THE COMPANY OF **Biologists**

RESEARCH ARTICLE

Evidence for vocal learning in juvenile male killer whales, *Orcinus orca*, from an adventitious cross-socializing experiment

Jessica L. Crance[1,*], Ann E. Bowles[1,‡] and Alan Garver[2]

ABSTRACT

Killer whales (*Orcinus orca*) are thought to learn their vocal dialect. Dispersal in the species is rare, but effects of shifts in social association on the dialect can be studied under controlled conditions. Individual call repertoires and social association were measured in three adult female killer whales and three males (two juveniles and an adult) during two periods, 2001–2003 and 2005–2006. Three distinct dialect repertoires were represented among the subjects. An adventitious experiment in social change resulted from the birth of a calf and the transfer of two non-focal subjects in 2004. Across the two periods, 1691 calls were collected, categorized and attributed to individuals. Repertoire overlap for each subject dyad was compared with an index of association. During 2005–2006, the two juvenile males increased association with the unrelated adult male. By the end of the period, both had begun producing novel calls and call features characteristic of his repertoire. However, there was little or no reciprocal change and the adult females did not acquire his calls. Repertoire overlap and association were significantly correlated in the first period. In the second, median association time and repertoire similarity increased, but the relationship was only marginally significant. The results provided evidence that juvenile male killer whales are capable of learning new call types, possibly stimulated by a change in social association. The pattern of learning was consistent with a selective convergence of male repertoires.

KEY WORDS: Vocal repertoire, Killer whale, Behavior, Vocal dialect, Social association

INTRODUCTION

Killer whales, *Orcinus orca* (Linnaeus 1758), live in long-term, stable matrilines (Bigg et al., 1990) that produce unique repertoires of discrete, stereotyped calls characterized as dialects (Ford and Fisher, 1982; Ford, 1989; Yurk et al., 2002). Because social dialects appear to be rare in mammals, the selective pressures giving rise to the killer whale dialect are of interest from an evolutionary perspective. They also have management implications. Dispersal among populations is extremely low based on observational and genetic evidence (Bigg et al., 1990; Hoelzel et al., 2007), to the extent that some populations may be classed ultimately as separate species (Morin et al., 2010). Repertoire differences correlate with genetic distance (Yurk et al., 2002; Hoelzel et al., 2007; Deecke et al., 2010). Strategies for management will differ if the dialect is an important driver of genetic isolation as opposed to the reverse.

[1]Hubbs-SeaWorld Research Institute, 2595 Ingraham Street, San Diego, CA 92109, USA. [2]SeaWorld San Diego, 500 SeaWorld Drive, San Diego, CA 92109, USA.
*Present address: National Marine Mammal Laboratory, 7600 Sand Point Way NE, Seattle, WA 98115, USA.

‡Author for correspondence (abowles@hswri.com)

The prevailing model for dialect development and persistence may be described as follows: young killer whales develop their dialect through vertical production learning (Janik and Slater, 1997; Janik and Slater, 2000); and the dialect is maintained by strong motivation to match calls, i.e. to emit calls similar to those of other group members in social exchanges (Deecke et al., 2000; Miller and Bain, 2000; Grebner et al., 2011; Filatova et al., 2011). However, use of calls by individuals is still poorly understood. Call types are thought to change through cumulative drift (Miller et al., 2004), but the source and directionality of the innovation is debated (Filatova et al., 2013). In one case, there was evidence of divergence instead of convergence in acoustic features of matched calls (Grebner et al., 2011).

Three limited lines of evidence support the learning model. Killer whales acquire their discrete call repertoire by gradual matching of the mother's repertoire (Bowles et al., 1988). They are capable of cross-dialect learning based on one observation under controlled conditions from a 10 year old Icelandic female that matched novel calls from a 13 year old Northern Resident female (Bain, 1986). They can match sounds produced by other species based on one observation of an abandoned juvenile killer whale that began producing barks sounding like sea lions (Foote et al., 2006).

There is also evidence for motivation to match calls. Free-ranging killer whales match calls of social partners with greater than random chance (Deecke et al., 2000). Where drift in call features occurs, it seems to happen at the level of matrilines or groups of related matrilines called pods (Miller et al., 2004; Deecke et al., 2010). Playback experiments with groups of killer whales off Kamchatka found that the whales matched calls of pod members but not calls from different pods (Filatova et al., 2011). It is possible the killer whales are also capable of matching calls at the level of regional clans (Weiß et al., 2011).

Experimental observations from other odontocetes are consistent with data from killer whales. Under controlled conditions, a beluga (*Delphinapterus leucas*) produced vocalizations similar to human underwater voice communication, a behavior quantitatively different from the natural repertoire (Ridgway et al., 2012). Several species of social odontocetes altered frequency characteristics of their vocalizations during short-term playback of military sonar signals (DeRuiter et al., 2013), although the matches were not close. Bottlenose dolphins (*Tursiops truncatus*) trained to imitate arbitrary human-made sounds (Richards et al., 1984) showed poor matches after short-term exposure, but bottlenose dolphins acquire close matches of novel, individually distinctive species-typical whistle contours (signature whistles) after changes in social association (Reiss and McCowan, 1993; Janik, 2000; Sayigh et al., 2007; King et al., 2013). They also add features of artificial sounds to their whistles (Miksis et al., 2002). Although exposure time may be a factor in learning, there has been no evidence of an age- or sex-related limitation on learning.

A comparison between call learning in the fluid social system of the bottlenose dolphin (Wells, 2003) and stable social groupings of

List of symbols and abbreviations

AI	association index (Eqn 3)
LFPs	low-frequency pulses
RU	'resting under' state
SEL	sound exposure level
SS	synchronous swimming state
VAR	variable pulsed calls
WSI	Whittaker similarity index (Eqns 1, 2)

killer whales could provide an important test of the hypothesis that learning is related to social association. If vocal learning is an adaptation that promotes social relationships in changing bottlenose dolphin associations, then killer whales in stable societies might experience less pressure to learn and more pressure to conform. The evidence for bottlenose dolphin vocal learning is strong thanks to experimental tests on individuals under free-ranging, semi-controlled and controlled conditions. Similar experimental tests are needed in killer whales, but only one experimental study has exposed groups to short-term playback of calls under free-ranging conditions (Filatova et al., 2011). However, individuals can be exposed to different dialects in adventitious cross-fostering and cross-socializing experiments when killer whales from different vocal traditions are housed together in oceanaria (Bain, 1986). We used this experimental setting to collect data on repertoires of juvenile to adult killer whales representing three vocal traditions over a 6 year period, looking for evidence that varying social association can result in learning of new calls, usage patterns or call features.

RESULTS

A total of 1691 vocalizations were attributed with high reliability to the six subject whales (Table 1) resident at Facility 1, 52% in 2001–2003 and 48% in 2005–2006. The adult females (F1, F3 and F4) produced 6–13 call types during 2001–2003, and the males (M4, M6 and M8) produced from 10 to 15 (Table 1). No individual whale produced all the stereotyped call types in the ethogram for the three dialects, nor did whales sharing call types use them with equal frequency (supplementary material Tables S1–S3). In every subject, a subset of vocalizations predominated in the sample, usually stereotyped pulsed calls (call types). However, categorized vocalizations [variable pulsed calls (VAR) low-frequency pulse trains (LFPs) and whistles] could be used frequently as well.

Repertoire similarity, 2001–2003

Whale F3's predominant calls were two variants of the SD1-1 call (supplementary material Fig. S1), used by all four whales with the Icelandic-origin dialect. Animal Care staff characterized her as the dominant whale in the facility, and F4 as her immediate subordinate

(the other whales had more complex relationships). F3 also commonly produced the Icelandic-origin resting call SD-RST (43%). F3 and F4 were less likely than the other whales to produce categorized vocalizations (4% overall).

F4 shared F3's predominant calls during the first period, with the Icelandic resting call making up 56% of her sample. Only F3 and F4 used resting calls commonly.

F1 produced six call types originating from the Northern Resident A5-pod repertoire (Table 1; supplementary material Fig. S2). Three calls made up 71% of her sample, including one A5-pod resting call (N3). F1 did not produce any of the stereotyped calls of Icelandic origin, nor did F3 or F4 produce any of F1's calls during either period. All the overlap in F1's repertoire was the result of using categorized vocalizations.

Based on previous experience (Bowles et al., 1988), M4's repertoire was expected to be closest to that of his dam, F3. By the summer of 2002, M4 had acquired 10 stereotyped calls in the Icelandic-origin dialect (supplementary material Table S1) and shared seven of his dam's calls, but the Whittaker similarity index (WSI) indicated substantially lower overlap (0.36) because his predominant calls were less common in her repertoire and her three predominant calls comprised only 16% of his sample. His repertoire was more similar to that of the other juvenile male, M6 (WSI 0.62), even though his dam was his closest associate in 2001–2003.

