theory have been discussed (e.g., Findlay, 1991;
91   Laland, 1992; Lachlan and Slater, 1999), empirical evidence of stable gene-culture coevolution
92   outside of humans is limited (Rendell & Whitehead, 2001). Although research in this area is
93   increasing (e.g., Foote et al., 2016), the relationship between ecology, culture, and genetics is
94   poorly understood in all species (Laland et al., 2010).

95   Short-finned pilot whales, due to their social nature, may be affected by this reciprocal
96   link between social structure and genetic structure. Stable social units (Mahaffy et al., 2015) and
97   a long period of post-reproductive senescence in females (Marsh and Kasuya, 1986) may
98   contribute to gene-culture divergence in this species, both at the population and sub-population
99   level, as is true of killer whales (Brent et al., 2015). In the Pacific Ocean, two types of short-
100 finned pilot whale have been identified, distinct in their morphology, genetics, distribution and
101 vocal repertoire (Kasuya et al., 1988; Oremus et al., 2009; Van Cise et al., 2016, 2017). Little is
102 known of the mechanism of divergence between these two types, but due to their similarity to
103 killer whales in several life history characteristics (e.g., stable social units, reproductive
104 senescence in females, and distinct vocal repertoires), we hypothesize that cultural adaption to
105 distinct ecological environments (e.g., diet preference or foraging techniques) promoted the
106 divergence of the two types (Riesch et al., 2006), which may be distinct sub-species or species.

107 The Hawaiian archipelago is home to one of these types, the Naisa-type short-finned pilot
108 whale (Van Cise *et al.,* 2016). Their density is highest around the Main Hawaiian Islands (MHI),
109 but they are also found in the Northwestern Hawaiian Islands (NWHI) and pelagic waters
110 surrounding the archipelago. Photo ID and observations suggest little overlap between these
111 three regions (Baird, 2016).

112 Longitudinal observations and photo identification (photo ID) data collected since 2000
113 have been used to calculate the rate of association between pairs of individuals (called the
114 association index, and ranging from 0-1), using a half-weight index to control for effort
115 (Whitehead, 2008; Mahaffy et al., 2015). This revealed that short-finned pilot whales in Hawaiʻi
116 form stable social units of approximately 12 individuals for periods of at least a decade, and that
117 these social units will often associate with a number of other social units in affiliations called
118 clusters, with an average of 23 individuals (Mahaffy et al., 2015). Social units, the smallest group

119 in the social hierarchy, have a mean association index of 0.76. Clusters, the next hierarchical

120 level, comprise one or more social units with mean association index of 0.48.

121   Additionally, satellite tag and photo ID data indicate that, within the MHI, three island-

122 associated communities may exist: an eastern MHI community, around Hawai'i Island, a western

123 MHI community around O'ahu and Kaua'i Islands, and central MHI community around O'ahu

124 and Lāna'i Islands (Baird, 2016). The presence of these communities suggests that, in regions

125 with highly heterogeneous habitat such as island archipelagos, habitat preference may be an

126 important driver of local structure. Individuals are philopatric to their island communities,

127 although some social units have been observed on rare occasions visiting other communities, and

128 there is some overlap in geographic range among communities (Baird, 2016). Communities

129 represent the highest level of social organization, comprised of multiple clusters (Mahaffy et al.,

130 2015, Baird, 2016), therefore habitat preference may be a socially-learned behavior.

131   Based on studies from short-finned pilot whale populations in the Atlantic Ocean, social

132 units are thought to be matrilineal (Heimlich-Boran, 1993; Alves et al., 2013). These two studies

133 suggest that males remain in their natal social unit but mate outside of that group. However, in at

134 least some cases, all-male groups have been observed (Baird, 2016), suggesting that males do not

135 always exhibit natal philopatry. It is unknown whether males' extra-unit mate choices are

136 random or socially-driven, or whether genetic relatedness affects association or social structure

137 at any level higher than that of social units.

138   In this study, we aim to improve our understanding of local population structure and

139 divergence in Hawaiian short-finned pilot whales. We analyze genetic differentiation between

140 three geographic strata: the MHI, Northwestern Hawaiian Islands and pelagic waters surrounding

141 the Hawaiian Islands; we then examine genetic differentiation between observed island

142 communities within the MHI, test for sex-biased dispersal between those communities, and look

143 for evidence that individual island preference is a driver of the amount of time that individuals

144 spend together.

145   We further hypothesize that relatedness drives social structure, and that, in turn, social

146 structure affects genetic divergence among groups, for example by affecting mate selection. If

147 genetic structure affects social structure, insomuch as close relatives form lifelong associations

148 and travel in close-knit groups, we would expect to see higher relatedness within social units

149 than expected at random. Similarly, if social structure affects genetic structure we might expect

150   to see genetic divergence in the allele frequency among clusters. These patterns, along with a
151   statistical relationship between genetic and social structure, could indicate a reciprocal
152   relationship between genetic and social structure in Hawaiian pilot whales.

153

154   **Methods**

155   *Genetic data collection*

156   Skin samples (n=254) were collected from wild short-finned pilot whales throughout the
157   MHI and NWHI using biopsy darts, in collaboration with Cascadia Research Collective (CRC)
158   and NOAA's Southwest Fisheries Science Center (SWFSC). Biopsy darts are deployed using a
159   crossbow, and collect a tissue sample approximately 8 mm in diameter and up to 20 mm in
160   length, from the area below the dorsal fin. Samples were collected opportunistically, as social
161   groups were encountered in the field, with priority given to sampling as many adults in each
162   social group as possible. Samples were archived in the SWFSC Marine Mammal and Sea Turtle
163   Research Collection, and were either stored at -80°C, or preserved in either a salt-saturated 20%
164   DMSO solution or 100% ethanol and stored in a -20°C freezer. In the MHI, known social units
165   were heavily sampled in order to test for relatedness; additional samples were chosen randomly,
166   with consideration given to ensuring that samples represented unrelated individuals from
167   multiple social groups per stratum.

168

169   *Photo ID/social network data collection*

170   Photographs, used to generate social stratification data as well as pairwise association
171   indices between individuals, were collected according to Mahaffy et al. (2015). Photo
172   identification data from that publication and from subsequent field observations, between 2003
173   and 2015 (Baird et al., 2013), are included in this study. Association indices were calculated
174   using SOCPROG 2.4, with a sampling period of one day and a half-weight index (HWI) of
175   association to control for effort (Whitehead, 2008, 2009). We used the photo identification,
176   association indices, and terms (social units, clusters, and communities) used by Mahaffy et al.
177   (2015) to characterize the hierarchical nature of short-finned pilot whale social organization in
178   the MHI.

179

180   *Genetic sequencing and assembly*

181   DNA was extracted from skin and muscle samples as previously described (Martien et
182 al., 2014). The hypervariable mtDNA control region was amplified and sequenced in two parts
183 of approximately 420 bp and 560 bp, with approximately 20 bp of overlap between the two
184 sequences. Primers, PCR, and sequencing methods have been previously described by Martien et
185 al., (2014). The resulting combined sequence was 962 bp, and was assembled using SEQED,
186 version 1.0.3 (ABI), Sequencher software (versions 4.1 and 4.8; Gene Codes, Ann Arbor, MI,
187 USA) or Geneious (Kearse et al., 2012).

188   Mitochondrial sequences were aligned using a MAFFT alignment with default
189 parameters (Scoring Matrix: 200PAM/k=2, Gap open penalty: 1.53, Offset value: 0.123) in the
190 Geneious software package (Katoh & Kuma, 2002). Once the alignment was completed,
191 sequences were re-examined. Any haplotypes represented by only a single sequence or
192 haplotypes with a single base-pair difference from the most similar haplotype were reviewed for
193 accuracy. Unique haplotypes were repeat sequenced in order to ensure the accuracy of the
194 sequence. Sequences were combined with previously published sequences from Van Cise et al.
195 (2016) to generate the final mtDNA data set.

