Can evidence for spatial and/or temporal genetic heterogeneity of North Pacific minke whales be explained by different mixture fractions of the same two core stocks, or is it necessary to postulate an additional stock (s)?

ROBIN S. WAPLES
Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. East, Seattle WA 98112, USA, robin.waples@noaa.gov.

ABSTRACT
The analyses described here take advantage of the specificity of the two-stock mixing hypothesis for North Pacific common minke whales, which leads to some testable hypotheses based on standard population genetics theory. If the hypothesis is true (heterogeneity can be explained by different mixture fractions of the same two stocks, O and J, in different areas), loci with the largest allele frequency differences between J and O stock should show the largest departures from equilibrium when mixed samples are analyzed. Allele frequency differences were characterized by θ, an analogue of F_{ST}, and departures from equilibrium were indexed by F_{IS} for single loci and by r^2 (a measure of linkage disequilibrium) at pairs of loci. Samples from SA6 and SA9 were used to characterize putatively pure J and O stocks, respectively. Artificial mixtures of equal fractions of J and O individuals showed the expected strong correlations between θ and F_{IS} (or θ_i θ_j and r^2), but these were reduced somewhat when split-sample cross-validation was used, and when unequal mixtures were analyzed. Analysis of putatively mixed samples in general showed weaker correlations than were expected from mixtures of only J and O individuals. This novel type of analysis appears to hold some promise for informing conclusions about stock structure, but more evaluations are needed to determine how robust the results are.

INTRODUCTION

Elucidating stock structure in North Pacific common minke whales is challenging because the breeding grounds have not been identified and samples have to be taken from migrating individuals. Based on morphological and genetic data, it is generally agreed that at least two well-differentiated stocks exist: the so-called "J" and "O" stocks, with the J stock being more coastal and the O stock occurring more offshore. However, it has become increasingly apparent in recent years that the distributions of these two stocks overlap considerably in space and time. Some areas are thought to support mostly or entirely one stock or the other, while other areas show clear evidence for a mixture of stocks.

Scientists differ in their interpretation of the causes of the heterogeneity that is apparent in these mixed areas. In one view (Pastene et al. 2010), the heterogeneity in each mixed area can be explained by different mixing proportions of the same two stocks, J and O. An alternative view (Baker et al. 2010) is that the patterns seen in at least some of these mixed areas are better explained by the presence of one or more additional stocks. At the December 2010 "First Intersessional Workshop for western North Pacific common minke whales", some additional genetic analyses were outlined that might help resolve outstanding question regarding stock structure (see Annex N of the draft report). In this paper, I report results of analyses described in Section A.2. of Annex N of that draft report.

The analyses described here take advantage of the specificity of the two-stock mixing hypothesis, which leads to some testable hypotheses based on standard population genetics theory. It has long been known that samples taken for genetic analysis that include individuals from more than one population will exhibit a Wahlund effect (Wahlund 1928), which is a deficiency of heterozygotes compared to the expected Hardy-Weinberg proportions. This effect is largest when mixture fractions are approximately equal and when allele frequencies differences between populations are large. Therefore, the two-stock mixing hypothesis leads to the testable prediction that the Wahlund effect in the putative mixtures should be strongest at the loci that show the largest allele frequency differences between J and O stocks. An analogue to the Wahlund effect at individual gene loci also occurs for analyses of linkage disequilibrium (LD = non-random associations of alleles at pairs of loci). This provides another testable prediction: all else being equal, LD should be higher for pairs of loci that show strong allele frequency differences between O and J stocks.

To test these predictions, I first used samples of putatively 'pure' J and O stocks to quantify the frequency differences at each locus, and then created artificial mixtures in known proportions to establish the expected relationship between allele frequency differences and the one- and two-locus Wahlund effect. I then computed those

same statistics for samples from mixed areas to see whether results agreed with those obtained from known mixtures of J and O stock.

