Crowd intelligence can discern between repertoires of killer whale ecotypes

Anastasya Yu. Danishevskaya[1], Olga A Filatova[1*], Filipa I P Samarra[2], Patrick J O Miller[3], John K B Ford[4], Harald Yurk[5], Craig O Matkin[6], Erich Hoyt[7]

corresponding author email: alazorro@gmail.com, ORCID 0000-0003-1533-4807

[1]*Department of Vertebrate Zoology, Faculty of Biology, Moscow State University, Moscow 119991, Russia*

[2]*Marine and Freshwater Research Institute, Skulagata 4, 101 Reykjavik, Iceland*

[3]*Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, Fife KY168LB, Scotland*

[4]*Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T1K6, Canada*

[5]*JASCO Research Ltd., 2305-4464 Markham Street, Victoria, British Columbia V8Z7X8, Canada*

[6]*North Gulf Oceanic Society, Homer, Alaska 99603, USA*

[7]*Whale and Dolphin Conservation, Park House, Allington Park, Bridport, Dorset DT65DD, United Kingdom*

23   Abstract

24   Call classifications by human observers are often subjective yet they are critical to studies of

25   animal communication, because only the categories that are relevant for the animals themselves

26   actually make sense in terms of correlation to the context. In this paper we test whether

27   independent observers can correctly detect differences and similarities in killer whale repertoires.

28   We used repertoires with different a priori levels of similarity: from different ecotypes, from

29   different oceans, from different populations within the same ocean, and from different local

30   subpopulations of the same population. Calls from nine killer whale populations/subpopulations

31   were pooled into a joint sample set, and eight independent observers were asked to classify the

32   calls into separate categories. None of the observers' classifications strongly followed the known

33   phylogeny of the analyzed repertoires. However, some phylogenetic relationships were reflected

34   in the classifications substantially better than others. Most observers correctly separated the calls

35   from two North Pacific ecotypes. Call classifications averaged across multiple observers

36   reflected the known repertoire phylogenies better than individual classifications, and revealed the

37   similarity of repertoires at the level of subpopulations within the same population, or closely

38   related populations.

39

40   Keywords: crowd intelligence; categorization; killer whale; dialect.

41

Introduction

Perception of animal sounds by human observers can vary depending on different factors. Call classifications are often subjective yet they are critical to studies of animal communication, because only the categories that are relevant for the animals themselves actually make sense in terms of correlation to the context. Human observer classifications rely on patterns that are relevant for humans, but not necessarily for the animals that produce the classified sounds. For example, Nelson and Soha (2004) found that the reaction of male sparrows *Zonotrichia leucophrys pugetensis* to the song was inconsistent with human observer classifications. The response to playback mostly depended on the variation in the song's trill, but no evidence was found that the birds recognized 'superdialects' categorized by human observers.

The killer whale (*Orcinus orca*) has a broad acoustic repertoire that includes stereotyped call types with a distinct structure. Their calls are transmitted across generations through learning rather than genetically. Different pods have specific call repertoires, and pods with similar repertoires comprise acoustic clans. Different populations normally have completely different call sets and can include one or more clans. Calls change over time through learning errors and innovations (Ford, 1991; Deecke et al., 2000), and consequently the repertoires of closely related, recently diverged groups are generally more similar, than in the more distantly related groups (Ford, 1991; Yurk et al., 2002; Filatova et al., 2017).

Killer whales occur in a number of local forms, or ecotypes, in various regions of the world. In the North Pacific, three ecotypes have been described. Two coastal ecotypes include 'residents' that mostly feed on fish and usually travel in large stable social units of maternally related whales (Ford et al., 2000; Ivkovich et al., 2010), and 'transients' that mostly feed on marine mammals and typically travel in smaller, more fluid social groups (Baird and Dill, 1996; Ford et al., 1998). There is also an offshore ecotype with a pelagic distribution, that appears to feed on sharks (Ford et al., 2011). The terms 'resident' and 'transient' killer whales have been

historically used to identify killer whale ecotypes in British Columbia, because the occurrence of 'resident' killer whales was more predictable, and researchers initially thought that the residency pattern was a key difference between them (Ford et al., 2000). However, more important differences were discovered later, including feeding preferences and genetic divergence (Ford et al., 1998). The terms 'residents' and 'transients' proved to be misleading because in other species they were typically used to refer to territory owners vs. vagrants (Burt, 1943). In this paper we refer to 'residents' as 'R-type' killer whales and to 'transients' as 'T-type' killer whales.

Ecotype differentiation can influence call similarity. Foote and Nystuen (2008) found that the minimum frequency of calls from a T-type population was significantly lower than that of a sympatric R-type population. Filatova et al., (2015) showed that frequency of calls from four North Pacific R-type and two North Atlantic populations did not differ significantly, while frequency of calls of two T-type populations was significantly lower. The difference in call frequencies could be related to ecological specialization or to the phylogenetic history of these populations.

Call categorization in killer whales has been traditionally performed by human observers based on characteristic features of sonograms (Ford, 1989; Strager, 1995; Yurk et al., 2002). No specific call types were found to occur exclusively in particular contexts (Ford, 1989; Filatova et al., 2013) except herding calls of Icelandic killer whales which are used to herd herring (Simon et al., 2006) and the cross-population excitement call (Rehn et al., 2011). However, the system of group dialects provides consistent and easily observed variation in repertoires both within and between populations. Therefore, different groups of whales produce different sets of call types, confirming that differences between repertoires are perceived by the whales and are relevant to them.