196   Sequencing of 78 targeted nuclear loci for SNP analysis was completed using a custom
197 capture enrichment array designed at SWFSC based on common bottlenose dolphin (*Tursiops*
198 *truncatus*) genome sequences (Supplemental File S1), followed by highly-parallel sequencing
199 (Hancock-Hanser et al., 2013; Morin et al., 2015). Four libraries of genomic DNA were prepared
200 using protocols described in Meyer and Kircher (2010) and Hodges et al. (2009), with
201 modifications described in Hancock-Hanser et al. (2013). Up to 400 ng of extracted DNA in 80
202 μL total volume was sonicated using a Bioruptor UCD-200 (Diagenode). Blunt-ends of the DNA
203 were repaired using 20 μL of the sonicated product, adaptors were ligated to the DNA, and
204 indexes were added to each sample library via PCR with indexed primers (Meyer & Kircher,
205 2010). Once indexed, each sample was quantified using qPCR to estimate the number of nuclear
206 DNA copies in each sample, and approximately 100,000 copies per sample were pooled and
207 hybridized to a capture array. The capture-enriched product was amplified, then sequenced on
208 Illumina HiSeq (1 x 100 bp) or NextSeq (1 x 75 bp) instruments by The DNA Array Core
209 Facility (The Scripps Research Institute, La Jolla, CA).

210   Nuclear sequences were assembled as in Morin et al. (2015), using common bottlenose
211 dolphin reference sequences (used for capture enrichment) for sequence assembly and SNP

genotyping. The cutoff for calling a genotype at any position was set to 10 reads for both homozygous and heterozygous positions, to minimize genotype error (Fountain et al., 2016). Potential SNPs were identified using scripts developed at SWFSC (Dryad data repository doi:10.5061/dryad.cv35b) in the R computing environment (R Core Team 2016). From the pool of sequenced loci, candidate SNPs were selected if at least five individuals were heterozygous at that locus. Those SNPs with coverage at fewer than 55% of samples were removed, and samples with coverage at fewer than 70% of the SNP loci were also removed. Next, sequenced regions with multiple SNP loci were examined for signs of paralogous reads within the assembly (e.g., excess heterozygosity across multiple SNPs in a region, discrete regions of high coverage), and SNPs were removed if assembly of paralogous loci was determined to have occurred. Finally, quality control analyses were performed on this set of SNPs and samples using the strataG package for R (Archer et al., 2017). SNPs were removed if the quality control analysis indicated that the locus was an outlier for homozygosity (>80% homozygous, based on the distribution of homozygous genotypes across all loci), and we additionally tested for outliers from HWE, using a Bonferonni adjustment for multiple tests. Loci that deviated significantly from HWE equilibrium were closely re-examined for evidence of assembly of paralogous loci. Additionally, samples that had highly similar SNP genotypes and could be duplicates were checked against photo ID records to confirm that they were distinct individuals; if this could not be determined, one from each pair of duplicate samples was removed. Loci with multiple SNPs (see Supplemental Table S1) were phased based on allele frequencies in the three regional strata, with a phase cutoff probability of 0.5, to generate a single multi-SNP genotype per sample at each locus for analyses of genetic differentiation (Morin et al., 2012). For analysis of relatedness within Hawaiian social units, the highest heterozygosity SNP at each locus (N = 51 after removal of one locus that was invariant in these populations) was chosen for the analysis.

Data analysis: Population structure and diversity

We tested for both geographic and socially-driven genetic structure using both mitochondrial control regions and nuclear SNPs. Supplemental Table S2 lists sample stratifications used for data analysis in this study. For mitochondrial DNA analysis, samples were divided into three strata: Main Hawaiian Islands (MHI), Northwestern Hawaiian Islands (NWHI), and pelagic samples (Fig 1). Samples were placed in one of these three strata primarily

243 based on their sampling location, with the exception that samples collected near the MHI were
244 placed in the pelagic stratum if photo ID data verified that the individuals did not associate with
245 MHI communities. MHI mtDNA samples were not further stratified because all samples except
246 one have the same haplotype. We placed samples from the NWHI in a separate stratum because
247 several studies have shown strong differentiation between the MHI and NWHI for other marine
248 mammals (Andrews et al., 2010; Courbis et al., 2014; Martien et al., 2014).

249 SNP data were only available for the MHI and pelagic strata. Using previous knowledge
250 of the social structure, habitat use, and movements (Baird 2016; Mahaffy et al., 2015), SNP
251 samples were divided into two strata within the MHI (eastern and western MHI communities)
252 based on photo ID data, visual observations of social units, and satellite tag data (Figure 1).
253 Several social units were heavily sampled in order to test for relatedness within social units.
254 Therefore, in order to remove any potential bias due to sampling regime, we randomly
255 subsampled the dataset using a random number generator to include no more than two
256 individuals from each social unit before conducting tests of genetic differentiation among
257 geographic strata.

258 Molecular diversity indices for all samples and for each region were calculated for both
259 mtDNA (Theta (θ_H), haplotypic diversity (h), and mean nucleotide diversity (π)) and SNP
260 genotypes (average number of alleles per locus, expected and observed heterozygosity (H_e, H_o)).
261 Pairwise genetic differentiation was calculated among geographic strata using F_{ST} and Φ_{ST} for
262 mtDNA. For SNP genotypes, geographic differentiation (F_{ST} only) was calculated only between
263 island communities within the MHI. All estimates of divergence and genetic diversity were
264 conducted using the strataG package for R except haplotypic diversity, which was calculated in
265 Arlequin (Excoffier & Lischer 2010).

266 We tested for sex-biased dispersal among island communities using the Hierfstat package
267 in R (Goudet 2005), which looks for first-generation immigrants within the sample set. To do
268 this, we tested for differences among males and females in F_{ST}, F_{IS}, or the mean or variance of
269 assignment probability (Goudet et al., 2002).

Data analysis: genetic structure, social structure, and island preference

271 In order to test the hypothesis that there are links between genetic structure, social
272 structure, and island preference in Hawaiian short-finned pilot whales, we first calculated

273 pairwise genetic relatedness among individuals, as well as pairwise genetic differentiation among
274 clusters, which represent one or more social units.

275 To calculate genetic relatedness within and among social units in the MHI, samples were
276 stratified according to previously inferred social structure (Mahaffy *et al.,* 2015), and social unit
277 relatedness was calculated if at least five individuals from a social unit had been sampled.
278 Pairwise relatedness was estimated using a dyadic maximum likelihood estimator (Milligan,
279 2003) in the R package Related (Pew et al., 2014), which implements the software program
280 COANCESTRY (Wang & Summers, 2010). Within-unit relatedness was compared to the
281 expected relatedness by permuting a random sample 1,000 times and calculating relatedness.
282 From one cluster, we were able to sample two social units, and we used this cluster to test the
283 hypothesis that genetic relatedness is a driver of association among social units by comparing
284 within-cluster relatedness with the distribution of relatedness between 1,000 randomly selected
285 pairs of social units.

286 Pairwise genetic differentiation (F_{ST}) was estimated among clusters using SNP genotypes
287 only due to the lack of mtDNA haplotypic diversity. Clusters were only included if there were at
288 least five samples collected from that cluster. To characterize the overall degree of differentiation
289 among social clusters, we performed this test using all available samples from clusters. Next, to
290 characterize the extent to which gene differentiation has been affected by social structure, we
291 removed highly related (r > 0.6) samples to reduce bias due to genetic relatedness and
292 recalculated F_{ST} among social clusters, now considering differences in the allele frequency
293 within each cluster.