METHODS

Theory

Consider first a single gene locus in a single population (i), with alleles A and a at frequencies p_i and $q_i = 1-p_i$, respectively. Then, the Hardy-Weinberg expected frequency of heterozygotes (individuals with the Aa genotype) is $E(Aa) = 2p_iq_i$. In a second population (j), frequencies of the same alleles are p_j and q_j. Now consider a mixture of individuals derived from both populations. In that mixed population, the expected frequency of heterozygotes is reduced by the Wahlund effect: $E(Aa) = 2\overline{pq} - 2Var(p)$. The Wahlund effect thus increases with the variance of p among subpopulations. F_{ST} is a measure of allele frequency difference among subpopulations, and F_{IS} is a measure of departures from Hardy-Weinberg genotypic proportions, with positive F_{IS} values indicating a deficiency of heterozygotes. Therefore, according to standard population genetics theory, loci showing large F_{ST} values between populations should be the loci for which F_{IS} values are most strongly positive in population mixtures.
A two-locus analogue to the Wahlund effect was described by Nei and Li (1973) and Sinnock (1975). The premise is that, even when alleles at different gene loci assort randomly within each population, in a population mixture these associations will appear to be non-random when they involve loci that differ in frequency between the populations. Based on work by Nei and Li (1973) and Waples and Smouse (1990), Waples and England (in review) showed that the expected magnitude of mixture disequilibria is a simple function of the mixture fraction and the product of F_{ST} values for the two loci. A common measure of linkage disequilibrium is r^2, the squared correlation of alleles at different gene loci. Therefore, pairs of loci for which the product of single-locus F_{ST} values are large should be the loci for which r^2 values are largest in population mixtures.

Samples

These analyses focused on the microsatellite data for 16 gene loci described by Kanda et al. (2010), which were kindly provided by Institute of Cetacean Research, Tokyo, under Procedure A of the IWC Data Availability Agreement. I identified one geographic area believed to contain pure or nearly pure samples of each stock: SubArea 6 (SA6) for J stock and SA9 for O stock (Table 1). As these areas did include a few individuals possibly belonging to the 'wrong' stock, as well as a number of unassigned individuals, I also considered 'trimmed' versions of the datasets from these areas, which excluded individuals found by Kanda et al. (2010) to have less than a 90% probability of belonging to J stock (SA6) or less than a 90% probability of belonging to O stock (SA9). I also analyzed putative mixtures from five different areas: SA2, SA11, SA7W(bycatach), SA7W-Kushiro, and SA7W-Sanriku (Table 1).

Statistical analyses

The software FSTAT (Goudet 1995) was used to compute F_{IS} and Weir and Cockerham's (1984) θ, which is a widely-used analogue to F_{ST}, and the program LDNe (Waples and Do 2008) was used to calculate r^2 for each pair of gene loci.

Artificial mixtures

θ values were computed for each of the 16 gene loci between reference samples from SA6 and SA9, and locus-specific F_{IS} values were computed for artificial mixtures of individuals from these two areas. The strength of the correlation between the θ and F_{IS} values provided a baseline measure of what can be expected in mixtures of J and O stock in known proportions. For the LD analyses, the 16 loci provided a total of 16*15/2 = 120 different pairwise comparisons of loci. For each locus pair i,j, we computed r^2 and the product $\theta_i \theta_j$. The strength of the correlation between r^2 and $\theta_i\theta_j$. provided a baseline measure of what can be expected for the two-locus Wahlund effect in mixtures of J and O stock in known proportions. We performed these analyses using all individuals from SA6 and SA9, as well as only the trimmed datasets.

Natural mixtures

Samples from five subareas presumably represent mixture of J and O stock individuals, with perhaps some individuals from additional (uncharacterized) stocks as well. We compared the θ values obtained above with F_{IS} and r^2 values for the putative natural mixtures, and compared results with those obtained above for known mixtures.

RESULTS

In the presumably pure samples from both SA6 and SA9, F_{IS} values fluctuated randomly around 0 and none were significant. This result supports conclusions of previous researchers that minke whales collected in these areas have genotypic frequencies that are in approximate Hardy-Weinberg equilibrium. For the full datasets, single-locus θ values ranged from near 0 to just over 0.1, and averaged 0.041—again in agreement with previously reported results. θ values for the trimmed datasets were consistently slightly higher, with an average of 0.047 (Figure 1).

Artificial mixtures

When we used the full SA6 and SA9 datasets to calculate both F_{IS} and θ, we found a very strong correlation between the two measures in an artificial mixture of 410 individuals from each area (Figure 2; r = 0.83). Similarly, we found a very strong correlation between r^2 and $\theta_i\theta_j$ in the analyses of linkage disequilibrium (Table 2; r = 0.84). However, this approach has a potential flaw: the same individuals were used to calculate both θ and F_{IS} or r^2, and this lack of cross-validation can lead to overly optimistic results (in this case, perhaps inflated correlation coefficients). The 'gold standard' for cross-validation is to split the data into two parts—one used to develop an algorithm, the other to test it (Anderson 2010). Accordingly, we redid the artificial mixture analyses using the split-sample method. This allowed two replicates: θ was computed using the first half of the data and F_{IS} (or r^2) the second half, and vice-versa. Results for the θ -F_{IS} analyses are shown in Figure 3: in the first replicate the correlation remained very high (r = 0.84), but in the second replicate it fell to 0.53 (still significant at $P < 0.05$). For the LD analyses, under the split-sample method the correlations fell slightly to 0.72 and 0.73, both still highly significant.