M6 produced 13 call types (supplementary material Table S1) and had the most broadly shared repertoire of any subject. However, he did not produce stereotyped calls of F1 or M8.

M8 had the largest repertoire in the first period (Table 1; supplementary material Fig. S3) and produced the most calls (N=387). Tonal ('scream') components in most of his calls were warbled. Killer whales warble the trailing end of their calls in highly active social states (Ford, 1984; Ford, 1987), but M8's calls were consistently and stereotypically warbled regardless of behavioral state. The origins of this behavior were not known, but he had an unusual background, having been held with bottlenose dolphins prior to arriving at Facility 1. M8 also produced trains of clicks that overlapped stereotyped call components and repeated or recombined components (supplementary material Table S3).

M8's sample was skewed by a sequence of 225 calls produced on a single day when he was in the test pool alone. Most of the novel repetitions and recombinations were observed during this bout. If calls of a given type with repetitions and combinations were pooled, his repertoire would have included just five call types in 2001–2003. He also used categorized vocalizations commonly (mainly LFPs and whistles), amounting to 21% of his sample.

The hierarchical cluster analysis (Fig. 1A) showed that repertoires of all subjects were distinguishable. They were clustered along

Table 1. Counts of vocalization samples and call types attributed to the six study subjects in two periods

Subject	Vocalization samples		Call types (categorized vocalizations)		
	2001–2003	2005–2006	2001–2003	2005–2006	Cumulative
F3	86 (10%)	25 (3%)	13 (2)	4	14 (2)
F4	91 (10%)	106 (13%)	9 (1)	11	14 (1)
F1	72 (8%)	163 (20%)	6 (1)	6 (2)	6 (2)
M4	58 (7%)	258 (32%)	10	25 (2)	26 (2)
M6	182 (21%)	176 (21%)	13 (2)	16 (1)	19 (2)
M8	387 (44%)	87 (11%)	15 (3)	11 (2)	24 (3)

Vocalization samples within the two periods are shown as the number and the percentage (in parentheses) of the total sample produced by each whale. The number of stereotyped call types within the two periods and the cumulative data are shown with the count of categorized vocalization types (VAR, LFPs, whistles) produced by each subject in parentheses.
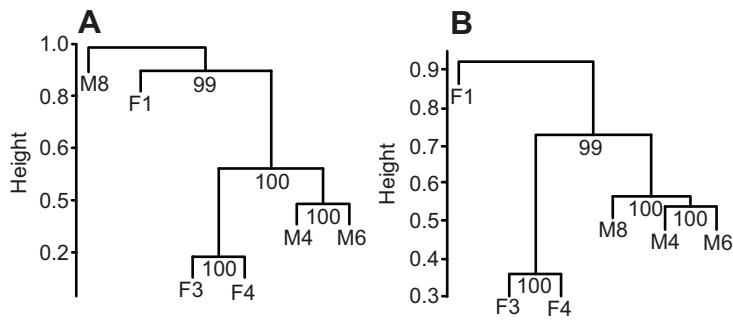
Fig. 1. Dendrograms illustrating the results of the hierarchical cluster analysis of repertoire similarity. Data for the six killer whales (F1, F3, F4, M4, M6 and M8) were used from the period 2001–2003 (A) and 2005–2006 (B). Percentage of bootstrapped simulations exceeding chance expectation is indicated by the number under each split.

dialect boundaries, with the whales having Icelandic-origin repertoires separated from F1 and M8, who were also separated from one another.

Repertoire similarity, 2005–2006

During 2005–2006, F3's attributed sample was small (25 calls) and limited to four call types (Table 1). The small sample and reduction in count of call types was the product of challenges in attribution posed by the presence of her closely affiliated calf. In the first period, her sample was collected after M4 had begun swimming on his own and was attributed using multiple methods. In 2005–2006, her calls were attributed largely by bubbling or nodding. The most common call types in her 2005–2006 repertoire were disproportionately attributed with these behaviors (SD1-1 and SD-RST, respectively).

F4 produced a similar count of Icelandic-origin calls in both periods (9 and 11 call types, respectively). SD-RST comprised 40% of her sample, which explained the high WSI for the F3–F4 dyad (Fig. 2). She produced a somewhat different subset of other Icelandic-origin call types in 2005–2006 (supplementary material Table S1). However, she did not produce calls from the repertoires of either F1 or M8.

F1's repertoire was unchanged in the second period (supplementary material Table S2) and dominated by the same three call types. Although her association index (AI) with M4 and F3 increased, she did not produce any calls of Icelandic origin. Categorized vocalizations (VAR and LFP) accounted for 10% of her sample.
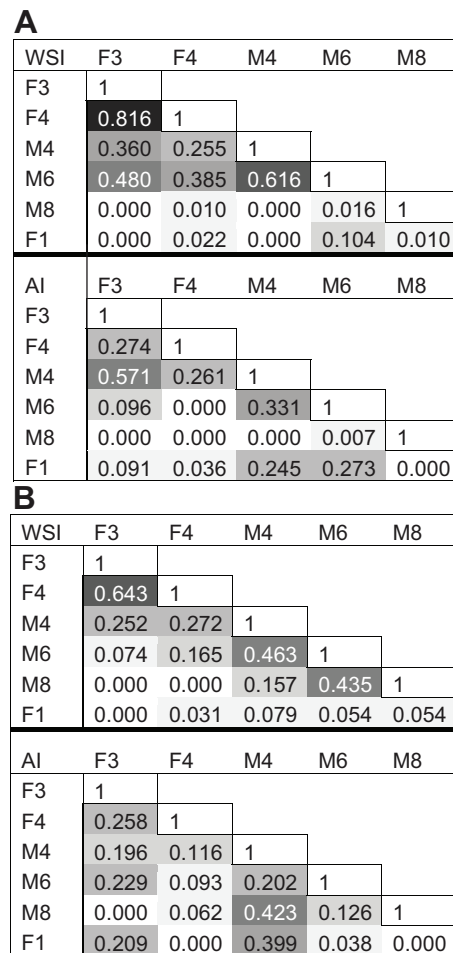
While the repertoires of the female whales remained consistently within dialect in the second period, there were cross-dialect changes in the repertoires of the two young males, M4 and M6. Their repertoires increased substantially, partly through the acquisition of novel call types. M4's repertoire jumped from 10 to 25 call types, giving him the largest repertoire of any whale in the study. There was also a shift in usage; for example, two calls that had comprised 58% of his sample during the first period (SD3-1 and SD4-1) comprised only 17% in the second.

M4's repertoire increased in part because he had learned to produce six of M8's stereotyped calls (supplementary material Table S3), including U1(ABC). None of these had been attributed to any whale other than M8 before this period. M4 also incorporated warbling, not detected during the 2001–2003 period. He produced more LFPs, the third most common sound in M8's repertoire in 2001–2003. Finally, in one instance M4 matched M8's U3(DEBC[R]) call with many repetitions (Fig. 3; supplementary material Audio 1). All of these changes indicated matching of calls and call features from M8's repertoire.

The matches were not precise. M4 warbled the U1(ABC) matched call, but M8's modulation rate in warbled calls averaged 8.5 modulations per second (mod s^{-1}), while M4's was 4.3 mod s^{-1}.

Based on data presented by Ford (Ford, 1984; Ford, 1987), excitement calls in Northern Resident whales had a modulation rate of ~5 mod s^{-1}. Thus, M4's warbling could be considered typical, while M8's was nearly twice as fast. Also, while the time–frequency profile of the matched U3(DEBC[R]) call was recognizable, M4 did not match it closely (Fig. 3A,B). Calls from M8's repertoire accounted for only 3% of M4's sample (nine calls), all produced in one bout in 2006.