294 To determine whether genetically similar social units and clusters were more likely to
295 associate, we compared pairwise cluster genetic differentiation (F_{ST}) with mean pairwise
296 association between clusters, using a fixed effect linear model with cluster ID controlled as a
297 fixed effect. Association between pairs of clusters was calculated by taking the mean of
298 association between individuals in the first cluster and individuals in the second cluster.

299 We used Mantel tests and linear models to examine the relationship between geographic
300 distance, genetic relatedness, and associations between individuals. To do this, we first
301 calculated geographic distance (d) as the straight-line distance between sampling locations for
302 each sample. Three Mantel tests were calculated between all pairs of individuals, comparing

genetic distance (defined as 1 – genetic relatedness, r), geographic distance (d), and the amount of time a pair spends together (association index, AI).

We compared linear, exponential, and logarithmic models to test the importance of geographic distance (d), genetic relatedness (r), and an interaction term (r*d) as potential drivers of association (AI) between individuals, and also between clusters. For these models, we converted geographic distance to a categorical variable with two categories (inter-island, d < 300 mi and intra-island, d > 50 mi), due to the fact that, within each island community, sampling location is not representative of an individual's habitat use or distance to other individuals in the community. Further, in order to account for multiple observations of each individual, we included fixed effects for each pairwise individual (I). We iteratively built models by adding one predictor variable with each iteration, for a final model that included all possible predictor terms:

$$E(f[AI_{ij}]) = \alpha + \beta_1 r_{ij} + \beta_2 d_{ij} + \beta_3 r_{ij} d_{ij}, + G(I_i) + G(I_j)$$

Significant parameters of the model that minimized Akaike's Information Criterion (AIC) were considered to be potential drivers of association among pairs of individuals.

Results

The mtDNA dataset consisted of 242 samples from throughout the Hawaiian Islands (125 previously reported in Van Cise et al., 2016). A total of 163 SNPs at 50 nuclear loci from 112 individuals were successfully genotyped from four capture-enriched library pools. The SNP and mtDNA datasets overlapped by 100 samples. Six samples were determined to be duplicates and removed from the dataset, so that the final SNP dataset included 106 individuals (Dryad Digital Repository http://dx.doi.org/10.5061/dryad.xxxxx). Forty-four SNPs were removed during the quality analysis phase due to possible assembly of paralogous loci, resulting in 119 SNPs at 49 nuclear loci (Supplemental Table S1).

Sample stratifications can be found in Figure 1 and Supplemental Table S2. Only eight samples with SNP data were available from the pelagic stratum, and no samples were successfully genotyped from the Northwestern Hawaiian Islands. Cluster assignments were made for 93 of the samples; analyses of differentiation among social clusters were performed using a dataset that included related individuals (n=93) and a dataset with individuals removed from pairs with relatedness estimates > 0.6 (n=85). Finally, pairwise relatedness based on the 51

unlinked SNPs was calculated for the full 106 sample SNP dataset, and group relatedness was calculated for three social units, five clusters and two communities.

Population structure and diversity

We found very low mtDNA haplotype diversity in the Hawaiian Islands (Table 1). Six haplotypes were identified among the 242 samples (Table 2); 232 of the 242 samples had haplotype J. With the exception of one sample collected off Kaua'i, all samples from the MHI stratum had haplotype J. SNP genotypes were subsampled within each island community to control for non-random sampling of social groups, so that the dataset used for molecular diversity and geographic differentiation included 63 samples from the MHI. Observed and expected heterozygosity for the phased multi-SNP genotypes in the MHI were both 0.46, with slightly higher heterozygosity in the western MHI community than in the eastern MHI community (Table 1).

Mitochondrial differentiation was significant between the MHI ($N = 204$) and NWHI ($N = 17$) strata, as well as between the MHI and pelagic ($N = 20$) strata (F_{ST} and Φ_{ST} $P < 0.001$, Table 3). Within the MHI, SNP differentiation was small but significant between the eastern ($N = 42$) and western ($N = 21$) MHI communities (F_{ST} $P = 0.009$). SNP differentiation was not tested between other strata (pelagic, NWHI) due to low sample size. We did not find any evidence of sex-biased dispersal between communities in the MHI (*P*-values for all indices ranged from 0.2 to 0.9).

Genetic structure, social structure, and island preference

Average pairwise relatedness (r) among individuals was 0.11, with a range from 0 to 0.76. Within-unit relatedness estimates for each of three social units with five or more samples were all significantly higher than expected if groups were randomly organized (Figure 2). Within-cluster relatedness for cluster H20, comprised of three social units, was also significantly higher than relatedness between randomly selected pairs of social units (r = 0.33, $P < 0.03$), as well as being higher than mean relatedness at the community level (r = 0.11).

When pairs with r > 0.6 were removed, clusters with more than five individuals sampled were found to be significantly differentiated from each other in eight out of ten pairwise comparisons (Table 4). Global F_{ST} was also significant when tested using all samples with cluster assignments (n = 84, $F_{ST} = 0.02$, $P = 0.05$). When the same analysis is performed using all samples regardless of relatedness, the number of significant pairwise differences between

364 social clusters increases from eight to nine, likely due an increase in both sample size and
365 relatedness within groups (Supplemental Table S3).

366   Pairs of clusters that exhibited higher genetic differentiation associated less often (Figure
367 3), according to the results of a fixed effect linear regression, which indicated a negative causal
368 relationship between pairwise F_{ST} differentiation and association between clusters ($P = 0.01$). In
369 this model, genetic differentiation explained 68% of the variance in association between clusters
370 ($R^2 = 0.68$).

371   While there was no correlation between relatedness and geographic distance (Mantel test
372 $P = 0.13$), association index was significantly correlated with both relatedness and distance
373 (Mantel test $P < 0.001$ for both tests).

374   Regression model fits indicated that association between individuals increases with
375 genetic relatedness. Genetic relatedness was found to be a significant driver of association time
376 ($P < 0.0001$), while distance category (near or far), and the product of genetic relatedness and
377 distance category, were not found to be significant ($P = 0.9$ and 0.2, respectively). AIC was
378 minimized using a model in which association index increased with an exponential increase in
379 relatedness (AIC = -4169), but a linear relationship was similar (AIC = -4164). Relatedness
380 explained 21% of the variance in association time between pairs of individuals ($R^2 = 0.21$).

381

382 **Discussion**

383 *Genetics, sociality and island preference*

384   Our results show that short-finned pilot whales in Hawai'i exhibit links between their
385 genetic structure, social structure and island preference, which is likely a socially-learned
386 behavior. Similar links have been shown in other social animals, such as killer whales, sperm
387 whales, and elephants (Yurk et al., 2002; Archie et al., 2006; Wittemyer et al., 2009; Rendell et
388 al., 2012; Foote et al., 2016), and may have a stabilizing effect that promotes rapid genetic
389 divergence among groups. In Hawaiian pilot whales, it seems that island preference and social
390 structure influence genetic structure in the absence of any physical barriers to gene flow, based
391 on genetic differentiation of island communities and clusters. Genetic relatedness in turn affects
392 social organization, based on high genetic relatedness within social units and clusters.