The above analyses involved 1:1 mixtures of putative J and putative O individuals. Table 3 also shows results for artificial 3:1 mixtures of the two stocks. For the single-locus analyses, the 3:1 mixtures produced slightly lower correlations than the equal mixtures (0.73 ($P<0.01$) and 0.47 for 3J:1O, and 0.40 and 0.28 for 3O:1J). For the LD analyses, the correlation coefficients were also considerably reduced compared to the equal-mixture scenario (Table 2; all in the range 0.28-0.44 but all highly significant because of the large number of degrees of freedom).

Finally, I examined the correlation between θ calculated between areas SA6 and SA9 and F_{IS} or r^2 within a pure stock (Table 3). As expected, these correlations were small and non-significant for LD and the single-locus analysis for pure O, but for SA6 ('pure' J) the correlation between θ and F_{IS} was 0.60 (significant at $P < 0.05$).

Natural mixtures

The natural mixtures were chosen because they have relatively large sample sizes and estimated mixture proportions that fall in the range of those evaluated in the artificial mixtures. In theory, assuming the two-stock mixture hypothesis is correct, the SA7Bycatch sample (estimated fraction J = 0.54-0.55 by the two methods) should produce results comparable to those found for the artificial 1:1 mixtures, results for the SA7W-K, SA7W-S, and SA2 samples should be roughly comparable to those for the 1:3 and 3:1 mixtures, and results for the SA11 sample should be intermediate.

Empirical correlations for the presumed natural mixtures were lower than these expectations. For the single-locus analyses, the correlations were all non-significant and ranged from -0.15 to 0.32. The correlation for SA7Bycatch (0.30) was well below the split-sample values for 1:1 artificial mixtures (0.52-0.83), as was the correlation for SA11 (0.24). Correlations for SA7W-K and SA7W-S (0 and -0.15, respectively) were well below those for artificial 1:3 or 3:1 mixtures. The correlation (0.32) for SA11 (estimated to be ~ 80% J) was also below the range for 3J:1O found in the artificial mixtures, but was in the range for 1J:3O mixtures (0.28-0.40). Results changed only slightly when the trimmed datasets were used to calculate locus-specific θ values (Table 3).

Results for the LD analyses were roughly comparable, and again differed only slightly when the trimmed datasets were used (Table 3). The correlation between $\theta_i\theta_j$ and r^2 for the SA7Bycatch sample (0.26) was highly significant but much lower than for the artificial 1:1 mixtures (0.72-0.73, Table 2), while the correlation for SA11 was low and non-significant. However, the correlations for the other three samples (0.22-0.33) were within, or close to, the range found for artificial mixtures with 75% of one stock and 25% of the other (0.28-0.44, Table 2).

Figure 4 gives another perspective for the single-locus analyses for the SA7Bycatch sample. The F_{IS} values provide clear evidence of population mixture, as they are skewed toward high positive values compared to those found in 'pure' J or O samples. In this respect they mimic and even exceed the pattern seen in the artificial 1:1 mixtures.

However, relatively low θ -F_{IS} correlation for the SA7Bycatch sample indicates that the loci with high F_{IS} values in this sample are not consistently those with high θ values between SA6 and SA9.

DISCUSSION

The two-stock mixing hypothesis leads to some specific predictions that are testable in principle based on standard population genetics theory. Analysis of natural samples that are putative mixtures of J and O stock produced correlations between genetic indices that were generally weaker than those found in artificial mixtures of these two stocks in known proportions. This results lends some support to the idea that these natural mixtures do not only contain individuals from the J and O stocks, at least as they are currently characterized. However, this conclusion must be tempered with several caveats.

First, although the analyses used here are grounded in standard population genetics theory, application of this theory to testing hypotheses about mixture proportions is novel, and I am not aware of any published analyses that attempt to do exactly the same thing. Therefore, caution should be exercised in interpreting the results until the behavior of the approach is better characterized. Second, baseline results for the known mixtures in some cases were quite variable, which makes it difficult to determine what the expectation is if the null hypothesis is true. This result is probably due in part to reduced sample size required by the split-sample method for cross-validation. Other cross-validation methods that are not as costly in terms of data loss might be explored in this context. Finally, the empirical correlations for the natural mixtures are only meaningful if the allele frequencies in the pure O and J stocks have been accurately characterized. This is difficult to verify at present, given that individuals cannot be sampled on the breeding grounds. Comparison of the full and trimmed datasets indicates that the differences do not arise from large differences at a few loci but rather to a collection of small differences at nearly all loci (Figure 1). This would seem to argue against the idea that a few genetically divergent individuals are strongly affecting the analyses, which in turn suggests that the current characterization of 'pure' J and 'pure' O stocks might be approximately correct.