M6's sample provided clearer evidence of learning. His repertoire increased from 13 to 16 call types (supplementary material Tables S1, S3) as a result of acquiring M8's calls (Fig. 3C,D). Unlike M4, he produced M8's calls frequently, with

A

WSI	F3	F4	M4	M6	M8
F3	1				
F4	0.816	1			
M4	0.360	0.255	1		
M6	0.480	0.385	0.616	1	
M8	0.000	0.010	0.000	0.016	1
F1	0.000	0.022	0.000	0.104	0.010

AI	F3	F4	M4	M6	M8
F3	1				
F4	0.274	1			
M4	0.571	0.261	1		
M6	0.096	0.000	0.331	1	
M8	0.000	0.000	0.000	0.007	1
F1	0.091	0.036	0.245	0.273	0.000

B

WSI	F3	F4	M4	M6	M8
F3	1				
F4	0.643	1			
M4	0.252	0.272	1		
M6	0.074	0.165	0.463	1	
M8	0.000	0.000	0.157	0.435	1
F1	0.000	0.031	0.079	0.054	0.054

AI	F3	F4	M4	M6	M8
F3	1				
F4	0.258	1			
M4	0.196	0.116	1		
M6	0.229	0.093	0.202	1	
M8	0.000	0.062	0.423	0.126	1
F1	0.209	0.000	0.399	0.038	0.000

Fig. 2. Relationship between repertoire similarity and association among subjects. The boxes are heat plots of repertoire similarity (Whittaker similarity index, WSI) and association index (AI) across subject dyads in the two periods: (A) 2001–2003 and (B) 2005–2006.
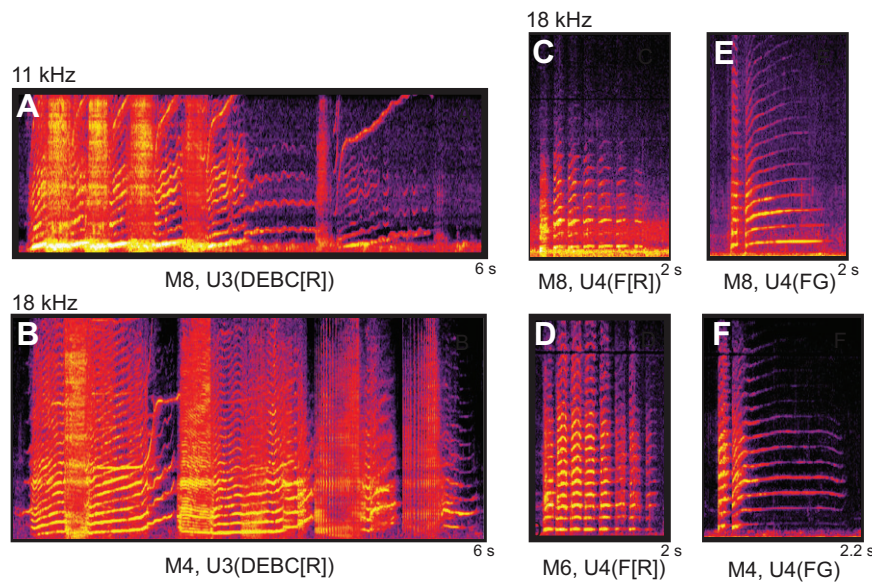
Fig. 3. Spectrograms of three call types produced by M8, with corresponding matched calls by M4 and M6. U3(DEBC[R]) by M8 (A) and M4 (B). U4(F[R]) calls by M8 (C) and M6 (D). U4(FG) calls by M8 (E) and M4 (F).

three comprising 43% of his sample (supplementary material Table S3, Audio 2). U4(FG) was his most common call in this period (28%).

However, M8 provided only limited evidence for a reciprocal change in repertoire and call features (supplementary material Table S3). He produced many fewer recombined and repeated variants, a calling pattern more typical of the other whales, but he only produced one Icelandic-origin call type, SD5-2 (Fig. 4A). This had been attributed to M4 and M6 frequently in the first period. It provided the only evidence of reciprocal call matching across dialects among the adult whales. It was of Icelandic origin and had been used by whales in Facility 1 for many years prior to the start of the study. M8 produced it rarely (four calls, <5%), with all instances recorded when he was alone. His version incorporated warbling (Fig. 4A; supplementary material Audio 3) even though the young males did not warble SD5-2 calls (Fig. 4C).

The hierarchical cluster analysis (Fig. 1B) showed a shift in clustering consistent with these observations. All repertoires were still significantly differentiated, but M8 was included in the cluster with the juvenile males and separate from another cluster with F3 and F4.

Association and repertoire, 2001–2003

Time spent in association varied greatly among the dyads (Fig. 2), ranging from the dam–calf (F3–M4) dyad, which had the highest AI of any pair (57%), to an AI of zero. As a guideline, dyads with an AI of 20% or more were considered social associates (Fig. 2).

The relationship between WSI and AI fitted study predictions broadly, but there were also exceptions. F3 was associated with F4 (AI 0.27), and F4 associated with M4 (0.26) largely as a result of frequent bouts of trio swimming with F3 (95% of joint F4–M4 observations). The F3–F4 dyad had the highest WSI during this period (0.82), despite their moderate AI. The M4–M6 dyad had an
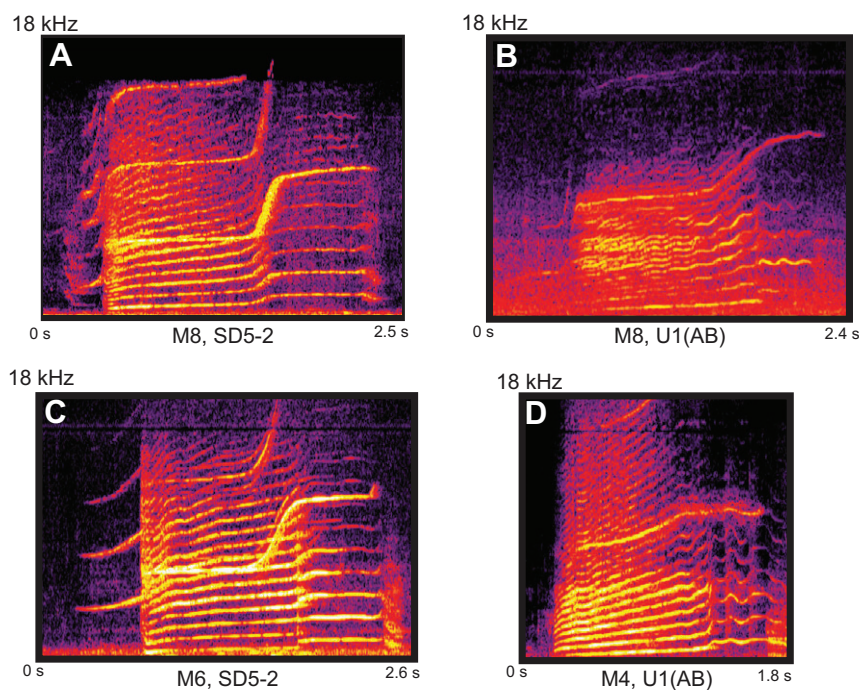
Fig. 4. Calls of M8 and matching calls by the two juveniles males. (A) SD5-2 and (B) U1(AB) calls produced by M8. Both calls show rapid modulation. Versions of the same calls by M6 (C) and M4 (D) are shown.

AI of 0.33, also moderate, but the second highest WSI during the period (0.62).

AI and WSI values were more consistent with expectation in dyads that crossed dialect boundaries. F3 and F4 had low indices of association with F1 and low WSI values (<0.10). M8 had no strong associations during this period and overlapped repertoires only based on the use of categorized vocalizations. AI values showed association between F1 and the two juvenile males, M4 and M6 (AI 0.24 and 0.27, respectively), but she shared no call types with them.

The exceptional cases were consistent with sex-linked convergence. M4 had a higher WSI with M6 than with his dam, with whom he was most closely associated. F3 and F4 had moderate AI values, but the highest WSI in both periods. Even though they did not share call types in 2001–2003, M6 and M8 shared use of categorized calls, accounting for over 86% of these calls. In 2005–2006, the two converged in use of both stereotyped and categorized calls.

Association due to the resting under (RU) state was limited to the two young whales and their associates. M4's RU time was spent with his dam (1.61 h, 98%) and F1 (0.034 h, 2%). The bout with F1 was observed at the end of the 2001–2003 period, as he became increasingly emancipated. M6 spent 0.234 h in RU with F1.

The results of the Mantel test on the 2001–2003 whale×WSI and whale×AI matrices supported a relationship between repertoire overlap and association. The probability of a positive relationship between WSI and AI was significant, but the relationship explained only a moderate proportion of the variance (Mantel r=0.5478, P=0.045) and the resampled range estimate of Mantel r was large (0.2270 to 0.7022).

Association and repertoire, 2005–2006
A change in the study population in 2004 triggered a cascade of changes in association among subjects in 2005–2006. For example, F4 remained a social associate of F3 (AI 0.26), largely as a result of trio-swimming with the new calf, but there was a drop in association between F3 and M4 (from 0.36 to 0.20). M4's association with F1 strengthened (AI 0.40), which also increased the association between F3 and F1 (AI 0.21). F1 spent less time in association with M6 (AI decreased from 0.27 to 0.04), but F3 spent more time with him (AI 0.23 versus 0.09).

Overall, the changes resulted in more uniform AI values across all subjects. Median AI increased from 0.091 to 0.13 and interquartile range decreased from 0.27 to 0.19. However, the three males underwent a large shift in association with one another. The largest change for any dyad during the study was the increase between M4 and M8 (AI 0.00 to 0.42). The increase was from 0.01 to 0.13 for M6.

A shift in the incidence of RU also suggested changed associations. M4 was most often observed resting under F1 (0.349 h, 61.5% of RU), but he was seen in RU with M8 (0.196 h, 38.5%). It was the first time this behavior had been seen with an adult male. M6's association with F3 strengthened (AI 0.10 to 0.23), and he spent 0.112 h in RU with her.