393   The importance of genetic relatedness to social organization is evident when we examine
394 the high level of relatedness within social units as compared to random (Fig. 2), a pattern that has

395 been demonstrated in pilot whales from other regions of the world (Alves et al., 2013), and may

396 result from matrilineal fidelity. We additionally found that relatedness was higher within clusters

397 than throughout the Hawaiian population, suggesting that relatedness plays a role in determining

398 how groups are organized at hierarchical levels above the immediate family unit. We saw the

399 same pattern in the regression comparing relatedness with association in pairs of individuals,

400 which showed that animals that were more closely related were also more likely to associate.

401   If relatedness does not affect social structure at any level higher than that of the social

402 unit, we would expect relatedness at the cluster level to fall to the level of relatedness within the

403 entire population. Our results indicate that relatedness continues to drive social structure and

404 association at higher levels in the hierarchical organization than just the matrilineal social unit.

405 This may indicate that clusters are groups of related social units that underwent fission, similar to

406 elephants (Archie et al., 2006) and killer whales (Williams & Lusseau, 2006). Genetic

407 relatedness between groups can decay quickly in time due to the death of kin, and would be

408 consistent with the lower relatedness within clusters than social units that we observed in this

409 study.

410   In elephants, social units that associate more often were shown to have recently split from

411 each other due to the death of a matriarch (Archie et al., 2006). A larger, more comprehensive

412 sample that includes all or most clusters, and a greater number of SNPs, would increase the

413 resolution of the genetic structure among socially-divided units, clusters and communities, and

414 may allow us to determine which clusters are more genetically similar, and whether specific

415 clusters are facilitating gene flow between island communities.

416   On the other hand, we were able to show significant genetic differentiation among

417 sympatric clusters even when highly related individuals were removed from our analyses,

418 indicating restricted gene flow among sympatric clusters. Clusters that were more genetically

419 differentiated also spent less time together (Fig. 3). This would suggest that social structure

420 inhibits gene flow among clusters, which could accelerate genetic divergence among clusters

421 compared to a group of randomly mating individuals. It is important to note, however, that the

422 observed genetic differentiation among clusters may also be caused by low effective population

423 size, sampling stochasticity, or a combination of these factors.

424   This bi-directional influence between social structure and genetic structure creates a

425 positive feedback between the two that may be self-stabilizing, thus encouraging continued

426 genetic and social divergence. Similar patterns have been seen in other social animals; for
427 example, in some bird species, social song learning has been argued to restrain genetic
428 divergence soon after a dispersal event, but promote divergence at later stages in the process
429 (Slabbekoorn & Smith, 2002). In killer whales, social structure and social learning are thought to
430 have promoted rapid sub-species divergence into novel ecological niches (Foote et al., 2016). In
431 a similar way, social structure in pilot whales may promote genetic divergence, and in turn
432 genetic relatedness helps maintain a familial social structure.

433 Geographic distance is significantly correlated with association between individuals, or
434 social structure, although it was not found to be a significant driver of association between
435 individuals. Since geographic distance (d) cannot be interpreted as a continuous variable, due to
436 the geographic overlap of social units within island communities, it instead represents
437 individuals that were encountered in the same island community (d < 50 mi) or different island
438 communities (d > 300 mi). The correlation between geographic distance and association among
439 individuals likely indicates that individual preference for one island community and association
440 with other individuals are both driven by similar mechanisms.

441 While the present study did not examine genetic or social structure as drivers of
442 ecological behaviors such as island preference, there is evidence for social and parental (i.e.,
443 genetic) learning of ecological and other behaviors in other highly social cetaceans, such as killer
444 whales and sperm whales (Cantor et al., 2015; Foote et al., 2016). Indeed, social learning of
445 ecological behaviors may be important to the long-term resilience of oceanic predators
446 (Whitehead, 2007b). Further studies of ecological and social behaviors in pilot whales, such as
447 diet preference, foraging strategies, mating strategies, group movements, and vocal repertoire
448 would help elucidate whether social and genetic structure also contribute to the learning and
449 practice of these behaviors.

450

451 *Population structure and diversity*

452 Mitochondrial diversity is very low in Hawaiian short-finned pilot whales: of the six
453 haplotypes reported in this study, haplotype J made up the majority of individuals, and although
454 sampling was increased in the MHI from previous Pacific-wide studies (Van Cise et al., 2016),
455 no new haplotypes were found in this study. The MHI stratum was distinct from the pelagic and
456 NWHI strata, indicating the presence of an insular population around the MHI, as well as a

457   pelagic/NWHI population. Insular or coastal populations have been observed in other
458   odontocetes, such as false killer whales (Martien et al., 2014), bottlenose dolphins (Allen et al.,
459   2016) and spinner dolphins (Andrews et al., 2010). Pilot whales exhibit strong site fidelity
460   (Mahaffy et al., 2015), and it is possible that the MHI population has become adapted to the
461   slope habitat it prefers (Baird, 2016; Abecassis et al., 2015), and may have different dietary
462   preferences from the pelagic population. However, tagging data indicate that pelagic social
463   groups will sometimes travel through the slope region of the MHI (Baird, 2016). The lack of
464   mtDNA gene flow between these two populations suggests that social structure prevents
465   dispersal of females between these two populations when they come in contact with each other.

466   Although mtDNA differentiation between the pelagic and NWHI strata was non-
467   significant, we expect that a larger sample size will differentiate the two populations. Samples
468   from the pelagic stratum had haplotypes also found in SE Asia, the South Pacific, the Indian
469   Ocean, and southern Japan, while NWHI haplotypes were either J (MHI) or an endemic
470   haplotype with 4 bp difference from J, suggesting that the NWHI group may have diverged from
471   the MHI insular population, possibly due to geographic isolation. This is similar to the pattern
472   observed in Hawaiian false killer whales (*Pseudorca crassidens*), where photo-identification,
473   tagging, and mtDNA suggest three populations, with shared maternal ancestry between the MHI
474   and NWHI, but nuclear data showing contemporary gene flow is highest between the NWHI and
475   pelagic populations (Martien et al., 2014). However, our nuclear SNP sample size was not large
476   enough to test for geographic differentiation between these strata, therefore the possibility still
477   remains for male-mediated gene flow between the NWHI and Pelagic strata. A large dataset of
478   both mtDNA haplotypes and SNP genotypes from the NWHI and pelagic strata may provide
479   greater insight into the historical and contemporary rates of gene flow among these geographic
480   areas.

481   Within the insular MHI population, there are at least two genetically distinct island
482   communities, with some continued gene flow between them. This may be driven by cluster
483   philopatry to island communities, with some clusters key to gene flow between communities.
484   Satellite tag data indicate a third possible community, around Oʻahu/Lānaʻi, known as the central
485   MHI community (Baird, 2016). Additional samples from that community are needed to test
486   whether it is genetically distinct from the eastern and western MHI communities. Individuals
487   rarely leave their island community, instead spending the majority of their time around one

island; however, on rare occasions clusters have been observed outside their island community

ranges (Baird 2016), and mating may occur during these rare excursions.

Within small groups, such as social units or clusters, inbreeding depression can be avoided through mechanisms such as sex-biased dispersal (Prout, 1981). We found no detectable difference in genetic diversity indices at the regional, MHI population, or community level, indicating a lack of inbreeding, though there was no nuclear evidence for sex-biased dispersal among communities. Sugg et al., (1996) use a socially-structured population of prairie dogs to show that an increase in coancestry within a breeding group is countered by divergence among groups, which works to maintain genetic diversity at the population level. This can happen through kin recognition and behavioral avoidance of mating within a group, or if one sex remains philopatric to the group while the other sex is more likely to disperse. The advantages of social living, such as cooperative behaviors and increased genetic fitness, are thought to outweigh the costs if inbreeding can be avoided (Sugg et al., 1996). In Main Hawaiian Island pilot whales, high levels of coancestry, or relatedness, within social units and clusters may be countered by genetic divergence among these groups, thus maintaining genetic diversity at the community and population level. However, Parreira & Chikhi (2015) found that randomly permuting social unit membership within a population always produces an excess of heterozygotes, and concluded that it is not necessary to use inbreeding-avoidance mechanisms to explain outbreeding signatures in small groups, but rather that social structure itself generates outbreeding signatures that can have advantageous fitness traits.