91   Various methods of similarity measurements of call repertoires have been used. Ford (1991)

92   used the Dice index based on the number of calls shared between pods; however, this method

93   works only if the measured repertoires share some call types, which typically occurs only within

94   clans of the same population. Deecke and Janik (2006) suggested a method of dynamic time

95   warping of fundamental frequency contours, which was used later by Filatova et al. (2012) to

96   compare repertoires of different populations across the North Pacific. Shamir et al. (2014)

97   compared calls of killer and pilot whales from different locations using computer methods of

98   image analysis and 'citizen scientists'. The computer analysis automatically categorized the calls

99   according to species and locations. Citizen scientists were able to categorize the calls to

100   locations, but the separation was somewhat less accurate compared to the computer method.

101   In this paper we test whether independent observers can detect differences and similarities in

102   killer whale repertoires that arise from the long history of independent cultural evolution. We

103   used repertoires with different a priori level of similarity: from different ecotypes, from different

104   oceans, from different populations within the same ocean, and from different local

105   subpopulations of the same population. Based on genetic studies (Morin et al., 2010) and

106   previous acoustic analyses (Filatova et al., 2015), we hypothesize that the repertoires of T-type

107   groups would show the highest divergence, while R-type and North Atlantic repertoires would

108   display a higher level of similarity to each other." Among the latter, we hypothesize that

109   subpopulations from within the same population will display a higher level of repertoire

110   similarity compared to subpopulations from different populations.

111

112   **Methods**

113   ***The study populations***

114   'Southern Resident' R-type killer whales inhabit the coastal waters of southern British Columbia

115   and Washington State. The core area of this population is located in the waters off southern

116 Vancouver Island, but they sometimes range as far south as Monterey Bay, California (Fig. 1).

117 The population consists of the single acoustic clan (Ford, 1991).

118 'Northern Resident' R-type killer whales inhabit the coastal waters of south-central British

119 Columbia to southeastern Alaska. The population is comprised of three acoustic clans: A, G and

120 R clans (Ford, 1991).

121 Alaskan R-type killer whales range from southeastern Alaska to Kodiak Island and possibly into

122 the Aleutian Islands and Bering Sea (Muto et al., 2017). This population includes two acoustic

123 clans: AB and AD (Yurk et al., 2002).

124 Some overlap exists in the ranges of the northeastern Pacific R-type populations: Southern

125 Residents overlap with Northern Residents in the waters off Vancouver Island, and Northern

126 Residents overlap with the Alaskan population in southeastern Alaska (Ford et al., 2000).

127 Despite this overlap, members of different populations do not associate (Ford et al., 2000) or

128 interbreed (Barrett-Lennard, 2000).

129 Eastern Kamchatka R-type killer whales are encountered along the eastern coast of Kamchatka

130 peninsula from Avacha Gulf to Karaginsky Gulf and east to the Commander Islands. This

131 population includes three acoustic clans. Killer whales frequenting the Avacha Gulf sometimes

132 visit the Commander Islands and socialize with the local whales, and vice versa, but most groups

133 typically occur within their normal range in either of these regions (Shabalina et al., 2015).

134 Therefore, we consider resident killer whales from Avacha Gulf and the Commander Islands

135 belong to the same population, but to different local subpopulations.

136 Eastern Aleutian T-type killer whales range from Gulf of Alaska to Aleutian Islands and Bering

137 Sea (Muto et al., 2017). West Coast Transients range from southern Alaska to California (Muto

138 et al., 2017).

139 Icelandic killer whales occur in the waters around Iceland and some whales seasonally travel to

140 Scotland (Samarra and Foote, 2015). Norwegian killer whales occur in the coastal and offshore

141 waters off Norway. Earlier comparisons of photo-identified individuals suggested no movements

142 between Iceland and Norway, but genetically they belong to two populations, one occurring both

143 in Iceland and Norway, and another – in Iceland only (Foote et al., 2011). Given the genetic

144 connection between Icelandic and Norwegian killer whales, here we consider them as

145 subpopulations of the same population.

146 *Data collection*

147 Acoustic recordings used for this study were collected over various field projects by means of a

148 variety of equipment. However, all recordings were made at a sampling rate of 44.1 kHz or

149 higher, enabling systematic evaluation of call features <22.05 kHz.

150 Recordings of the Northern, Southern and Alaskan R-type and the West Coast T-type killer

151 whale populations were taken from existing long-term recording databases. Recordings of

152 Southern Residents were made from 1980 to 2009, Northern Residents from 1988 to 1999,

153 Alaskan whales from 1984 to 2008.

154 Calls from Kamchatkan R-type whales were recorded in 2000-2014 off the southeastern coast of

155 Kamchatka peninsula, and calls from the Commander Islands R-type whales were recorded in

156 2008-2015 off the southwestern coast of Bering Island as part of the Far East Russia Orca

157 Project.

158 Eastern Aleutian T-type killer whales were recorded between 2003-2008 in False Pass (eastern

159 Aleutian Islands) as part of a project conducted by the North Gulf Oceanic Society.

160 Calls from Icelandic killer whales were obtained between 2008 and 2014 using both digital

161 acoustic recording tags (Dtags, Johnson and Tyack, 2003) and various other recording systems at

162 two main locations: in Grundarfjörður and Kolgrafarfjörður on the Snæfellsnes Peninsula (West

163 Iceland) and in Vestmannaeyjar, off the southern coast of Iceland.

164   Calls from Norwegian killer whales were obtained between 2005-2009 using Dtags in the

165   northern Norwegian Vestfjord fjord system and off Vesterållen.