The changes in male association patterns (Fig. 2) paralleled clustering in repertoire similarity (WSI) shown by the dendrogram analysis (Fig. 1B). However, as a result of increased homogeneity in AI, the statistical relationship between WSI and AI was only marginally significant in the second period (Mantel r=0.4277, P=0.0709).

DISCUSSION
The adventitious experiment described here tested two hypotheses: (1) that killer whales can learn calls or call features and (2) that the acquisition or usage of the repertoire is shaped by social affiliation. The study provided evidence that production learning (Janik and Slater, 1997) occurs in killer whales, demonstrating matching across dialect boundaries even when the model repertoire was unusual. However, learning was strongly supported only in the repertoires of the two juvenile male subjects. While social association had a detectable effect in the first period, AI did not explain a high proportion of the repertoire overlap, and was not significantly associated in the second period. Possible interpretations of this pattern of learning are discussed below.

Juvenile repertoire change
Across-dialect change provided stronger evidence of learning than within-dialect change. All four whales sharing calls of Icelandic origin used somewhat different subsets of the dialect and there were shifts in relative call usage within dialect between the two periods, particularly for F4. These shifts could not be explained by the available measure of association (AI). The hierarchical cluster analysis based on the proportional WSI showed a breakdown by dialect, and beyond that by sex, in the first period, but because of differences in usage, all the repertoires of individuals were also significantly different. Within-dialect shifts in usage in 2005–2006 had little effect on clustering.

Some of the differences in repertoire usage could have been explained by sampling. In territorial birds, asymptotic counts characterize repertoires well (Strager, 1995; Dias et al., 2009; Peshek and Blumstein, 2011), but in these species hundreds of individually attributed calls can be collected in short sampling blocks. In contrast, samples of attributed killer whale calls collected in this study were small and spanned longer periods. Although cumulative call repertoires of individuals reached asymptotes, uncommon calls might not have been detected and the method was insensitive to large imbalances in the count of calls in bouts. These would have affected WSI. In addition, social context had an influence on attribution, despite efforts to balance samples by method of attribution as much as possible. F3's decrease in repertoire diversity in 2005–2006 was best explained by a skew in attribution.

However, bias could not explain the evidence for cross-dialect learning by the juvenile males. In 2001–2003, the four whales with Icelandic-origin dialects produced a total of 417 calls, of which none matched calls in the dialects of M8 or F1. In the second period, 44% of M6's sample matched M8's repertoire. This shift was too great to be explained by insufficient sampling of rare calls, attribution bias or within-bout bias. Social association was the most reasonable explanation.

M8 had been held with one or more bottlenose dolphins for over a decade prior to transfer to Facility 1, so it is possible that his unusual repertoire was shaped by extra-species exposure. Unfortunately, no recorded evidence from this earlier period was available. Regardless of the reasons, his calls and calling patterns were sufficiently unlike those of other killer whales that when M6 and M4 acquired them, the behavior could not be explained by genetic relatedness or accidental similarity. This type of evidence has been used to support call learning in other contexts and species (Foote et al., 2006; Ridgway et al., 2012).

It was more difficult to interpret M4's use of M8's calls as evidence of selective learning. Born at the start of the study, he was producing adult pulsed calls by mid-2002 and by the end of the period he was using more of the Icelandic-origin repertoire than his dam. By the end of 2005–2006 he was producing more calls than any other whale. The pattern suggested a steady progression of

learning, possibly with a sex-linked bias, after he became independent of his mother. In 2002 and 2003 (i.e. when he was less than 2 years of age), WSI suggested a greater overlap with M6's repertoire than with his dam's, even though she was his closest associate. In the second period, he began producing calls matching those of M8, although the sample was small, just 9 of 258 calls. All were attributed in 2006 when interacting with M8, and he did not begin producing calls of F1 under the same conditions, even though he associated with her almost as much in 2005–2006 as M8.

The evidence for learning was consistent with other data from young killer whales and other odontocetes (Bain, 1986; Foote et al., 2006; Ridgway et al., 2012; DeRuiter et al., 2013; King et al., 2013). However, the ability to learn appeared to have limitations. Learned calls were not matched precisely by either juvenile. It was not clear why killer whales might produce imperfect matches, but weak matches have been seen in other species when exposure has been limited (Richards et al., 1984; DeRuiter et al., 2013). Additional samples collected over time will be needed to show whether the imprecise matches were a question of practice or limitations on vocal plasticity.

Adult repertoire change

Although significant in the first period, the relationship between repertoire similarity (WSI) and AI explained only half the observed variance at best. Other factors must have influenced repertoire overlap. One possibility was an age and sex bias in learning. The contrast between the clear evidence of learning by the two juvenile males and its absence among the adult females or males associating with the females was suggestive. The only evidence for cross-dialect learning among the adults was an Icelandic-origin call that M8 produced four times in 2005–2006. The call had been present in Facility 1 repertoires since M8 arrived as a young adult. He could have learned it on arrival, or it could have been present in his unknown Icelandic natal repertoire.

Observations in this study were consistent with Bain's (Bain, 1986) report of unidirectional, cross-dialect learning between a subadult and an adult; in that case, the social associates were female.

Association and repertoire

A relationship between repertoire overlap and association has been reported for free-ranging killer whales (Deecke et al., 2010). The results in this study were consistent with such a relationship during 2001–2003, but the relationship weakened after the social environment changed in 2004, and there was high unexplained variance in both periods.

Preferential convergence in repertoire among the males contrasted with a lack of convergence among the females and between the males and the females, even when association increased. This supported motivation in favor of sex-linked repertoire convergence. If males associate to improve prey capture or increase reproductive success, as has been seen in bottlenose dolphins (Smolker and Pepper, 1999; Connor et al., 2000; King et al., 2013), and if these associations are mediated by convergence in call repertoires, repertoire convergence would be favored. In killer whales, gene flow is male mediated (Pilot et al., 2010) and variance in paternity is greater than expected (Ford et al., 2011). Data on free-ranging killer whales should be re-examined for evidence of behavioral and repertoire convergence within age and sex classes; even in highly stable killer whale societies, alliances may form.

The question of vocal plasticity in both sexes is important because a behavioral isolating mechanism that contributed to social or genetic isolation would affect population resilience. Among killer whale populations, failure to recover from local declines (Matkin et al., 2012) contrasts with resilience in odontocetes with more fluid social structures (Ansmann et al., 2012). If limits on learning make it difficult for a reduced population to merge with others, management decisions for local populations must be made based on social relationships as well as population genetics and resources.

MATERIALS AND METHODS
Subjects and experimental conditions

The research protocols were approved by the Hubbs-SeaWorld Research Institute (HSWRI) Institutional Animal Care and Use Committee. Data were collected in two blocks (2001–2003, 2005–2006) from six killer whales at Facility 1: three adult females (F1, F3 and F4) and three males (M4, born to F3 at the start of the study; M6, a juvenile; and M8, an adult). The adults had all been together in the facility for at least 7 years. Three dialects with distinct stereotyped call repertoires were in use from before the start of the study through to its end. F1 used the A-5 pod dialect of the Northern Resident community in the Pacific Northwest. F3, F4, M4 and M6 used a dialect of Icelandic origin. Although originally from Icelandic waters, M8 was not related to them and he had a distinctive repertoire that may have been shaped by early experience with bottlenose dolphins in another oceanarium.

There was a hiatus in observations from summer 2003 to spring 2005, during which F3 gave birth to a second calf and two non-subject whales were transferred to another facility. These two whales had used the Icelandic-origin dialect. The change in group composition provided the experimental manipulation for the study.

The experiment was adventitious rather than planned, but provided two important conditions necessary to study the relationship between association and repertoire change. First, all subjects were exposed to the three distinct repertoires throughout the study. Second, the composition of the group changed halfway through the study, triggering changes in social association. If they began producing novel call types, particularly across dialect, then the hypothesis that killer whales can learn calls would be supported. If subjects with changed associations began producing different call types, then the hypothesis that association drives repertoire change would be supported.

Data collection

The subjects were housed in a 5-pool system, connected by visually and acoustically transparent gates. Recordings were made in a pool with underwater viewing (test pool) from September of 2001 to November of 2006 (dimensions 24.7×41.2×6.1 m). An 8-element array of ITC 6050H and ITC 8212 hull-mounted hydrophones (ITC, Santa Barbara, CA, USA) was used to locate callers acoustically. Multichannel recordings from the hydrophone array were recorded on a laptop computer with customized hardware and software (described below). Video recordings were collected using the multiplexed signal from two underwater and two overhead video cameras recorded onto a Panasonic SVHS or DMR-E95H digital video recorder (Panasonic USA, Newark, NJ, USA). The stereo audio track was recorded from hydrophones at opposite ends of the pool (50 Hz to 16–22 kHz).

Audio tracks were digitized from video using a Sound Blaster A/D card (Creative Labs Americas, Milpitas, CA, USA) at 44.1 kHz (16 bit sampling). Spectrograms and acoustic measurements were made with Raven Pro (v 1.3+ or later, Cornell Laboratory of Ornithology, NY, USA).