ACKNOWLEDGMENTS

We thank Kanda-san and collaborators for sharing unpublished genetic data used in these analyses.

REFERENCES

Goudet J (1995) FSTAT version 1.2: a computer program to calculate *F*-statistics. Journal of Heredity 86, 485–486.

Kanda, N., Park, J.-Y., Goto, M., An, Y.-R., Choi, S.-G., Moon, D.-Y., Kishiro, T., Yoshida, H., Kato, H. and Pastene, L.A. 2010. Genetic analyses of western North Pacific minke whales from Korea and Japan based on microsatellite DNA. Paper SC/62/NPM11 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 13pp.

Nei, M., and W. Li., 1973 Linkage disequilibrium in subdivided populations. Genetics 75: 213-219.

Pastene, L.A., M. Goto, and N. Kanda. 2010. Progress in the development of stock structure hypotheses for western North Pacific common minke whales. Paper SC/62/NPM12 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 9 pp.

Sinnock, P., 1975 The Wahlund effect for the two-locus model. Am. Nat.109:565-570.

Wade, P.R., R.L. Brownell, and T. Kasuya. 2010. A review of the biology of western North Pacific minke whales relevant to stock structure. Paper SC/62/NPM13 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 16 pp.

Wahlund, S., 1928 Zuzammensetzung von populationen und korrelation-serscheiunungen von standpunkt der vererbungslehre aus betrachtet. Hereditas 11: 65-106.

Waples, R. S., and C. Do, 2008 *LdNe*: A program for estimating effective population size from data on linkage disequilibrium. Mol. Ecol. Res. 8:753-756.

Waples, R. S., and P. R. England. Effects of migration on single-sample estimates of contemporary N_e based on linkage disequilibrium. Genetics (in review)

Waples, R. S., and P. E. Smouse, 1990. Gametic disequilibrium analysis as a means of identifying mixtures of salmon populations. Amer. Fish. Soc. Symposium 7:439-458.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38, 1358–70.

Table 1. Samples analyzed in this paper, which are a subset of those shown in Table 10 of Kanda et al. (2010). The number of individuals assigned by Kanda et al. to either J or O stock, as well as the number of unassigned (?) individuals, is shown. Two methods were used here to estimate the proportion of the sample that is J stock. In Method 1, Proportion J = J/(J+O). In Method 2, Proportion J = (J + 0.5*Unassigned)/(J + O + Unassigned). The 'trimmed' subsets for SA6 and SA9 included only those individuals determined (on the basis of STRUCTURE analyses conducted by Kanda et al. 2010) to have at least a 90% probability of belonging to J or O stock, respectively.

					Proportion J stock	
	J	?	O	Total	Method 1	Method 2
'Pure' samples						
Japan SA6	345	63	3	411	0.99	0.92
Trimmed SA6	382	0	0	382	1.00	1.00
Japan SA9W+E	6	88	371	465	0.02	0.11
Trimmed SA9	0	0	418	418	0.00	0.00
Mixtures						
Japan SA2	130	33	20	183	0.87	0.81
Japan SA11	19	18	43	80	0.31	0.35
Japan SA7W-BC	92	45	75	212	0.55	0.54
Japan SA7W-K	34	51	168	253	0.17	0.24
Japan SA7W-S	33	55	139	227	0.19	0.27

Table 2. Correlation coefficients between θ and F_{IS} (top) and between $\theta_i\theta_j$ and r^2 (LD; bottom) in samples of 'pure' J and O (areas SA6 and SA9, respectively), as well as for artificial mixtures in proportions 1:1 and 3:1. In the split-sample analyses, different individuals were used to calculate θ and r^2 or F_{IS}. *P < 0.05; **P < 0.01. Single-locus analyses have 14 df; those for LD have 118 df.