166   *Call categorization*

167   Stereotyped calls of killer whales can be organized into discrete categories – call types – based

168   on distinctive structural characteristics, and some call types show structural variation allowing to

169   distinguish several subtypes (Ford 1989). Calls from four R-type populations (Northern and

170   Southern Resident, Alaskan and Kamchatka-Commander) and two T-type populations (West

171   Coast and Eastern Aleutian) from the North Pacific and one population of killer whales from the

172   North Atlantic (Icelandic-Norwegian) were assigned to types and subtypes according to existing

173   catalogues. We selected two call samples from each type and/or subtype from each population

174   and subpopulation to cover its entire repertoire of stereotyped calls. All these samples were

175   pooled into a joint sample set, so that calls had no indication of population and type. In total, the

176   sample set included 780 calls (120 calls from Northern Resident R-type, 62 from Southern

177   Resident R-type, 96 from Alaskan R-type, 112 from Kamchatkan R-type, 108 from the

178   Commander Islands R-type, 34 from Eastern Aleutian T-type, 36 from West Coast T-type, 72

179   from Norwegian, and 140 from Icelandic killer whales).

180   To measure the similarity of repertoires of different populations/subpopulations, we used

181   independent human observers that were asked to divide the whole sample set into separate

182   categories (the number of categories was not specified). The categories were not supposed to

183   match any of the real call types in the repertoire catalogues – their only purpose was to measure

184   how many calls from each population/subpopulation fell into the same category as calls from

185   each other population/subpopulation, i.e. how many calls of each pair of

186   populations/subpopulations the observers perceived as similar. On the next stage, the similarity

187   of each pair of population/subpopulation repertoires was calculated according to the number of

188   shared calls.

189   Eight independent observers were asked to classify the calls. One observer had previous

190   experience with killer whale calls, one had experience with bird sounds, and the other six

191   observers had no previous experience with animal sounds. During categorization, the observers

192   could perceive the calls both as sounds and sonograms simultaneously, which was achieved

193   using the program ACDSee 8.1 that allows to view images and play the corresponding sound

194   files if they are placed in the same folder and have matching file names. Sonograms were saved

195   through AviSoft SASLab Pro as images with the following settings: cut-off frequency 12 kHz,

196   Hamming window, FFT-length – 512, overlap – 87.5%. See Fig. 2 for an example of call

197   categorisation.

198   After the categorization was complete, we counted the number of calls from each

199   population/subpopulation in each category. The distances between populations/subpopulations

200   were calculated using the Dice coefficient with the formula $2 * N_{shared} /(N1 + N2)$, where N_{shared}

201   is the number of categories shared by two populations/subpopulations, and N1 and N2 are the

202   repertoire sizes of each population/subpopulation. Based on the Dice coefficients, we created a

203   similarity matrix for each observer and constructed dendrograms using the average linkage

204   method in R (R Core Team 2015).

205   To create the average dendrogram across all eight observers, we averaged the similarity value

206   between each pair of populations/subpopulations across all observers. From the resulting

207   similarity matrix, we constructed the dendrogram using the average linkage method.

208   **A priori relatedness across populations**

209   Killer whale dialects are learned mostly within matrilineal groups, so we calculated a priori

210   relatedness across populations/subpopulations using maternally transmitted mitochondrial

211   genomes. We selected one most common haplotype of complete mitochondrial DNA per

212   population/subpopulation (Morin et al. 2015; Genbank accession numbers: Alaskan and

213   Northern Resident R-type – GU187193; Southern Resident R-type – GU187195; Kamchatka and

214   Commander Islands R-type – GU187196; West Coast T-type – GU187162; Eastern Aleutian T-

215   type – GU187174; Icelandic – GU187180; Norwegian – GU187186). A priori relatedness

216   between populations/subpopulations was calculated as phylogeny of complete mitochondrial

217   genomes using program Mr. Bayes (Ronquist et al. 2012). Phylogenetic tree (Fig. 3) and matrix

218   of pairwise distances (Table 1) were assembled in R package *ape* (Paradis et al. 2004).

219   Matrix of pairwise phylogenetic distances was compared to the averaged similarity matrix of call

220   repertoires with Mantel test in R package *vegan* (Oksanen et al. 2018).

221   **Results**

222   In agreement with the a priory measures of population/subpopulation relatedness (Table 1, Fig.

223   3), four out of eight observers divided the repertoires into two clusters, one containing T-type

224   and another – R-type and North Atlantic repertoires (Fig. 4). In all but one observers, the two T-

225   type repertoires were placed at the adjacent branches of the same cluster.

226   Three observers classified calls from the different local stocks of the same population (from

227   Kamchatka and the Commander Islands) to the adjacent branches of the same cluster. Three

228   observers classified calls from the geographically close and related Alaskan and Northern

229   Resident R-type populations to the adjacent branches of the same cluster. North Atlantic

230   populations from Iceland and Norway never occurred at adjacent branches.