The whales vocalized in bouts, separated by unpredictable periods of silence. They were most likely to be vocal during the period from daybreak (when they were first visible) to the start of husbandry activities for the day, so recordings were collected preferentially at this time. Vocalizations were collected until the cumulative count of call types by individual had been constant for at least three calling bouts on separate days (asymptotic call count). Asymptotic counts have proven to be good estimators of repertoire size (Peshek and Blumstein, 2011).

Husbandry considerations, such as keeping dam and calf together, ensured that some dyads were more likely to be found together in the test pool than others. Observations of each dyad were collected until the measure of association, the proportion of available time spent in close proximity, reached an asymptote.

Vocal data extraction and repertoire assessment

Several methods were used to attribute calls to individuals. First, multichannel recordings from the hydrophone array were processed to localize callers using custom-written programs in LabView (National Instruments, Austin, TX, USA) and Matlab (MathWorks, Torrance, CA, USA) developed by BAE Systems in San Diego (C. Greenlaw, D. McGehee).

When an amplitude threshold was exceeded anywhere on the array within a 500 Hz–10 kHz band, where most of the energy in killer whale calls is concentrated, the LabView program collected a 2 s sample at 22 kHz. It is possible that some higher frequency whistles and click trains could have been missed, but the band covered the frequency range of the stereotyped repertoire. Sound exposure levels (SELs) of the events at each hydrophone were calculated. SEL was selected as a metric because complex propagation patterns in the pool had a large effect on peak levels, but affected the total energy in the integrated 2 s samples less.

To protect both equipment and whales, the hydrophones were embedded with the receiving elements 4 cm inside 30 cm cubic niches built into the walls and rockwork and covered by grates. In this configuration, vocalizations had substantially higher SELs when the caller was close to the niche. The custom-written BAE software provided a spectrogram, oscillograms from all hydrophones, and a color-coded map showing the pattern of SEL values across the array.

All calls produced by whales in the pool alone and not in the gate channel could be attributed correctly. If there were two to three whales in the pool, the likely caller had to be within a body length of the hydrophone reporting a single peak in SEL and separated by several body lengths from other whales for an attribution to be reliable. Simultaneous multiplexed video gave the identity of the caller. Approximately 40% of calls emitted with two to three whales in the pool could be attributed with high reliability.

Only calls attributed with high reliability were included in the study dataset. As a result, array-based localizations were biased against attributions when whales were in the pool gateway or swimming close to other whales. Other attribution methods were used to balance these biases. When whales were at the surface and at close range (e.g. in gateways and channels), observers could identify callers by ear. Calls could also be attributed behaviorally using bubble streams ('bubbling') and nodding. Behavioral attributions were potentially biased with respect to call type (Fripp, 2005), but were reliable when emitted by adult whales and not overlapping other calls. Because bubbling and nodding provided attributions in the most challenging context, when whales were swimming close together, they were included. None of the available methods of attribution were completely unbiased, so attribution methods were balanced as much as possible within subject.

In addition to discrete, stereotyped pulsed calls, killer whales produced unstereotyped categories of vocalizations: variable pulsed calls (VAR), clicks and low-frequency pulses (LFPs), and whistles (Ford, 1989; Ford, 1991; Rehn et al., 2007; Riesch et al., 2008). Variable calls (supplementary material Fig. S1) were characterized by unusual modulations, long strings of components and intermixed whistles or pulse bursts. The whales produced high frequency click trains infrequently, but LFPs (supplementary material Fig. S1) and whistles were more common. Whistles were distinguished by high fundamental frequency, rough timbre and an irregular time–frequency profile (supplementary material Fig. S1).

Discrete stereotyped calls were composed of one or more pulsed components (Ford, 1987; Watkins et al., 1998) separated by transitions lasting a few milliseconds and differing in time–frequency profile, pulse structure and timbre. Call types were incorporated into the study ethogram after they had been cataloged at least five times, either within the study period or before it (supplementary material Tables S1–S3). Author J.L.C. extracted and classified calls by ear and by comparison with spectrograms of calls with the assistance of volunteers. The reliability of call identification was tested with 10 untrained volunteers, yielding an average inter-observer agreement of better than 97% (Crance, 2008). This method has been used successfully in identifying stereotyped cetacean calls previously (Janik, 1999; Sayigh et al., 2007).

Stereotyped calls with rapid modulation in tonal components (warbling) are considered aberrant or excitement calls in Ford's terminology (Ford, 1984). These were pooled with the corresponding unmodified call type. All other pulsed calls were categorized as variable.

Although attribution criteria were chosen for reliability, it was still possible for a call to be attributed to the wrong caller. Until use of a call type could be confirmed at least three times, calls were not accepted into the dataset unless the whale was well isolated or the call attributed by two independent methods. These constraints may have reduced the count of uncommon calls in repertoires but they were applied systematically to all subjects and were preferable to inflating repertoires with uncommon calls.

Discrete, stereotyped calls were given an accession number and a code indicating variants when they were entered into the ethogram. For example, SD1-1a indicated the first variant of the first call identified at Facility 1, with a short chirp indicated by 'a' at the end (supplementary material Fig. S1). Where they matched call types classified in previous studies, the published designation was used, e.g. the N3 call produced by Northern Residents (supplementary material Fig. S2).

M8 produced a large number of stereotyped pulsed calls (supplementary material Table S3) that met criteria for inclusion in the ethogram but were unlike those of the other killer whales. They were characterized by warbling, overlapping pulse trains and multiple repetitions or combinations of components. For example, call U1(ABC) had three components (A, B and C), any of which could be elided or repeated (encoded with '[R]' in the ethogram).

Repertoire similarity index

Previous studies of birds and killer whales have used a Dice-Sørensen index of similarity to compare repertoire composition (McGregor and Krebs, 1982; Molles and Vehrencamp, 1999; Yurk et al., 2002). This index is based on the presence/absence of call types in the repertoires to be compared. It has been a useful index for songbirds, which produce all the elements in their repertoire in the course of a few hours of sampling, and in recordings collected from killer whale matrilines with multiple callers. It may not produce good results for samples from individual odontocetes. Odontocetes call in bouts, in which call types tend to be clustered, and there are imbalances in the rate of calling across social contexts (Fripp, 2005).

Data similar to repertoire counts are collected by community ecologists (Garamszegi et al., 2002). The WSI (Whittaker, 1975) has been a good predictor of similarity in community ecology because it takes into account relative occurrence and sample size (Kohn and Riggs, 1982). WSI was calculated for each dyad in both time periods as follows:

$$\text{WSI} = \sum_{i=1}^{N} \min(p_{i,a}, p_{i,b}), \tag{1}$$

where N is the total number of call types, i is the individual call type, a and b are the whales in the dyad being compared, and p is the proportion of the repertoire composed of a particular call. The value of p was calculated by:

$$p_{i,a} = \frac{\text{calls of type } i \text{ produced by whale } a}{\text{total calls by whale } a}. \tag{2}$$

Behavioral association index

The most common measure of association in small odontocetes is synchronous swimming (e.g. Würsig and Würsig, 1979; Gero et al., 2005; Connor et al., 2006), which has also been applied to the killer whale (Ford, 1984; Jacobsen, 1986). We defined synchronous swimming (SS) as time spent swimming in the same direction within one body length. In addition, a second behavior proved to be a strong indication of association, termed 'resting under' (RU). In this state, one whale rested quietly under another with the head in contact with the partner's ventral surface. This behavior was reminiscent of bumping performed by calves to elicit nursing, but occurred outside the dam–calf dyad as well. Morton et al. described the behavior as 'motionless one above the other' (Morton et al., 1986).

The time spent in SS and RU was related to the total time members of a dyad were together. Proportional AI was calculated as follows:

$$\text{AI} = (\text{SS} + \text{RU}) / T, \tag{3}$$

where SS is the total time spent swimming within one body length, RU is the total time spent resting under and T is the total time the dyad was together.

Analysis

Statistical analyses were conducted in the R statistical programming environment (v 2.12.1, R Development Core Team, University of Auckland, New Zealand). Repertoire overlap was assessed using hierarchical cluster analysis in the R package *pvclust* (Suzuki and Shimodaira, 2006), using WSI as the basis for the distance matrix and simple clustering. An autonomous unbiased bootstrapping procedure determined the probability that each split in the resulting dendrogram exceeded the probability expected by chance.

The logical sampling units for the comparison between repertoire overlap and association, dyads, were not independent of one another and neither AI nor WSI could be considered independent predictors. The Mantel test (Sokal and Rohlf, 2012) was used to measure similarity between the two matrices of interest (whale×AI versus whale×WSI). Both AI and WSI were arcsin transformed before analysis. The distribution of the Mantel correlation coefficient was determined with the non-parametric Kendall coefficient of concordance by running 1000 random permutations over the matrices (range from 2.5% to 97.5% of the distribution).