		1J:1O	3J:1O	3O:1J	Pure J	Pure O
F_{IS}						
	All data	0.83**	0.71**	0.56*	0.60*	0.07
	Split sample	0.84**	0.73**	0.40		
		0.52*	0.47	0.28		
LD						
	All data	0.84**	-	-	0.01	0.07
	Split sample	0.72**	0.36**	0.38**		
		0.73**	0.44**	0.28**		

Table 3. As in Table 2, but showing results for analysis of putative natural mixtures from five different sampling areas. Values for 'Fraction J' are from Method 2 in Table 1. In the 'trimmed' analyses, θ values were computed using only the trimmed datasets.

			F_{IS}		LD	
	n	Fraction J	Full	Trimmed	Full	Trimmed
SA7 Bycatch	212	0.54	0.30	0.31	0.26**	0.27**
SA7W-K	253	0.24	0.00	-0.04	0.22*	0.22*
SA7W-S	227	0.27	-0.15	-0.10	0.33**	0.34**
SA2	184	0.80	0.32	0.35	0.26**	0.26**
SA11	80	0.35	0.24	0.23	0.09	0.11

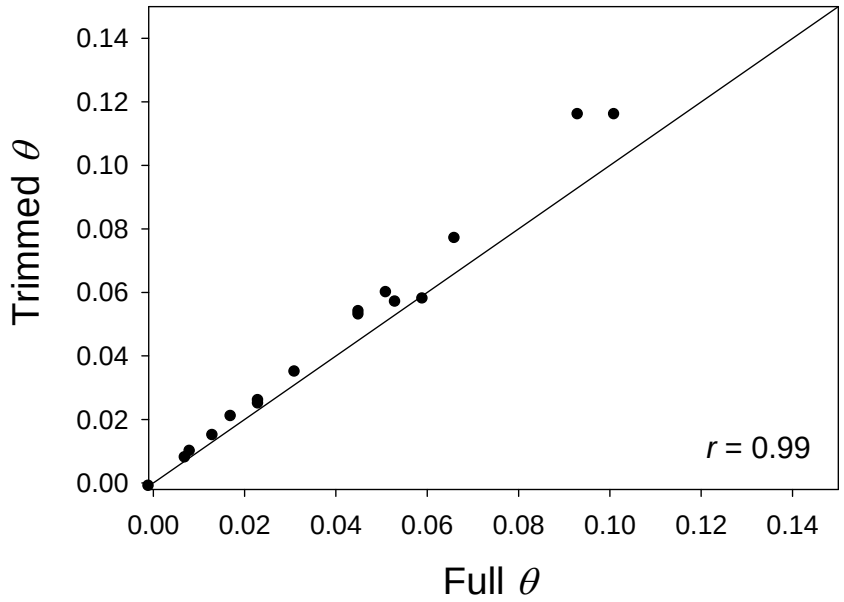

Figure 1. Relationship between Weir and Cockerham's (1984) theta computed for a comparison of areas SA6 and SA9, using the full and trimmed datasets. Each symbol represents one gene locus.

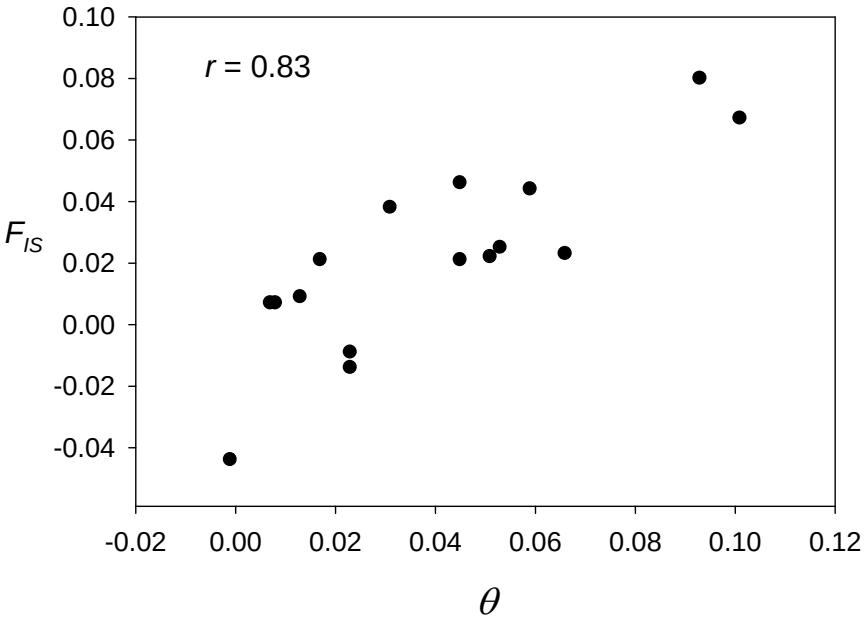

Figure 2. Relationship between locus-specific θ and F_{IS}. θ was measured using 410 individuals each from Area 9 (presumed pure O) and Area 6 bycatch (presumed pure J). F_{IS} was measured in an artificial 50:50 'mixtures' of 410 J and 410 O individuals.