231   The averaged dendrogram summarized the tendencies of the individual observers (Fig. 4, Table

232   2). Repertoires of T-type populations were classified as a separate cluster from all other, being

233   20% similar to one another and only 15.6% similar to the closest of the other repertoires –

234   Alaskan R-type. Alaskan and Northern Resident R-type populations were the most similar, with

235   similarity level of 21.9%. Kamchatka and the Commander Islands stocks were placed at the

236   adjacent branches of the same cluster with similarity level 19.1%. The averaged distance matrix

237   correlated significantly to the relatedness matrix calculated from complete mitochondrial

238   genome phylogeny (Mantel test, $r = 0.662$, $p < 0.01$).

10

Discussion

None of the observers' classifications followed strongly the a priory phylogeny of the analyzed repertoires. However, some phylogenetic relationships were reflected in the classifications substantially better than others. All but one of the observers correctly aggregated the calls from the two T-type populations, and four of them reproduced the expected separation between T-type vs R-type and North Atlantic repertoires. Three observers recognized the repertoire similarity of different stocks of the same population (Kamchatka and the Commander Islands) and of geographically close populations (Alaskan and Northern R-type). On the other hand, none of the observers identified the similarity of the geographically close Northern and Southern R-type populations and the two subpopulations from the North Atlantic.

The averaged dendrogram was closer to the phylogeny of complete mitochondrial genomes of the study populations/subpopulations than any of the individual dendrograms. The most diverged cluster consisted of two T-type populations, while the repertoires of the R-type and North Atlantic populations were mixed, with aggregated Kamchatka/the Commander Islands and Alaskan/Northern R-type repertoires. The averaged dendrogram summarized the patterns that were consistent among several observers and discarded the inconsistent patterns, suggesting that the 'wisdom of crowds' (Surowiecki, 2004) principle works well for the crowdsource analyses of call repertoire similarities.

T-type killer whales are currently considered the most diverged group of killer whales (Morin et al., 2010), while R-type and North Atlantic killer whales are more closely related to each other. In our study, untrained observers could discern between calls of T-type vs R-type and North Atlantic killer whales. Some previous studies also demonstrated the differences between the repertoires of T-type whales and other ecotypes. Foote and Nystuen (2008) found that the minimum frequency of calls from a T-type population was significantly lower than that of a

11

264   sympatric R-type population. Filatova et al. (2015) showed that the frequency of calls from four

265   North Pacific R-type and two North Atlantic populations was similar, while the frequency of

266   calls of two T-type populations was significantly lower. This frequency difference could be the

267   key for the separation of T-type calls by individual observers in our study. However, humans

268   usually pay more attention to contour shape rather than to the absolute frequency values (Janik,

269   1999). The fact that human observers discerned T-type repertoires from all others suggests that

270   either T-type repertoires differ not only by frequency, but also by contour shape, or their

271   frequency difference is so pronounced that it is obvious even for humans.

272   None of the observers was able to make a clear distinction between R-type and North Atlantic

273   repertoires, though they formed distinct clusters on the a priory relatedness tree. Given the

274   sensitivity of human observers to frequency contour shape (Janik, 1999), this suggests that the

275   repertoires of R-type and North Atlantic killer whales do not differ significantly in the proportion

276   of different contour shapes. The frequency range in these two groups of killer whales was also

277   similar (Filatova et al., 2015).

278   Three observers classified Alaskan and Northern Resident R-type repertoires together, but none

279   of the observers classified Northern and Southern Resident R-type repertoires together. Despite

280   the significant range overlap with Southern Residents, genetically Northern Residents are closer

281   to the Alaskan R-type population (Barrett-Lennard, 2000) either due to common origin or due to

282   recent gene flow. Repertoire similarity is more consistent with the common origin hypothesis,

283   but gene flow and underlying social interactions could provide some cultural exchange, also

284   leading to the increase in repertoire similarity.

285   Three observers classified calls from different stocks of the same population (Kamchatka and the

286   Commander Islands) to adjacent branches of the same cluster. As some families were observed

287   both in Kamchatka and the Commander Islands and socialized with local families (Shabalina et

288   al., 2015), these stocks can be considered belonging to the same population. The fact that some

observers classified them separately suggests that even within a population call variation can be so pronounced that stock repertoires are perceived as different. Some R-type populations include several clusters of whales with totally different repertoires – so-called 'clans' (Ford, 1991). The Kamchatka-Commander population comprises at least three clans; one of them mostly occurs in Kamchatka, and two others are more common in the Commander Islands (our unpublished data). This differentiation can drive the perceived variation in repertoires among these two stocks.

Norwegian, Icelandic and Southern Resident populations were not classified in any consistent pattern. Killer whale calls change in time, but their structural variation is limited by the physical and/or cultural constrains (Filatova et al., 2016). Therefore, diverging populations reach the maximum repertoire divergence, and after that their similarity can no longer decrease, but instead it can increase due to random convergences. This process is most likely responsible for the inconsistent similarity pattern of Norwegian, Icelandic and Southern Resident populations.

In summary, it appears that R-type killer whales and North Atlantic killer whales do not have inherent differences, and due to fast change and variance saturation they often are perceived as different even when related and recently diverged. Transients, or T-type killer whales, apparently have some significant differences that are inherent (or culturally selected) to some extent (otherwise they would be saturated and potentially masked as in others).