Short-finned pilot whales in Hawaiian waters are subjected to a variety of anthropogenic impacts, including interactions with fisheries, vessel strikes, and exposure to high-intensity Navy sonars (Baird, 2016). Social species such as this can be more vulnerable to the removal of a single individual, as it may precipitate the loss of an entire group (Wade et al., 2012). If some clusters contribute more to gene flow between communities, the loss of those clusters could act to fragment communities within the MHI, which would decrease genetic diversity and increase demographic isolation in each region, thus making those communities more vulnerable to environmental or anthropogenic perturbations.

In order to avoid this vulnerability, conservation management of this species in the Hawaiian Islands could focus on maintaining gene flow between communities within the MHI populations, similar to migration corridors between fragmented terrestrial habitats. This would

require the use of photo-identification and satellite tag data to identify individuals or social groups that regularly move among communities, and movement patterns associated with these events. Once these corridors are established, fisheries interactions within them could be monitored in order to minimize fatal injuries or inhibition of movement.

References

Abecassis M, Polovina J, Baird RW, Copeland A, Drazen JC, Domokos R, …Andrews, RD (2015) Characterizing a foraging hotspot for short-finned pilot whales and Blainville's beaked whales located off the west side of Hawai'i island by using tagging and oceanographic data. *PLoS ONE*, **10**, e0142628.

Allen SJ, Bryant KA, Kraus RHS, Loneragan NR, Kopps AM, Brown AM, … Krützen, M (2016) Genetic isolation between coastal and fishery-impacted, offshore bottlenose dolphin (*Tursiops spp.*) populations. *Molecular Ecology*, **25**, 2735–2753.

Alves F, Quérouil S, Dinis A, Nicolau C, Ribeiro C, Freitas L, …Fortuna, C (2013) Population structure of short-finned pilot whales in the oceanic archipelago of Madeira based on photo-

548   identification and genetic analyses: implications for conservation. *Aquatic Conservation:*

549   *Marine and Freshwater Ecosystems*, **5**, 758–776.

550   Amos W, Schlotterer C, Tautz D (1993) Social structure of pilot whales revealed by analytical

551   DNA profiling. *Science*, **260**, 670–672.

552   Andrews KR, Karczmarski L, Au WWL, Rickards SH, Vanderlip CA, Bowen BW, … Toonen

553   RJ (2010) Rolling stones and stable homes: social structure, habitat diversity and population

554   genetics of the Hawaiian spinner dolphin *(Stenella longirostris)*. *Molecular Ecology*, **19**,

555   732–48.

556   Archer FI, Adams PE, Schneiders BB (2016) strataG: An R package for manipulating,

557   summarizing, and analyzing population genetic data. *Molecular Ecology Resources*, **17**, 5-

558   11.

559   Archie EA, Moss CJ, Alberts SC (2006) The ties that bind: genetic relatedness predicts the

560   fission and fusion of social groups in wild African elephants. *Proceedings of the Royal*

561   *Society B: Biological Sciences*, **273**, 513–522.

562   Baird RW (2016) The Lives of Hawai'i's Dolphins and Whales: Natural History and

563   Conservation. University of Hawai'i Press.

564   Baird RW, Webster DL, Aschettino JM, Schorr GS, McSweeney DJ (2013) Odontocete

565   cetaceans around the Main Hawaiian Islands: Habitat use and relative abundance from small-

566   boat sighting surveys. *Aquatic Mammals*, **39**, 253–269.

567   Baird RW, Whitehead H (2000) Social organization of mammal-eating killer whales: group

568   stability and dispersal patterns. *Canadian Journal of Zoology*, **78**, 2096–2105.

569   Beck S, Kuningas S, Esteban R, Foote AD (2011) The influence of ecology on sociality in the

570   killer whale *(Orcinus orca)*. *Behavioral Ecology*, **23**, 246–253.

571   Brent LJN, Franks DW, Foster EA, Balcomb KC, Cant MA, Croft DP (2015) Ecological

572   knowledge, leadership, and the evolution of menopause in killer whales. *Current Biology*, **25**,

573   746–750.

574   Cantor M, Shoemaker LG, Cabral RB, Flores CO, Varga M, Whitehead H (2015) Multilevel

575   animal societies can emerge from cultural transmission. *Nature Communications*, **6**, 1–10.

576   Connor RC, Smolker RA, Richards AF (1992) Two levels of alliance formation among male

577   bottlenose dolphins *(Tursiops sp.)*. *Proceedings of the National Academy of Sciences of the*

578   *United States of America*, **89**, 987–990.

579 Courbis S, Baird RW, Cipriano F, Duffield D (2014) Multiple populations of pantropical spotted
580   dolphins in Hawaiian waters. *Journal of Heredity*, **105,** 627–641.

581 Dobson FS, Chesser RK, Hoogland JL, Sugg DW, Foltz DW (1998) Breeding groups and gene
582   dynamics in a socially structured population of prairie dogs. *Journal of Mammalogy*, **79**,
583   671–680.

584 Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: A new series of programs to perform
585   population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**,
586   564--567.

587 Filatova OA, Deecke VB, Ford JKB, Matkin CO, Barrett-Lennard LG, Guzeev MA, Hoyt, E
588   (2012) Call diversity in the North Pacific killer whale populations: implications for dialect
589   evolution and population history. *Animal Behaviour*, **83**, 595–603.

590 Findlay CS (1991) Fundamental theorem of natural selection under gene-culture transmission.
591   *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 4874–
592   4876.

593 Foote AD, Newton J, Piertney SB, Willerslev E, Gilbert MTP (2009) Ecological, morphological
594   and genetic divergence of sympatric North Atlantic killer whale populations. *Molecular
595   Ecology*, **18**, 5207–5217.

596 Foote AD, Vijay N, Avila-Arcos M, Baird RW, Durban J, Morin PA, …Wolf J (2016) Genome-
597   culture coevolution promotes rapid divergence in the killer whale. *Nature Communications*,
598   **7**, 1–12.

599 Ford J, Fisher HD (1982) Killer whale (*Orcinus orca*) dialects as an indicator of stocks in British
600   Columbia. *Report of the International Whaling Commission*, **32**, 671–679.

601 Fountain ED, Pauli JN, Reid BN, Palsbøll PJ, Peery MZ (2016) Finding the right coverage: The
602   impact of coverage and sequence quality on single nucleotide polymorphism genotyping
603   error rates. *Molecular Ecology Resources*, **16**, 966–978.

604 Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics.
605   *Molecular Ecology Notes*, **5**, 184–186.

606 Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited
607   genetic markers. *Molecular Ecology*, **11**, 1103–1114.

608 Hancock-Hanser BL, Frey A, Leslie MS, Dutton PH, Archer FI, Morin PA (2013) Targeted

609   multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear

610   DNA sequencing for population genomics. *Molecular Ecology Resources*, **13**, 254–68.

611 Hazlitt SL, Sigg DP, Eldridge MDB, Goldizen AW (2006) Restricted mating dispersal and

612   strong breeding group structure in a mid-sized marsupial mammal (*Petrogale penicillata*).

613   *Molecular Ecology*, **15**, 2997–3007.