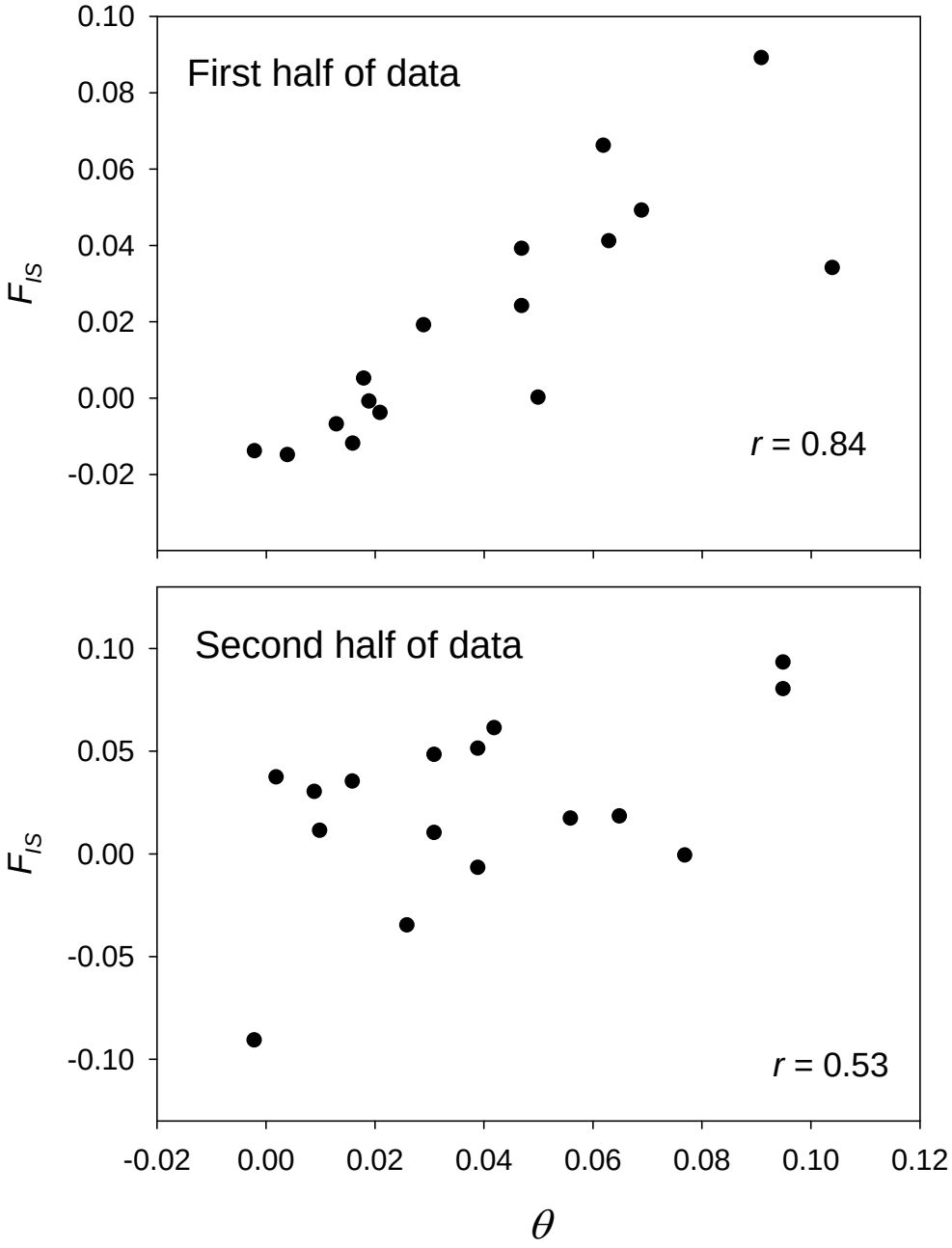

Figure 3. As in Figure 2, except using split-sample cross validation. θ and F_{IS} were calculated on different individuals. In the top panel, the first half of the data (n = 205 each for Areas 6 and 9) were used to compute θ and the other half were used to compute F_{IS}; the reverse procedure was used in the bottom panel.