Thus our results demonstrate that untrained observers can discern between repertoires of North Pacific T-type killer whales vs North Pacific R-type and North Atlantic killer whales, but cannot discern between R-type vs North Atlantic repertoires. Additionally, averaged classifications by multiple observers are more precise than individual classifications, and can reveal the similarity of repertoires at the level of subpopulations within the same population, or closely related populations.

323 **References**

324 Barrett-Lennard LG. 2000. Population structure and mating patterns of killer whales, *Orcinus*

325 *orca*, as revealed by DNA analysis. PhD dissertation, University of British Columbia

326 Baird RW, Dill LM. 1996. Ecological and social determinants of group size in transient killer

327 whales. Behav Ecol. 7:408-416.

328 Burt WH. 1943. Territoriality and home range concepts as applied to mammals. J Mammal. 24:

329 346–352.

330 Deecke VB, Ford JKB, Spong P. 2000. Dialect change in resident killer whales: implications for

331 vocal learning and cultural transmission. Anim Behav. 60:629-638.

332 Deecke VB, Janik VM. 2006. Automated categorization of bioacoustic signals: avoiding

333 perceptual pitfalls. J Acoust Soc Am 119(1):645-653.

334 Filatova OA, Deecke VB, Ford JKB, Matkin CO, Barrett-Lennard LG, Guzeev MA, Burdin AM,

335 Hoyt E. 2012. Call diversity in the North Pacific killer whale populations: implications for

336 dialect evolution and population history. Anim Behav. 83(3):595-603.

337   Filatova OA, Guzeev MA, Fedutin ID, Burdin AM, Hoyt E. 2013. Dependence of killer whale

338   (*Orcinus orca*) acoustic signals on the type of activity and social context. Biology Bulletin

339   40(9):790-796.

340   Filatova OA, Miller PJO, Yurk H, Samarra FIP, Hoyt E, Ford JKB, Matkin CO, and Barrett-

341   Lennard LG. 2015. Killer whale call frequency is similar across the oceans, but varies across

342   sympatric ecotypes. J Acoust Soc Am. 138:251-257.

343   Filatova OA, Samarra FIP, Barrett-Lennard LG, Miller PJO, Ford JKB, Yurk H, Matkin CO,

344   Hoyt E. 2016. Physical constraints of cultural evolution of dialects in killer whales. J Acoust Soc

345   Am. 140(5):3755-3764.

346   Filatova OA, Ivkovich TV, Guzeev MA, Burdin AM, Hoyt E. 2017. Social complexity and

347   cultural transmission of dialects in killer whales. Behav.154:171–194.

348   Foote AD, Nystuen JA. 2008. Variation in call pitch among killer whale ecotypes. J Acoust Soc

349   Am. 123:1747-1752.

350   Foote AD, Vilstrup JT, De Stephanis R, Verborgh P, Abel Nielsen SC, Deaville R, Pérez-Gil M.

351   2011. Genetic differentiation among North Atlantic killer whale populations. Mol Ecol.

352   20(3):629-641.

353   Ford JKB. 1989. Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver

354   Island, British Columbia. Can J Zool. 67(3):727-745.

355   Ford JKB. 1991. Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters

356   of British Columbia. Can J Zool. 69:1454-1483.

357   Ford JKB, Ellis GM, Balcomb KC. 2000. Killer whales: the natural history and genealogy of

358   *Orcinus orca* in British Columbia and Washington. UBC Press, Vancouver.

359   Ford JKB, Ellis GM, Barrett-Lennard LG, Morton AB, Palm RS, Balcomb KC. 1998. Dietary

360   specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British

361   Columbia and adjacent waters. Can J Zool. 76:1456–1471.

362   Ford JKB, Ellis GM, Matkin CO, Wetklo MH, Barrett-Lennard LG, Withler RE. 2011. Shark

363   predation and tooth wear in a population of northeastern Pacific killer whales. Aquat Biol.

364   11:213-224.

365   Ivkovich TV, Filatova OA, Burdin AM, Sato H, Hoyt E. 2010. The social organization of

366   resident-type killer whales (*Orcinus orca*) in Avacha Gulf, Northwest Pacific, as revealed

367   through association patterns and acoustic similarity. Mamm Biol. 75:198-210.

368   Janik VM. 1999. Pitfalls in the categorization of behaviour: a comparison of dolphin whistle

369   classification methods. Anim Behav. 57(1):133-143.

370   Johnson MP, Tyack PL. 2003. A digital acoustic recording tag for measuring the response of

371   wild marine mammals to sound. IEEE J Oceanic Eng. 28:3–12.

372   Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P, Durban J, Parsons K, Pitman R,

373   Li L, Bouffard P, Abel Nielsen SC, Rasmussen M, Willerslev E, Gilbert MT, Harkins T. 2010.

374   Complete mitchondrial genome phylogeographic analysis of killer whales (*Orcinus orca*)

375   indicates multiple species. Genome Res. 20:908-916.

376   Muto MM, Helker VT, Angliss RP, Allen BA, Boveng PL, Breiwick JM, Cameron MF,

377   Clapham PJ, Dahle SP, Dahlheim ME, Fadely BS, Ferguson MC, Fritz LW, Hobbs RC,

378   Ivashchenko YV, Kennedy AS, London JM, Mizroch SA, Ream RR, Richmond EL, Shelden

379   KEW, Towell RG, Wade PR, Waite JM, Zerbini AN. 2017. Alaska marine mammal stock

380   assessments, 2016. US Dept. Commer., NOAA Tech Memo NMFS-AFSC-355, 366 p.