614 Heimlich-Boran JR (1993) Social organization of the short-finned pilot whale, *Globicephala*

615   *macrorhynchus*, with special reference to the comparative social ecology of delphinids. PhD

616   Thesis, University of Cambridge.

617 Hodges E, Rooks M, Xuan ZY, Bhattacharjee A, Gordon DB, Brizuela L, … Hannon, GJ (2009)

618   Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively

619   parallel sequencing. *Nature Protocols*, **4**, 960–974.

620 Janik VM, Slater PJB (1997) Vocal learning in mammals. *Advances in the Study of Behaviour*,

621   **26**, 59–100.

622 Kasuya T, Miyashita T, Kasamatsu F (1988) Segregation of two forms of short-finned pilot

623   whales off the Pacific coast of Japan. *The Scientific Reports of the Whales Research Institute*,

624   **39**, 77–90.

625 Katoh M, Kuma M (2002) MAFFT: a novel method for rapid multiple sequence alignment based

626   on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059–3066.

627 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, … Drummond, Alexei

628   (2012) Geneious Basic: An integrated and extendable desktop software platform for the

629   organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

630 Kershenbaum A, Ilany A, Blaustein L, Geffen E (2012) Syntactic structure and geographical

631   dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological*

632   *Sciences*, **279**, 2974–2981.

633 Kessler SE, Radespiel U, Hasiniaina AIF, Leliveld LMC, Nash LT, Zimmermann E (2014)

634   Modeling the origins of mammalian sociality: moderate evidence for matrilineal signatures in

635   mouse lemur vocalizations. *Frontiers in Zoology*, **11**, 14.

636 de la Torre S, Snowdon CT (2009) Dialects in pygmy marmosets? Population variation in call

637   structure. *American Journal of Primatology*, **71**, 333–42.

638   Lachlan RF, Slater PJB (1999) The maintenance of vocal learning by gene-culture interaction:

639   the cultural trap hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, **266**,

640   701-706.

641   Laland KN (1992) A theoretical investigation of the role of social transmission in evolution.

642   *Ethology and Sociobiology*, **13**, 87–113.

643   Laland KN, Janik VM (2006) The animal cultures debate. *Trends in Ecology & Evolution*, **21**,

644   542–7.

645   Laland KN, Odling-Smee J, Myles S (2010) How culture shaped the human genome: bringing

646   genetics and the human sciences together. *Nature Reviews Genetics*, **11**, 137–148.

647   Luikart G, Cornuet JM (1999) Estimating the effective number of breeders from heterozygote

648   excess in progeny. *Genetics,* **151**:1211–1216

649   Mahaffy SD, Baird RW, McSweeney DJ, Webster DL, Schorr GS (2015) High site fidelity,

650   strong associations, and long-term bonds: Short-finned pilot whales off the island of Hawaiʻi.

651   *Marine Mammal Science*, **31**, 1427–1451.

652   Marsh H, Kasuya T (1986). Evidence for reproductive senescence in female cetaceans. *Reports*

653   *of the International Whaling Commission* **8**, 57–74.

654   Martien KK, Chivers SJ, Baird RW, Archer FI, Gorgone AM, Hancock-Hanser BL, …Taylor BL

655   (2014) Nuclear and mitochondrial patterns of population structure in North Pacific false

656   killer whales (*Pseudorca crassidens*). *Journal of Heredity*, **105**, 611–626.

657   McComb K, Semple S (2005) Coevolution of vocal communication and sociality in primates.

658   *Biology Letters*, **1**, 381–385.

659   Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed

660   target capture and sequencing. *Cold Spring Harbor Protocols*, **2010**, 1-10.

661   Milligan BG (2003) Maximum-likelihood estimation of relatedness. *Genetics*, *163*, 1153–1167.

662   Morin PA, Archer FI, Pease VL, Hancock-Hanser BL, Robertson KM, Huebinger RM, …Taylor,

663   BL (2012) An empirical comparison of SNPs and microsatellites for population structure,

664   assignment, and demographic analyses of bowhead whale populations. *Endangered Species*

665   *Research*, **19**, 1–27.

666   Morin PA, Parsons KM, Archer FI, Ávila-Arcos MC, Barrett-Lennard LG, Dalla RL, … Foote,

667   AD (2015) Geographical and temporal dynamics of a global radiation and diversification in

668   the killer whale. *Molecular Ecology*, **24**, 3964–3979.

669 Mundinger PC (1980) Animal cultures and a general theory of cultural evolution. *Ethology and*
670 *Sociobiology*, **1**, 183–223.

671 Oremus M, Gales R, Dalebout ML, Funahashi N, Endo T, Kage T, … Baker SC (2009)
672 Worldwide mitochondrial DNA diversity and phylogeography of pilot whales (*Globicephala*
673 *spp.*). *Biological Journal of the Linnean Society*, **98**, 729–744.

674 Parreira BR, Chikhi L (2015) On some genetic consequences of social structure, mating systems,
675 dispersal, and sampling. *Proceedings of the National Academy of Sciences*, **112**, E3318–
676 E3326.

677 Pew J, Muir PH, Wang J, Frasier TR (2014) related: an R package for analyzing pairwise
678 relatedness from codominant molecular markers. *Molecular Ecology Resources*, **15**, 557–
679 561.

680 Pope TR (1992) The influence of dispersal patterns and mating systems on genetic
681 differentiation within and between populations of the red howler monkey (*Alouatta*
682 *seniculus*). *Evolution*, **46**, 1112–1128.

683 Prout T (1981) A note on the island model with sex dependent migration. *Theoretical Applied*
684 *Genetics,* **59**:327–332.

685 R Core Team (2016) R: A Language and Environment for Statistical Computing (https://www.R-
686 project.org).

687 Rendell L, Mesnick SL, Dalebout ML, Burtenshaw J, Whitehead H (2012) Can Genetic
688 Differences Explain Vocal Dialect Variation in Sperm Whales, *Physeter macrocephalus.*
689 *Behavior Genetics*, **42**, 332–43.

690 Rendell L, Whitehead H (2001) Culture in whales and dolphins. *Marine Ecology Progress*
691 *Series*, **52**, 309–382.

692 Rendell LE, Whitehead H (2003) Vocal clans in sperm whales (*Physeter macrocephalus*).
693 *Proceedings of the Royal Society B: Biological Sciences*, **270**, 225–231.

694 Riesch R, Barrett-Lennard LG, Ellis GM, Ford JKB, Deecke VB (2012) Cultural traditions and
695 the evolution of reproductive isolation: ecological speciation in killer whales? *Biological*
696 *Journal of the Linnean Society*, **106**, 1–17.

697 Riesch R, Ford JKB, Thomsen F (2006) Stability and group specificity of stereotyped whistles in
698 resident killer whales, *Orcinus orca*, off British Columbia. *Animal Behaviour*, **71**, 79–91.

Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **357**, 493–503.

Sugg DW, Chesser RK Dobson FS, Hoogland JL (1996) Population genetics meets behavioral ecology. *Trends in Ecology and Evolution*, **11**, 338–342.

Van Cise AM, Morin PA, Baird RW, Lang AR, Robertson KM, Chivers SJ, …Martien, KK (2016) Redrawing the map: mtDNA provides new insight into the distribution and diversity of short-finned pilot whales in the Pacific Ocean. *Marine Mammal Science*, **32**, 1177–1199.

Van Cise AM, Roch MA, Baird RW, Aran Mooney T, Barlow J (2017) Acoustic differentiation of Shiho- and Naisa-type short-finned pilot whales in the Pacific Ocean. *The Journal of the Acoustical Society of America*, **141**, 737–748.