References

Ansmann, I. C., Parra, G. J., Chilvers, B. L. and Lanyon, J. M. (2012). Dolphins restructure social system after reduction of commercial fisheries. *Anim. Behav.* **84**, 575-581.

Bain, D. E. (1986). Acoustic behavior of *Orcinus*: sequences, periodicity, behavioral correlates, and an automated technique for call classification. In *Behavioral Biology of Killer Whales* (ed. B. C. Kirkevold and J. S. Lockard), pp. 335-371. New York, NY: Alan R. Liss, Inc.

Bigg, M. A., Olesiuk, P. F., Ellis, G. M., Ford, J. K. B. and Balcomb, K. C. (1990). Social organization and genealogy of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington State. *Rep. Int. Whal. Comm. Spec. Issue* **12**, 383-405.

Bowles, A. E., Young, W. G. and Asper, E. D. (1988). Ontogeny of stereotyped calling of a killer whale calf, *Orcinus orca,* during her first year. *Rit Fiskideildar* [Journal of the Marine Research Institute, Iceland] **11**, 251-275.

Connor, R. C., Wells, R., Mann, J. and Read, A. (2000). The bottlenose dolphin: social relationships in a fission-fusion society. In *Cetacean Societies: Field Studies of Whales and Dolphins* (ed. J. Mann, R. Connor, P. Tyack and H. Whitehead), pp. 91-126. Chicago, IL: The University of Chicago Press.

Connor, R. C., Smolker, R. L. and Bejder, L. (2006). Synchrony, social behaviour and alliance affiliation in Indian Ocean bottlenose dolphins, *Tursiops aduncus. Anim. Behav.* **72**, 1371-1378.

Crance, J. L. (2008). Changes in killer whale, *Orcinus orca*, stereotyped call repertoires as a function of time and social association. MSc thesis, University of San Diego, San Diego, CA, USA.

Deecke, V. B., Ford, J. K. B. and Spong, P. (2000). Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Anim. Behav.* **60**, 629-638.

Deecke, V. B., Barrett-Lennard, L. G., Spong, P. and Ford, J. K. B. (2010). The structure of stereotyped calls reflects kinship and social affiliation in resident killer whales (*Orcinus orca*). *Naturwissenschaften* **97**, 513-518.

DeRuiter, S. L., Boyd, I. L., Claridge, D. E., Clark, C. W., Gagnon, C., Southall, B. L. and Tyack, P. L. (2013). Delphinid whistle production and call matching during playback of simulated military sonar. *Mar. Mamm. Sci.* **29**, E46-E59.

Dias, P. A. D., Rangel-Negrin, A., Coyohua-Fuentes, A. and Canales-Espinosa, D. (2009). Behavior accumulation curves: a method to study the completeness of behavioural repertoires. *Anim. Behav.* **77**, 1551-1553.

Filatova, O. A., Fedutin, I. D., Burdin, A. M. and Hoyt, E. (2011). Responses of Kamchatkan fish-eating killer whales to playbacks of conspecific calls. *Mar. Mamm. Sci.* **27**, E26-E42.

Filatova, O. A., Burdin, A. M. and Hoyt, E. (2013). Is killer whale dialect evolution random? *Behav. Processes* **99**, 34-41.

Foote, A. D., Griffin, R. M., Howitt, D., Larsson, L., Miller, P. J. O. and Hoelzel, A. R. (2006). Killer whales are capable of vocal learning. *Biol. Lett.* **2**, 509-512.

Ford, J. K. B. (1984). Call traditions and dialects of killer whales (*Orcinus orca*) in British Columbia. PhD thesis, University of British Columbia, Vancouver, Canada.

Ford, J. K. B. (1987). A catalogue of underwater calls produced by killer whales (*Orcinus orca*) in British Columbia. *Can. Data Rep. Fish. Aquat. Sci.* **633**, 1-170.

Ford, J. K. B. (1989). Acoustic behavior of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. *Can. J. Zool.* **67**, 727-745.

Ford, J. K. B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.* **69**, 1454-1483.

Ford, J. K. B. and Fisher, D. (1982). Killer whale (*Orcinus orca*) dialects as an indicator of stocks in British Columbia. *Rep. Int. Whaling Comm.* **32**, 671-679.

Ford, M. J., Hanson, M. B., Hempelmann, J. A., Ayres, K. L., Emmons, C. K., Schorr, G. S., Baird, R. W., Balcomb, K. C., Wasser, S. K., Parsons, K. M. et al. (2011). Inferred paternity and male reproductive success in a killer whale (*Orcinus orca*) population. *J. Hered.* **102**, 537-553.

Fripp, D. (2005). Bubblestream whistles are not representative of a bottlenose dolphin's vocal repertoire. *Mar. Mamm. Sci.* **21**, 29-44.

Garamszegi, L. Z., Boulinier, T., Møller, A. P., Török, J., Michl, G. and Nichols, J. D. (2002). The estimation of size and change in composition of avian song repertoires. *Anim. Behav.* **63**, 623-630.

Gero, S., Bejder, L., Whitehead, H., Mann, J. and Connor, R. C. (2005). Behaviourally specific preferred associations in bottlenose dolphins, *Tursiops* spp. *Can. J. Zool.* **83**, 1566-1573.

Grebner, D. M., Parks, S. E., Bradley, D. L., Miksis-Olds, J. L., Capone, D. E. and Ford, J. K. (2011). Divergence of a stereotyped call in northern resident killer whales. *J. Acoust. Soc. Am.* **129**, 1067-1072.

Hoelzel, A. R., Hey, J., Dahlheim, M. E., Nicholson, C., Burkanov, V. and Black, N. (2007). Evolution of population structure in a highly social top predator, the killer whale. *Mol. Biol. Evol.* **24**, 1407-1415.

Jacobsen, J. K. (1986). The behavior of *Orcinus orca* in the Johnstone Strait, British Columbia. In *Behavioral Biology of Killer Whales* (ed. B. C. Kirkevold and J. S. Lockard), pp. 303-333. New York, NY: Alan R. Liss, Inc.

Janik, V. M. (1999). Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. *Anim. Behav.* **57**, 133-143.

Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science* **289**, 1355-1357.

Janik, V. M. and Slater, P. J. B. (1997). Vocal learning in mammals. *Adv. Stud. Behav.* **26**, 59-99.

Janik, V. M. and Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1-11.

King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W. and Janik, V. M. (2013). Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proc. Biol. Sci.* **280**, 20130053.

Kohn, A. J. and Riggs, A. C. (1982). Sample size dependence in measures of proportional similarity. *Mar. Ecol. Prog. Ser.* **9**, 147-151.

Matkin, C. O., Durban, J. W., Saulitis, E. L., Andrews, R. D., Straley, J. M., Matkin, D. R. and Ellis, G. M. (2012). Contrasting abundance and residency patterns of two sympatric populations of transient killer whales (*Orcinus orca*) in the northern Gulf of Alaska. *Fish. Bull.* **110**, 143-155.

McGregor, P. K. and Krebs, J. R. (1982). Song types in a population of great tits (*Parus major*): Their distribution, abundance, and acquisition by individuals. *Behaviour* **79**, 126-152.

Miksis, J. L., Tyack, P. L. and Buck, J. R. (2002). Captive dolphins, *Tursiops truncatus*, develop signature whistles that match acoustic features of human-made model sounds. *J. Acoust. Soc. Am.* **112**, 728-739.

Miller, P. J. O. and Bain, D. E. (2000). Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca. Anim. Behav.* **60**, 617-628.

Miller, P. J. O., Shapiro, A. D., Tyack, P. L. and Solow, A. R. (2004). Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca. Anim. Behav.* **67**, 1099-1107.

Molles, L. E. and Vehrencamp, S. L. (1999). Repertoire size, repertoire overlap, and singing modes in the banded wren (*Thryothorus pleurostictus*). *Auk* **116**, 677-689.

Morin, P. A., Archer, F. I., Foote, A. D., Vilstrup, J., Allen, E. E., Wade, P., Durban, J., Parsons, K., Pitman, R., Li, L. et al. (2010). Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Res.* **20**, 908-916.

Morton, A. B., Gale, J. C. and Prince, R. C. (1986). Sound and behavioral correlations in captive *Orcinus orca*. In *Behavioral Biology of Killer Whales* (ed. B. C. Kirkevold and J. S. Lockard), pp. 303-333. New York, NY: Alan R. Liss, Inc.

Peshek, K. R. and Blumstein, D. T. (2011). Can rarefaction be used to estimate song repertoire size in birds? *Curr. Zool.* **57**, 300-306.

Pilot, M., Dahlheim, M. E. and Hoelzel, A. R. (2010). Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *J. Evol. Biol.* **23**, 20-31.

Rehn, N., Teichert, S. and Thomsen, F. (2007). Structural and temporal emission patterns of variable pulsed calls in free-ranging killer whales (*Orcinus orca*). *Behaviour* **144**, 307-329.