381   Nelson DA, Soha JA. 2004. Perception of geographical variation in song by male Puget Sound

382   white-crowned sparrows, *Zonotrichia leucophrys pugetensis*. Anim Behav. 68(2):395-405.

383   Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara

384   RB, Simpson GL, Solymos PS, Henry M, Stevens H, Szoecs E, Wagner H. 2018. vegan:

385   Community Ecology Package. R package version 2.4-6. https://CRAN.R-

386   project.org/package=vegan

387   Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R

388   language. Bioinformatics 20:289-290.R Core Team. 2015. R: A language and environment for

389   statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

390   Rehn N, Filatova OA, Durban JW, Foote AD. 2011. Cross-cultural and cross-ecotype production

391   of a killer whale 'excitement' call suggests universality. Naturwissenschaften 98:1–6.

392   Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Hohna S, Larget B, Liu L,

393   Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference

394   and model choice across a large model space. Systematic Biology 61(3):539-42.

395   Samarra FIP, Foote AD. 2015. Seasonal movements of killer whales between Iceland and

396   Scotland. Aquat. Biol. 24(1):75-79.

397   Shabalina AO, Filatova OA, Ivkovich TV, Burdin AM, Hoyt E. 2015. Killer whales of

398   Southeastern Kamchatka and the Commander Islands: dynamics of occurrence and movement

399   between areas. Zoologichesky Zhurnal 94(3):352-364. [in Russian]

400   Shamir L, Yerby C, Simpson R, von Benda-Beckmann AM, Tyack P, Samarra F, ... and Wallin

401   J. 2014. Classification of large acoustic datasets using machine learning and crowdsourcing:

402   Application to whale calls. J Acoust Soc Am. 135:953-962

403   Simon M, Ugarte F, Wahlberg M, Miller LA. 2006. Icelandic killer whales *Orcinus orca* use a

404   pulsed call suitable for manipulating the schooling behaviour of herring *Clupea harengus*.

405   Bioacoustics. 16:57-74

406   Strager H. 1995. Pod-specific call repertoires and compound calls of killer whales, *Orcinus orca*

407   Linnaeus, 1758, in the waters of northern Norway. Can J Zool. 73:1037-1047.

408   Surowiecki J. 2004. The Wisdom of Crowds. Doubleday, New York.

409   Yurk H, Barrett-Lennard L, Ford JKB, Matkin CO. 2002. Cultural transmission within maternal

410   lineages: vocal clans in resident killer whales in southern Alaska. Anim Behav. 63:1103-1119.

411

412 Table 1. Pairwise distances between the pairs of killer whale populations based on branch
413 lengths of Bayesian phylogenetic tree calculated from complete mitochondrial DNA sequences.
414 AR – Alaskan R-type, NR – Northern R-type, SR – Southern R-type, Kam – Kamchatkan R-
415 type, Co – Commander Islands R-type, EAT – Eastern Aleutian T-type, WCT – West Coast T-
416 type, Ic – Iceland, Nw – Norway.

417

	EAT	WCT	AR	Kam	Co	NR	SR	Ic
WCT	0.0005							
AR	0.0038	0.0039						
Kam	0.0037	0.0038	0.0002					
Co	0.0037	0.0038	0.0002	0.0001				
NR	0.0038	0.0039	0.0001	0.0002	0.0002			
SR	0.0038	0.0038	0.0002	0.0001	0.0001	0.0002		
Ic	0.0041	0.0042	0.0013	0.0012	0.0012	0.0013	0.0012	
Nw	0.0041	0.0042	0.0013	0.0012	0.0012	0.0013	0.0012	0.0002

418

419　Table 2. Distance matrix of population repertoires averaged across all observers. AR – Alaskan
420　R-type, NR – Northern R-type, SR – Southern R-type, Kam – Kamchatkan R-type, Co –
421　Commander Islands R-type, EAT – Eastern Aleutian T-type, WCT – West Coast T-type, Ic –
422　Iceland, Nw – Norway.

	EAT	WCT	AR	Kam	Co	NR	SR	Ic
WCT	0.800							
AR	0.844	0.844						
Kam	0.871	0.906	0.788					
Co	0.893	0.914	0.795	0.809				
NR	0.904	0.890	0.781	0.830	0.848			
SR	0.893	0.882	0.834	0.856	0.888	0.837		
Ic	0.871	0.871	0.785	0.826	0.824	0.807	0.844	
Nw	0.884	0.893	0.834	0.879	0.887	0.858	0.869	0.854