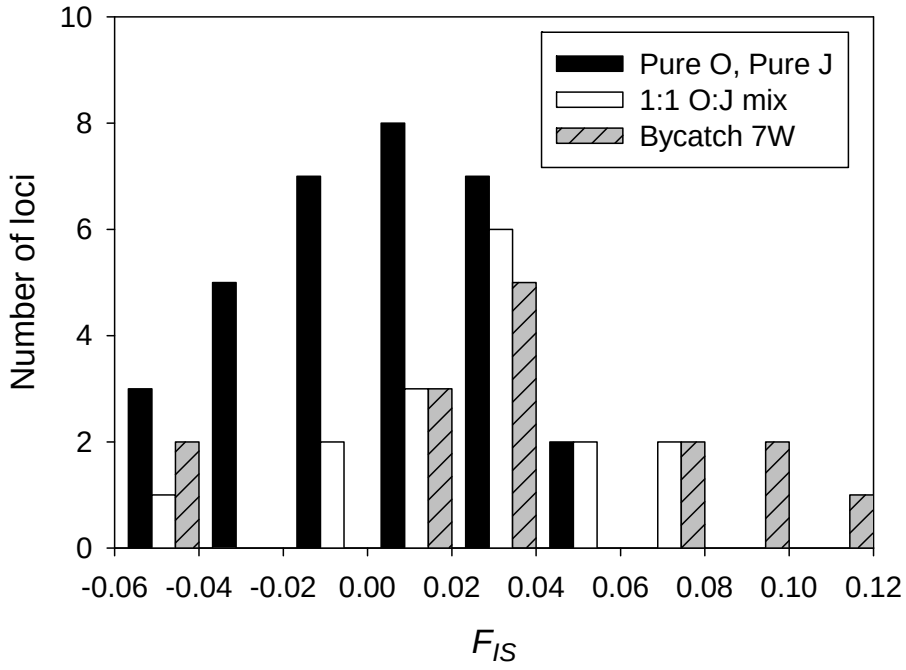

Figure 4. Distribution of single-locus F_{IS} values for presumably pure O and J, an artificial mix of O and J, and Area 7W bycatch.

APPENDIX

ADDENDUM TO SC/63/RMP7

In this Addendum I report results of two types of additional analyses that extend the material described in the main text:

1) Consideration of an alternative way to characterize the Pure O stock;
2) Treating the SA7W-K and SA7W-S samples as a combined sample rather than separately

Consideration of an alternative way to characterize the Pure O stock

The analyses described in the main text used animals collected in SA9 as a proxy for pure O stock. An alternative view is that pure O stock is best characterized by combining samples from SA8, SA7E, and SA7WR. The latter, which is referred to simply as 7W in the new terminology, excludes individuals from the old SA7W that were collected close to the Japanese coastline. The rationale for this view of pure O stock is explained in footnote 1 to Table 6 in the Report of the First RMP Intersessional Workshop for western North Pacific common minke whales.

Treating the SA7W-K and SA7W-S samples as a combined sample rather than separately

The following is a rationale for considering these two samples jointly (P. Wade, personal communication; see SC/63/RMP16 for details): "The coastal sub-areas of 7CN and 7CS (which correspond essentially to the Kushiro and Sanriku hunts, respectively) are significantly different in mtDNA from both bycatch samples along the coast and from the offshore subareas 8+9. However, 7CN and 7CS are not significantly different from each other. We consider 7CN and 7CS (>10nm from shore) to represent Ow stock."

RESULTS

The new version of pure O stock includes 342 animals collected from subareas 7WR, 7E, and 8. Analyses described in the main text were repeated after substituting the new pure O stock for the sample from SA9. Table A1 shows the locus-specific θ values for the new combined sample, as well as for SA6 and SA9. Values for the new comparison (SA6 vs 7WR+7E+8) were highly correlated with those for both the full and trimmed SA6 vs SA9 comparisons (r = 0.98 and 0.99, respectively). Overall mean θ for the new 6 vs 7-8combined comparison was 0.044, intermediate to the values for the full and trimmed comparisons of SA6 and SA9.

Table A2 repeats some of the information from Table 2 for context and adds new data (in **bold**) showing the correlation coefficients between F_{IS} (or r^2) and θ (or $\theta_i\theta_j$), in putatively pure samples and in artificial 1:1 mixtures of putative J and O individuals. In the new sample of putatively pure O stock from SA7E+7WR+8, the correlations for both F_{IS} and r^2 were close to 0 (-0.04 and -0.01, respectively). In the new artificial mixture of SA6 and SA7E+7WR+8, the correlation of F_{IS} and θ was 0.64 ($P < 0.01$) and the correlation for r^2 and $\theta_i\theta_j$ was 0.82 ($P < 0.01$).