Wade PR, Reeves RR, Mesnick SL (2012) Social and behavioural factors in cetacean responses to overexploitation: Are odontocetes less "resilient" than mysticetes? *Journal of Marine Biology*, **2012**, 1–15.

Wang IJ, Summers K (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology*, **19**, 447–458.

Weilgart L, Whitehead H (1997) Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. *Behavioral Ecology and Sociobiology*, **40**, 277–285.

Whitehead H (1998) Cultural Selection and Genetic Diversity in Matrilineal Whales. *Science*, **282**, 1708–1711.

Whitehead H (2007) Learning, climate and the evolution of cultural capacity. Journal of *Theoretical Biology*, **245**, 341–350.

Whitehead H (2008) Analyzing animal societies: quantitative methods for vertebrate social analysis. University of Chicago Press.

Whitehead H (2009) SOCPROG programs: Analysing animal social structures. *Behavioral Ecology and Sociobiology*, **63**, 765–778.

Williams R, Lusseau D (2006) A killer whale social network is vulnerable to targeted removals. *Biology Letters*, **2**, 497–500.

Wittemyer G, Okello JBA, Rasmussen HB, Arctander P, Nyakaana S, Douglas-Hamilton I, Siegismund HR (2009) Where sociality and relatedness diverge: the genetic basis for

730   hierarchical social organization in African elephants. *Proceedings of the Royal Society B:*

731   *Biological Sciences*, **276**, 3513–3521.

732 Yurk H, Barrett-Lennard LG, Ford JKB, Matkin CO (2002) Cultural transmission within

733   maternal lineages: vocal clans in resident killer whales in southern Alaska. *Animal*

734   *Behaviour*, **63**, 1103–1119.

735

Data availability

737   We have deposited the sequences used in these analyses in GenBank. Accession numbers

738 for mtDNA haplotypes are: KM624043, KM624044, KM624054, KM624055, KM624058, and

739 KM624059. Accession numbers for nuclear sequences generated for SNP discovery are

740 MG023261-MG023309. The *Tursiops truncatus* reference sequence and SNP genotype data are

741 available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.78521.

742

743 Figure 1. Sampling locations for samples used in this study. *Above:* samples used in mtDNA

744 analyses. Symbols represent their stratification for geographic structure analyses. Inset shows

745 additional samples from the NWHI and Pelagic strata. *Below:* samples used in SNP analyses.

746 Symbols represent their stratification for genetic structure analyses. Samples labeled "No Link"

747 are presumed to belong to the pelagic stratum, because they cannot currently be linked to any

748 social stratum within the Main Hawaiian Islands. Inset shows social units and clusters in the

749 Eastern Community that were used for relatedness analyses.

750

751
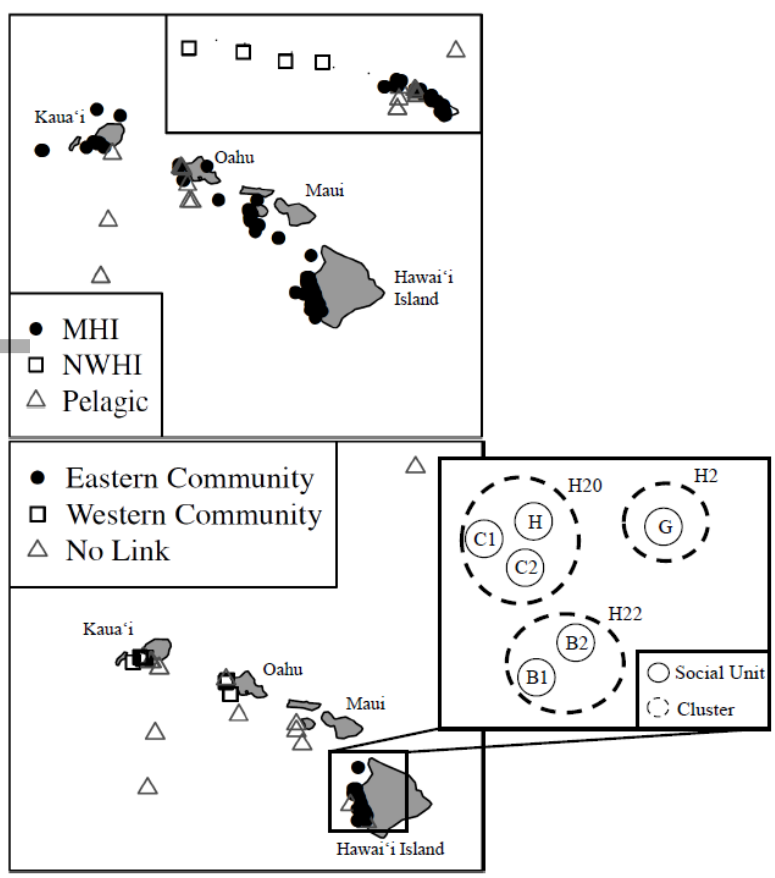
Figure 2. Relatedness analysis for three social units with at least five individuals sampled, and
overall relatedness within groups (bottom right). Arrows indicate average within-group
relatedness; histograms show the expected distribution of within-group relatedness values if
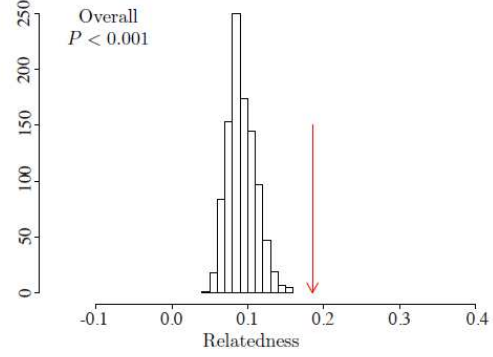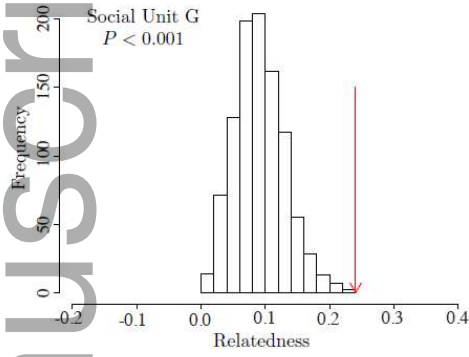groups were randomly organized but retained their original sample size.
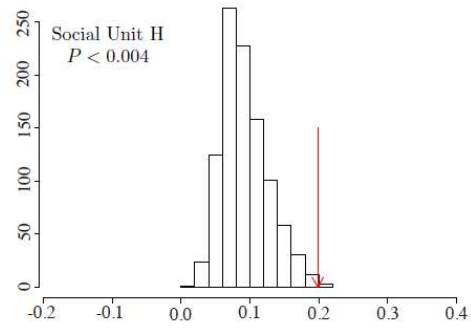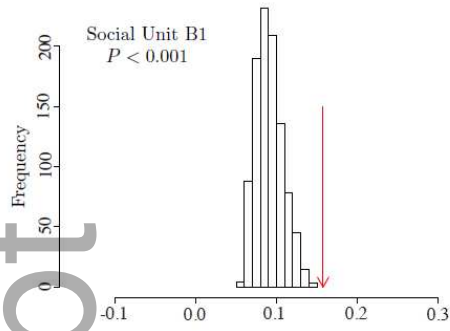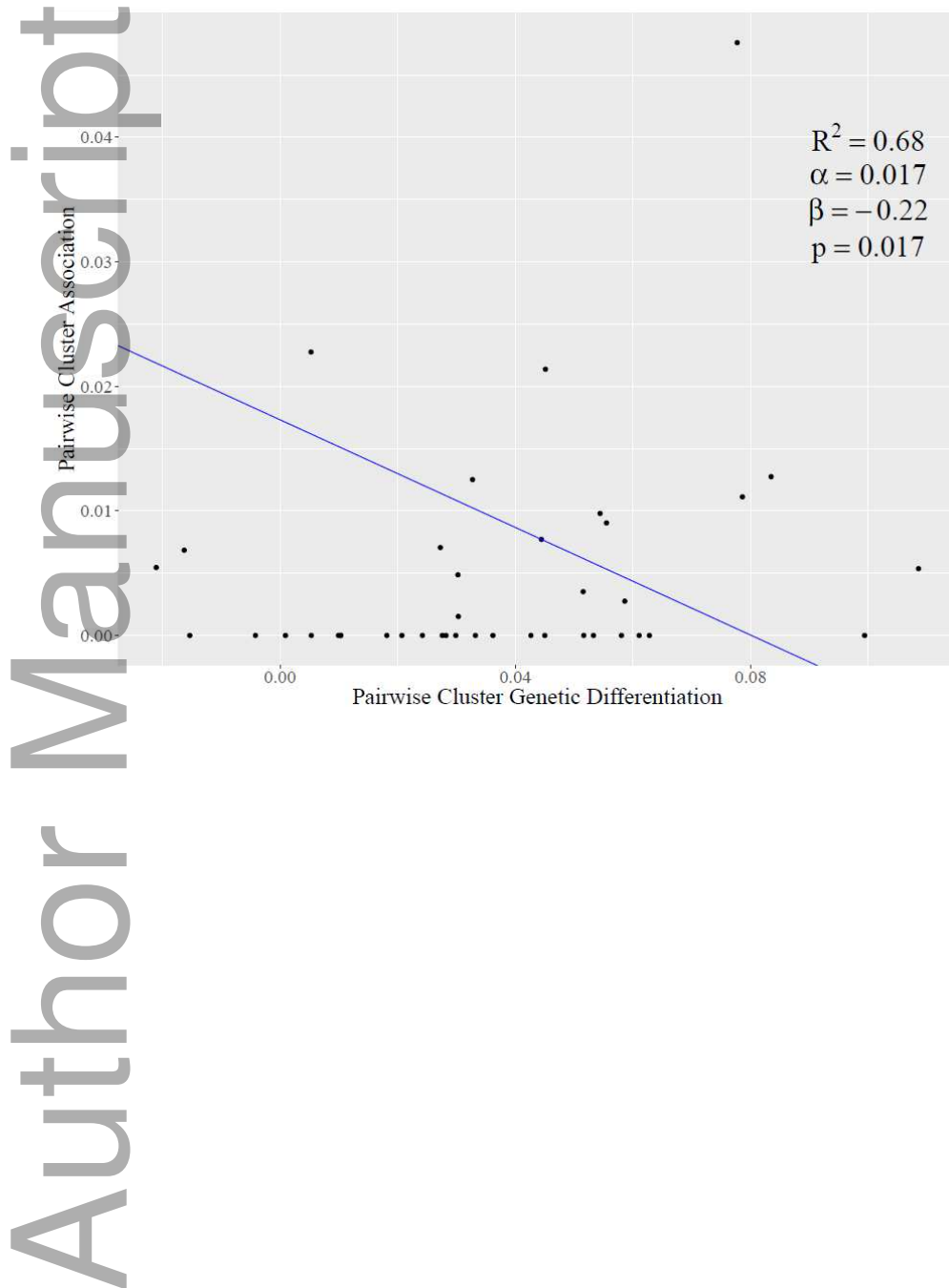
756