423

424

425   Figure captions

426   Fig. 1. Map showing the locations of each sampled population and subpopulation: AR – Alaskan
427   R-type, NR – Northern R-type, SR – Southern R-type, Kam – Kamchatkan R-type, Co –
428   Commander Islands R-type, EAT – Eastern Aleutian T-type, WCT – West Coast T-type, Ic –
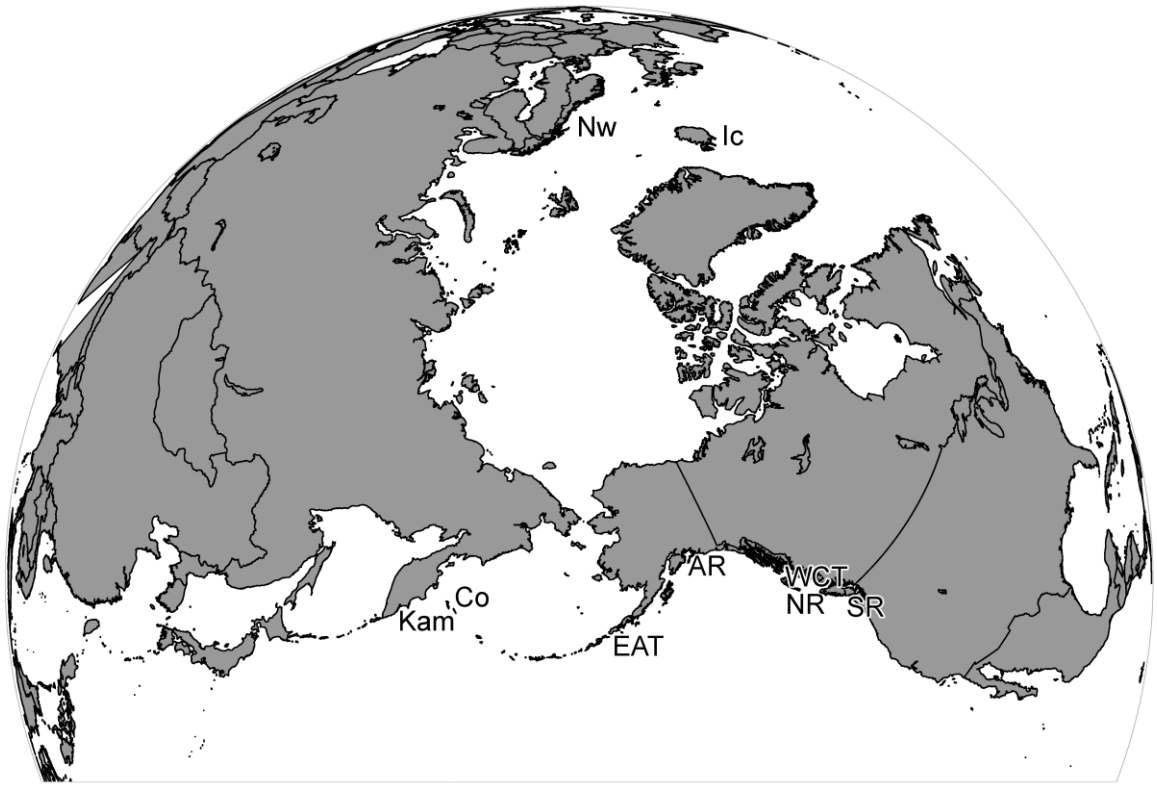429   Iceland, Nw – Norway.

430   Fig. 2. Examples of calls from different populations categorized by an observer into three
431   categories. Population affiliation is indicated above each call: AR – Alaskan R-type, NR –
432   Northern R-type, SR – Southern R-type, Kam – Kamchatkan R-type, Co – Commander Islands
433   R-type, EAT – Eastern Aleutian T-type, WCT – West Coast T-type, Ic – Iceland, Nw – Norway.

434   Fig. 3. Bayesian phylogenetic tree of complete mitochondrial genomes of the study populations.
435   AR – Alaskan R-type, NR – Northern R-type, SR – Southern R-type, Kam – Kamchatkan R-
436   type, Co – Commander Islands R-type, EAT – Eastern Aleutian T-type, WCT – West Coast T-
437   type, Ic – Iceland, Nw – Norway.