Reiss, D. and McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *J. Comp. Psychol.* **107**, 301-312.

Richards, D. G., Wolz, J. P. and Herman, L. M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus*. *J. Comp. Psychol.* **98**, 10-28.

Ridgway, S., Carder, D., Jeffries, M. and Todd, M. (2012). Spontaneous human speech mimicry by a cetacean. *Curr. Biol.* **22**, R860-R861.

Riesch, R., Ford, J. K. B. and Thomsen, F. (2008). Whistle sequences in wild killer whales (*Orcinus orca*). *J. Acoust. Soc. Am.* **124**, 1822-1829.

Sayigh, L. S., Esch, H. C., Wells, R. S. and Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Anim. Behav.* **74**, 1631-1642.

Smolker, R. and Pepper, J. W. (1999). Whistle convergence among allied male bottlenose dolphins (Delphinidae, *Tursiops* sp.). *Ethology* **105**, 595-617.

Sokal, R. R. and Rohlf, F. J. (2012). *Biometry,* 4th edn. New York, NY: WH Freeman.

Strager, H. (1995). Pod-specific call repertoires and compound calls of killer whales, *Orcinus orca* Linnaeus, 1758, in the waters of northern Norway. *Can. J. Zool.* **73**, 1037-1047.

Suzuki, R. and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540-1542.

Watkins, W. A., Daher, M. A., DiMarzio, N. A. and Reppucci, G. (1998). *Distinctions in Sound Patterns of Calls by Killer Whales (Orcinus Orca) From Analysis of Computed Sound Features. Technical Report WHOI-98-05*. Woods Hole, MA: Woods Hole Oceanographic Institution.

Weiß, B. M., Symonds, H., Spong, P. and Ladish, F. (2011). Call sharing across vocal clans of killer whales: Evidence of vocal imitation? *Mar. Mamm. Sci.* **27**, E1-E13.

Wells, R. S. (2003). Dolphin social complexity: lessons from long-term study and life history. In *Animal Social Complexity: Intelligence, Culture and Individualized Societies* (ed. F. B. M. de Waal and P. L. Tyack), pp. 32-56. Cambridge, MA: Harvard University Press.

Whittaker, R. H. (1975). *Communities and Ecosystems*, 2nd edn. New York, NY: Macmillan Publishing, Inc.

Würsig, B. and Würsig, M. (1979). Behavior and ecology of the bottlenose dolphin, *Tursiops truncatus*, in the South Atlantic. *Fish. Bull.* **77**, 399-412.

Yurk, H., Barrett-Lennard, L., Ford, J. K. B. and Matkin, C. O. (2002). Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Anim. Behav.* **63**, 1103-1119.
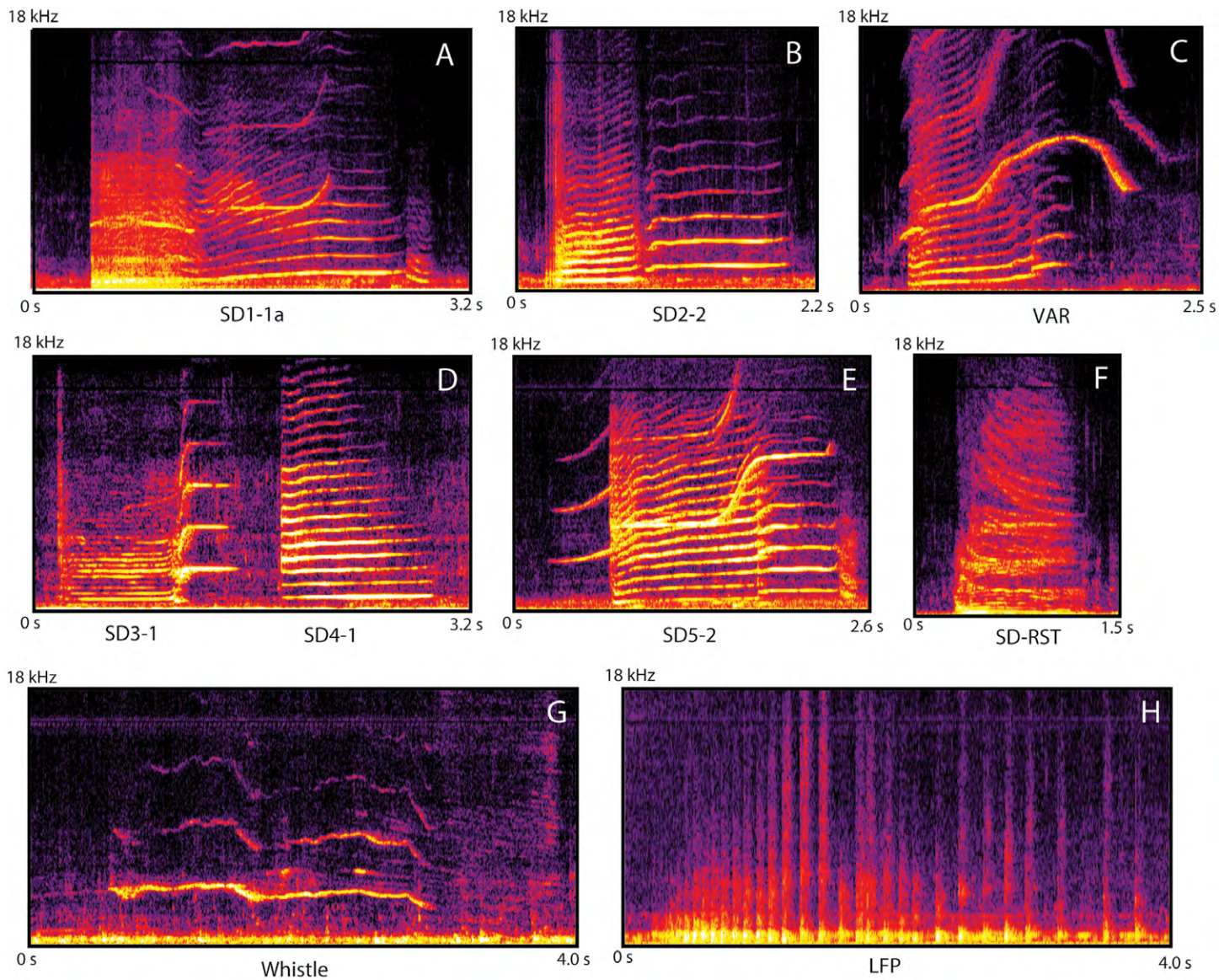
Fig. S1. **Examples of stereotyped calls and categorical vocalizations in the Icelandic-origin dialect.** Those indicated with an "SD" code are discrete, stereotyped calls. The remainder are the unstereotyped, categorized vocalizations, adult variable [VAR], whistle, and low-frequency pop trains [LFP].
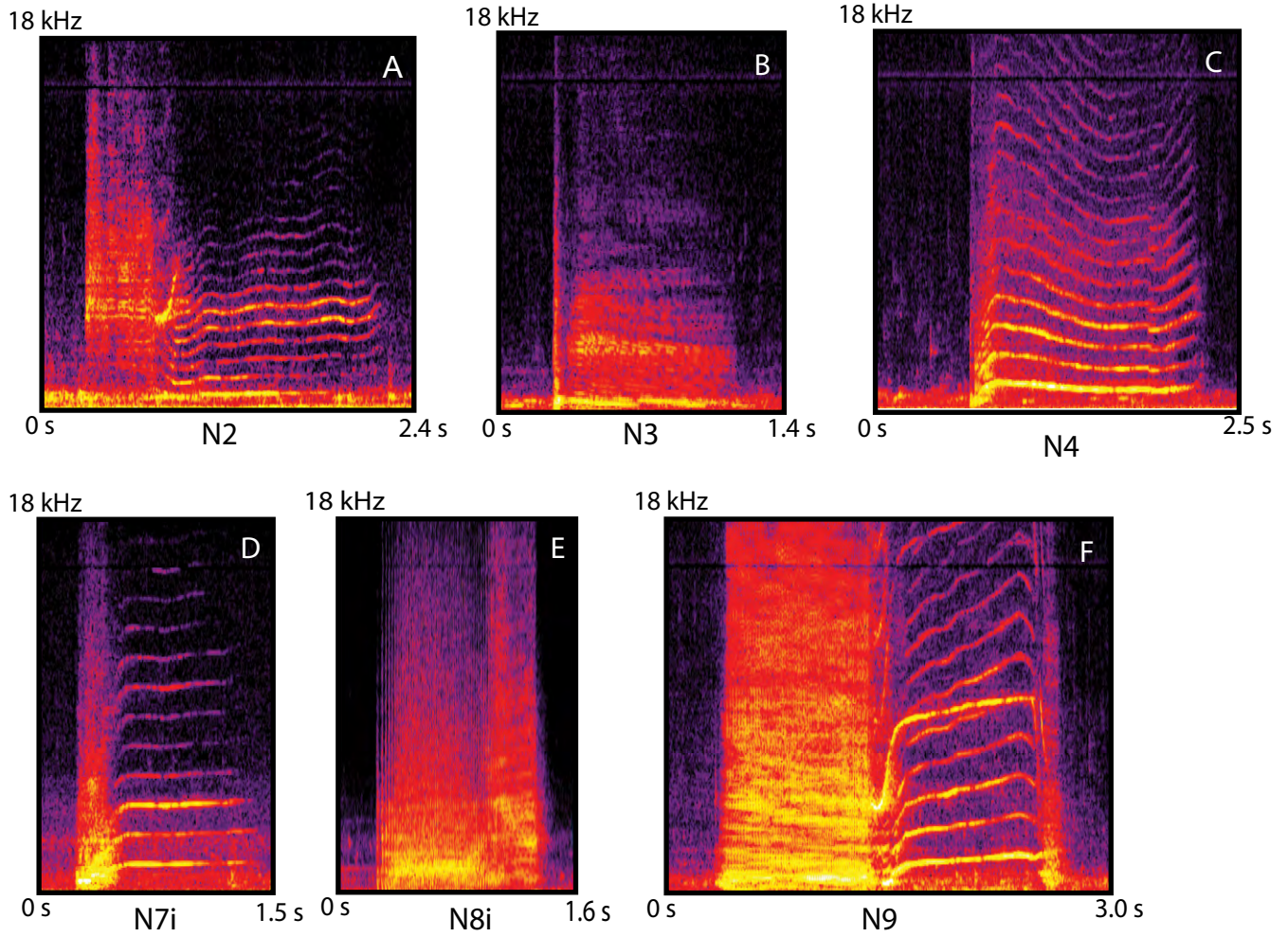
Fig. S2. **Spectrograms of discrete, stereotyped call types produced by F1.** (A) N2, (B) N3 (resting call), (C) N4, (D) N7i, (E) N8i, (F) N9. Call type codes correspond to published designations (Ford, 1984; 1987).
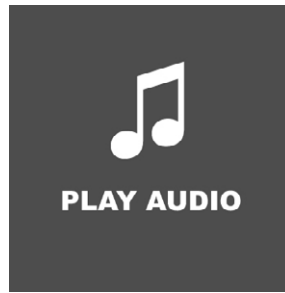
Audio 1. **U3(DEBC[R]) calls produced by M8 and M4**. Spectrograms of the calls in Fig. 3A,B. M8's call is followed immediately by a second call type, U4(FG). The soundfile was manipulated by low-pass filtering at 18 kHz with a 10-pole Butterworth filter in Adobe Audition and smoothing the amplitude envelopes of each clip.
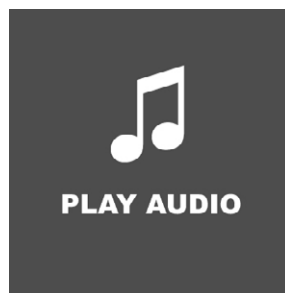
Audio 2. **U4(F[G]) calls produced by M8 and M6**. The sounds correspond to spectrograms in Fig. 3C,D. The soundfile was manipulated by converting to a common sample rate (22 kHz) and smoothing the amplitude envelopes of each clip.
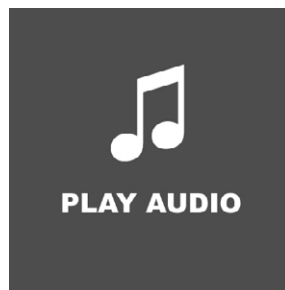
Audio 3. **SD5-2 calls produced by M8 and M6**. The sounds correspond to images in Fig. 4A, C. The soundfile was manipulated by low-pass filtering at 18 kHz with a 10-pole Butterworth filter in Adobe Audition and smoothing the amplitude envelopes of each clip.