Table A3 repeats some data in Table 2 for context and adds new data in **bold** for putative mixtures. For the $F_{IS} - \theta$ correlations, use of the new reference sample for pure O had only a small effect that varied in sign among samples. For the r^2 x $\theta_i\theta$ correlations, use of the new reference sample consistently increased the correlations, but by a modest amount (about 5-10% increase compared to the original value). For the new combined sample SA7W-K+S (estimated to be about 25% J using Method 2 in Table 1), the F_{IS}-θ correlations were small for all combinations of reference populations considered, but the r^2 -$\theta_i\theta$ correlations were all moderately high (0.45-0.49; all $P < 0.01$).

DISCUSSION

Use of SAs 7WR+7E+8 as the reference sample for pure O stock (rather than SA9) had little overall effect on the anlayses described in the main text. θ values for individual loci changed only slightly. Analyses of the single-locus Wahlund effect (based on F_{IS}) showed small and variable changes; those for linkage disequilibrium were consistently positive but also small in magnitude.

For the new combined sample SA7W-K+S, results differed for the single-locus and two-locus analyses, which reinforced and strengthened the pattern seen in Table 3 in the main text. The single-locus analyses showed no evidence of the positive correlation between F_{IS} and θ that would be expected if the sample were a mixture of J and O stocks: correlation coefficients for the new combined sample were negative, just as they were weak or negative for the individual samples (Table A3). In contrast, correlations between LD and $\theta_i\theta_j$ were moderate in the individual samples (0.22-0.37) and even higher in the new combined sample (0.45-0.49). An explanation for this difference between single-locus and two-locus analyses is not readily apparent.

Table A1. Locus-specific θ values for three different comparisons of putative 'pure' O and J stocks. First two columns show data for the comparisons discussed in the main text (and plotted in Figure 1). Last column show results of comparison of SA6 with combined samples from 7E, 7WR, and 8.

Locus	SA6 vs SA9	Trimmed 6 vs 9	6 vs 7EW+8
EV37	0.008	0.010	0.007
EV1	0.031	0.035	0.032
GT310	0.066	0.077	0.075
GATA28	0.013	0.015	0.014
GT575	0.045	0.053	0.051
EV94	0.023	0.026	0.028
GT23	0.017	0.021	0.018
GT509	0.045	0.054	0.051
GATA98	0.007	0.008	0.002
GATA41	0.023	0.025	0.029
GT211	0.051	0.060	0.058
EV21	-0.001	-0.001	-0.001
DlrFB1	0.053	0.057	0.038
EV14	0.093	0.116	0.109
GT195	0.101	0.116	0.105
TAA31	0.059	0.058	0.056
Mean	0.041	0.047	0.044

Table A2. Correlation coefficients between θ and F_{IS} (top row) and between $\theta_i\theta_j$ and r^2 (LD; bottom row) in samples of putatively 'pure' J and O, as well as for artificial mixtures in proportions 1:1 Some data from Table 2 are included here for reference; new data are shown in **bold.**

	SA6	SA9	SA7EW+8	6 and 9 Mix	6 and (7EW+8) mix
F_{IS}	0.60	0.07	**-0.04**	0.83	**0.64**
LD	0.01	0.07	**-0.01**	0.84	**0.82**

Table A3. Correlation coefficients between θ or $\theta_i\theta_j$ (shown in Table A2 for putative pure stocks) and F_{IS} or r^2 (LD) in samples from putative mixtures. Some of the same data shown in Table 3 in main text are repeated for reference; new data are shown in **bold.**

			F_{IS}			LD		
	n	% J	Full 6-9	Trim 6-9	6-7EW+8	Full 6-9	Trim 6-9	6-7EW+8
SA7 Bycatch	212	0.54	0.30	0.31	**0.37**	0.26	0.27	**0.29**
SA7W-K	253	0.24	0.00	-0.04	**0.00**	0.22	0.22	**0.23**
SA7W-S	227	0.27	-0.15	-0.10	**-0.01**	0.33	0.34	**0.37**
SA2	184	0.80	0.32	0.35	**0.28**	0.26	0.26	**0.26**
SA11	80	0.35	0.24	0.23	**0.23**	0.09	0.11	**0.12**
SA7W-K+S	**480**	**0.25**	**-0.08**	**-0.09**	**-0.02**	**0.45**	**0.47**	**0.49**