438   Fig. 4. Average linkage dendrograms of population/subpopulation similarity based on the
439   number of shared call categories classified by each observer, and dendrogram averaged by all
440   eight observers. AR – Alaskan R-type, NR – Northern R-type, SR – Southern R-type, Kam –
441   Kamchatkan R-type, Co – Commander Islands R-type, EAT – Eastern Aleutian T-type, WCT –
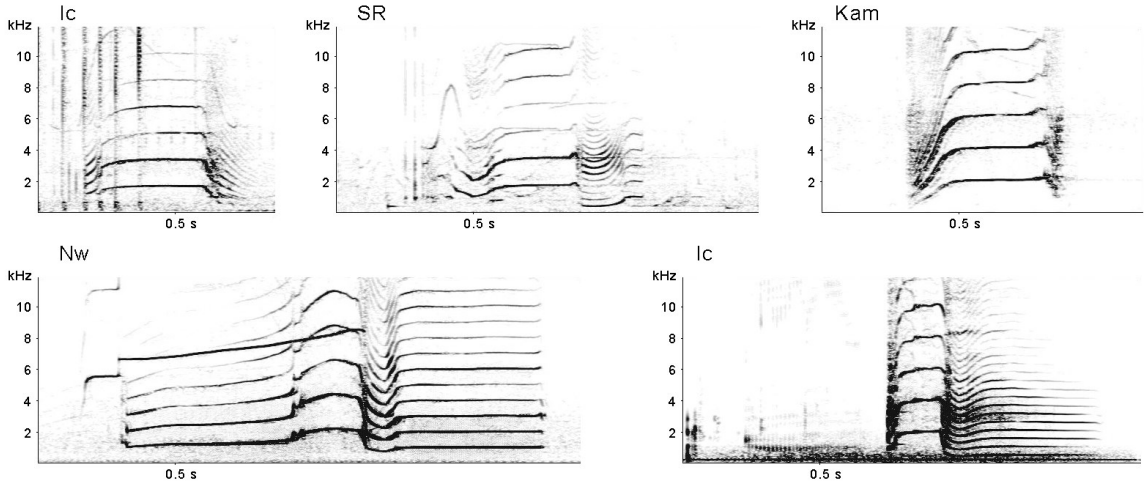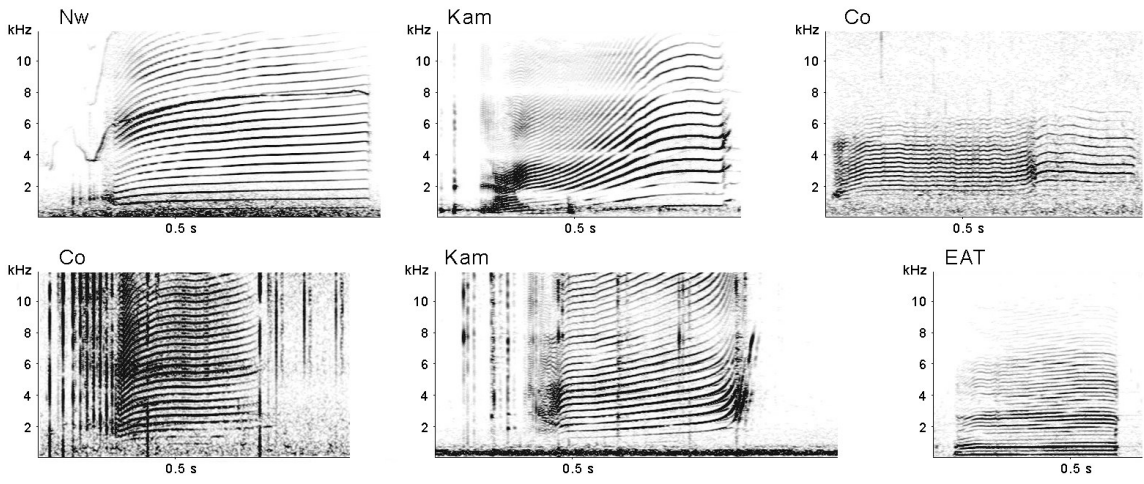442   West Coast T-type, Ic – Iceland, Nw – Norway.
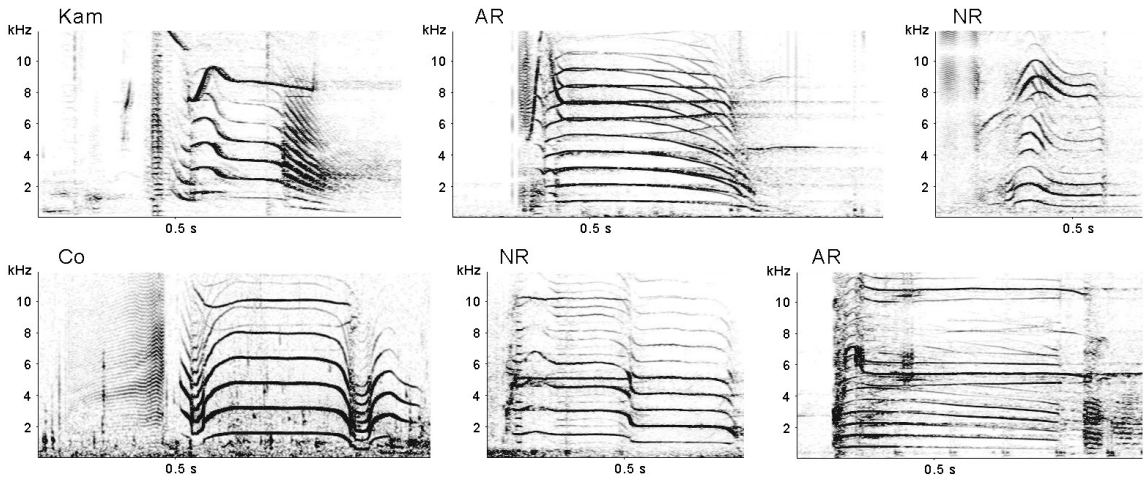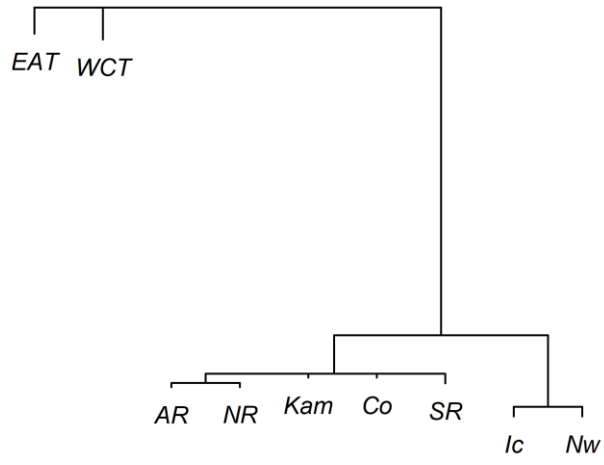

443

444

445

446

CATEGORY 1

Ic
SR
Kam
Nw
Ic

CATEGORY 2

Nw
Kam
Co
Co
Kam
EAT

CATEGORY 3

Kam
AR
NR
Co
NR
AR

447

448

449

450

451