Table S1. **Counts of the stereotyped call types in the Icelandic-origin dialect.**
Categorized vocalizations are also given by whale. Period indicated (01-03 for 2001-2003; 05-06 for 2005-2006). Column headings indicate the subject and period. Rank order of the five most common call types or categories indicated by shading (dark = most, light = least). Shading ordered across Table S1-S3. Non-zero cells are bolded.

Call/ Category	F3		F4		M4		M6		M8	
	01-03	05-06	01-03	05-06	01-03	05-06	01-03	05-06	01-03	05-06
Stereotyped pulsed calls										
SD1-1a	12	9	15	8	3	27	33	3	0	0
SD1-1b	10	3	12	2	6	42	18	10	0	0
SD1-2a	1	0	0	0	0	6	0	0	0	0
SD1-2b	0	0	0	1	1	16	0	9	0	0
SD1-3	0	0	0	28	0	2	0	0	0	0
SD1-4	0	0	0	0	0	9	0	0	0	0
SD1-5a	1	0	1	0	0	0	0	0	0	0
SD1-5b	1	0	0	0	0	2	1	2	0	0
SD2-1	2	0	2	2	1	1	4	0	0	0
SD2-2	0	1	0	14	0	1	2	0	0	0
SD3-1	6	0	3	5	18	42	27	13	0	0
SD4-1	6	0	2	1	14	13	30	9	0	0
SD5-1a	0	0	0	0	0	4	0	0	0	0
SD5-1b	1	0	0	0	6	15	4	2	0	0
SD5-2	0	0	0	0	3	17	22	8	0	4
SD6-1	0	0	2	0	0	0	0	1	0	0
SD7-1	0	0	1	0	0	1	0	0	0	0
SD8-1	2	0	0	1	0	2	2	3	0	0
SD8-2	1	0	0	2	0	6	8	7	0	0
SD9-1	0	0	0	0	2	0	7	8	0	0
SD10-1	1	0	0	0	0	4	0	1	0	0
SD-RST	37	12	51	43	1	6	4	0	0	0
*Categorized vocalizations**										
VAR	4	0	2	0	0	25	19	22	4	6
LFP	0	0	0	0	0	8	0	0	56	1
Whistle	1	0	0	0	0	0	1	0	21	0
Total	86	25	91	106	58	249	182	98	81	11[§]

Notes: * VAR –Variable pulsed calls; LFP – low-frequency pulse trains. [§]Counts of calls in Icelandic-origin repertoire and categorized vocalizations. Stereotyped pulsed calls in M8's repertoire are given in Table S3.

Table S2. **Counts of the most common stereotyped calls produced by F1**. Categorized vocalizations are also given. Period indicated (01-03 for 2001-2003; 05-06 for 2005-2006). Rank order of the five most common call types or categories indicated by shading (dark = most, light = least). Shading ordered across Table S1-S3. Non-zero cells are bolded.

Call/Category	F1 (01-03)	(05-06)
Stereotyped pulsed calls		
N9	**5**	**3**
N2	**2**	**5**
N8i	**15**	**27**
N7i	**17**	**41**
N4	**4**	**6**
N3 (Resting Call)	**19**	**64**
Categorized vocalizations*		
VAR	**10**	**7**
LFP	0	**10**
Whistle	0	0
TOTAL	**72**	**163**

Notes: * VAR –Variable pulsed calls; LFP – low-frequency pulse trains.

Table S3. **Counts of the most common stereotyped calls produced by M8, with matched calls produced by M4 and M6 in 2005-2006**. Categorized vocalization counts are given in Table S1; the sums in this table refer only to calls in M8's dialect. Repeated components encoded with an [R], e.g., U4(F[R]). Period indicated by whale (01-03 for 2001-2003; 05-06 for 2005-2006). Rank order of the five most common call types or categories indicated by shading (dark = most, light = least), with shading ordered across Table S1-S3. Non-zero cells are bolded.

Call Type	M8		M4		M6	
	(01-03)	(05-06)	(01-03)	(05-06)	(01-03)	(05-06)
Stereotyped pulsed calls						
U1(AB)	0	4	0	3	0	13
U1(ABC)	73	22	0	1	0	0
U1(ABC[R])	65	5	0	2	0	0
U1(ABC[RM])	19	1	0	0	0	0
U1(AC[R])	4	0	0	0	0	0
U1(C[R])	3	1	0	0	0	0
U2[R]*	55	0	0	0	0	0
U3(DEB)	0	1	0	1	0	0
U3(DEBC)*	6	0	0	0	0	0
U3(DEBC[R])*	19	0	0	0	0	0
U3(DEBC[RM])*	7	0	0	1	0	0
U3(D[R])*	9	0	0	0	0	0
U4(F[R])	5	14	0	0	0	14
U4(FG)	14	17	0	1	0	51
U4(FG[R])	8	2	0	0	0	0
U4(FG[RM])*	14	0	0	0	0	0
U Resting Call	5	9	0	0	0	0
TOTAL	306	76	0	9	0	78

Notes: * Indicates call types detected on a single day, 6/25/02.