757

758

Figure 3. Fixed effect linear regression comparing pairwise genetic differentiation (F_{ST}) among clusters with average association index, or rate of association, among clusters. Association index is calculated using a half-weight index and a sampling period of one day, to control for effort.

1   Supplemental Table S1. Summary metrics for 119 SNP loci included in this study.

2   Supplemental Table S2. Sample stratification levels used for statistical analyses in this study.

3   Bold values in the Cluster column indicate samples that were removed before cluster F_{ST}

4   analysis due to high relatedness to other samples in the study.

5   Supplemental Table S3. Genetic differentiation (F_{ST}) between five clusters with more than five

6   sampled individuals (related individuals included). Significant differentiation between clusters

7   (p-value, final column) is shown in bold.

Table 1. Molecular diversity indices for SNP and mtDNA datasets. MHI SNP data were tested using sub-sampled datasets so that diversity indices within strata were not biased by sampling technique. "All samples" includes all samples included in the study. Nuclear samples were subsampled within the eastern and western communities. N = sample size, H_o = observed heterozygosity, H_e = expected heterozygosity.

	mtDNA N	Theta (θ_H)	Haplotype diversity (h)	Nucleotide diversity (π)	SNP N	Ave. num alleles	H_o	H_e
All samples	242	0.06	0.08 ± 0.02	0.004	106	4	0.45	0.45
Regions								
MHI	205	0.007	0.01 ± 0.01	0.004	63	3.9	0.46	0.46
Western MHI Community	--	--	--	--	21	3.5	0.49	0.47
Eastern MHI Community	--	--	--	--	42	3.7	0.45	0.45
NWHI	17	0.33	0.44 ± 0.1	0.004	--	--	--	--
Pelagic	20	0.27	0.36 ± 0.1	0.004	--	--	--	--

Table 2. Mitochondrial haplotype distribution by stratum in the Hawaiian Islands.

Stratum	MHI	NWHI	Pelagic
Haplotype			
J	204	12	16
C	1	0	0
K	0	0	2
12	0	5	0
11	0	0	1
2	0	0	1

Table 3. Geographic population differentiation in Hawaiian Island short-finned pilot whales. For SNP data, only F_{ST} was calculated; for mtDNA data, both F_{ST} and Φ_{ST} were calculated. Sample sizes for each stratum are shown in parentheses. Significant values are shown in bold.

Stratum	F_{ST}	F_{ST} P-value	Φ_{ST}	Φ_{ST} P-value
mtDNA				
MHI (204) v. NWHI (17)	0.67	**<0.001**	0.58	**<0.001**
MHI (204) v. Pelagic (20)	0.39	**<0.001**	0.30	**<0.001**
NWHI (17) v. Pelagic (20)	0.08	0.07	0.01	0.28
SNP				
Eastern MHI Community (42) v. Western MHI Community (21)	0.01	**0.009**	NA	NA

Table 4. Genetic differentiation (F_{ST}) between five clusters with more than five sampled individuals (related individuals not included); sample sizes for each cluster are shown in parentheses. F_{ST} P-values (in parentheses) are shown below F_{ST} values; significant differentiation between clusters is shown in bold.

	Eastern Community Cluster 2 (8)	Eastern Community Cluster 20 (10)	Eastern Community Cluster 22 (10)	Western Community Cluster 13 (9)
Eastern Community Cluster 20 (10)	**0.05** **(<0.001)**			
Eastern Community Cluster 22 (10)	**0.06** **(<0.001)**	**0.04** **(0.002)**		
Western Community Cluster 13 (9)	**0.02** **(0.05)**	**0.02** **(0.04)**	0.01 (0.12)	
Western Community Cluster 24 (6)	**0.05** **(0.02)**	**0.03** **(0.03)**	**0.02** **(0.04)**	0.002 (0.39)

MHI
NWHI
Pelagic

Kaua'i
Oahu
Maui
Hawai'i Island

Eastern Community
Western Community
No Link

Kaua'i
Oahu
Maui
Hawai'i Island

H20
H2
H22
C1
H
C2
G
B2
B1

Social Unit
Cluster

$R^2 = 0.68$

$\alpha = 0.017$

$\beta = -0.22$

$p = 0.017$