Stock structure of North Pacific minke whales as revealed by Principal Component Analyses of microsatellite data

OSCAR GAGGIOTTI AND FANNY GASCUEL

Laboratoire d'Ecologie Alpine, UMR CNRS 5553, BP 53, 38041 Grenoble Cedex 9, France
Oscar.gaggiotti@ujf-grenoble.fr

SUMMARY

We present the results of a study aimed at exploring the genetic structure of North Pacific minke whales using Principal Component Analysis of genotypic data. We used microsatellite data provided by the Institute of Cetacean Research (Japan) corresponding to samples of common minke whales taken as bycatch and in scientific whaling programmes JARPN and JARPNII. The samples cover sub-areas 2, 6, 7, 8 and 9 in the western North Pacific. We explored three scenarios proposed by Gaggiotti and Durand (2010) comprising different degrees of purging the samples from J-stock individuals and added two additional scenarios (excluding putative O-stock individuals) proposed at the first intercessional workshop held in Busan in December 2010. The results suggest that minke whales in subareas 7-9 exhibit a hierarchical structure that comprises two main genetic clusters corresponding to the putative J and O stocks identified by STRUCTURE analyses and further substructuring within the O-stock. This substructuring, however, is not correlated with body length or geographic position so for the moment it has no clear biological interpretation. Results for areas 2 and 6 indicate that there is no genetic structuring among J-stock individuals East and West of Japan.

INTRODUCTION

There is a longstanding disagreement about the stock structure of western North Pacific minke whales. The number of alternative hypotheses has varied throughout the years but only three have been retained at the first intersessional workshop held in Busan in December 2010:

(I) there is a single J-stock distributed in the Yellow Sea, Sea of Japan, and Pacific coast of Japan, and a single O-stock in sub-areas 7, 8, and 9 (i.e. 'old' hypothesis 1);

(II) as for hypothesis (I), but there is a different stock (Y-stock) which resides in the Yellow Sea and overlaps with J-stock in the southern part of sub-area 6 (i.e. 'old' hypothesis 3); and

(III) there are five stocks, referred to Y, JW, JE, OW, and OE, two of which (Y and JW) occur in the Sea of Japan, and three of which (JE, OW, and OE) are found to the east of Japan (i.e., 'old' hypothesis 5).

One of the main differences between hypotesis III and the two others is that stocks J and O are each subdivided into two stocks. Some evidence for the distinction between O_W and O_E was presented at the intersessional workshop (SC/D10/NPM16). The study covered sub-areas 7, 8 and 9 and the same four scenarios proposed by Gaggiotti and Durand (2010), which comprise different degrees of purging the samples from J-stock individuals. A genetic cluster in the southwest of the study area (probably J-stock individuals) that is apparent when all individuals are included in the analyses disappears when J-stock individuals are excluded. However, a new cluster appears in the northwest. The genetic pool of this latter cluster is not intermediate between those of J- and O-stocks. Instead, it seems to be closer to that of the O-stock because it becomes apparent when J-stock individuals are removed from the analyses. Since this genetic cluster is found between 142.5°E and 147.5°E it is posited that it may represent the so-called O_W-stock.

In this report we present an update of SC/D10/NPM16 that considers both scientific whaling samples (JARPN and JARPNII) and adding bycatch samples. The objectives are two fold. Firstly, to further evaluate the plausibility of two alternative hypotheses concerning common minke whales in sub-areas 7, namely (a) they represent a mixture of O- and J-stock animals and (b) they represent a single stock with intermediate 'characteristics'. Secondly, to investigate the possibility that minke whales in subareas 2 and 6 represent a single stock or two distinct stocks (J_W and J_E).

MATERIALS AND METHODS

Data
We used microsatellite data provided by the Institute of Cetacean Research (Japan) corresponding to samples of common minke whales taken as bycatch between 2001-2007 and in scientific whaling programmes JARPN and

JARPNII from 1994-2007. The samples comprise 1946 individuals from sub-areas 7-9 and 603 samples from subareas 2 and 6. The microsatellite data included 16 loci and are described in Kanda et al. (2009).

Besides the genetic data, we also had at our disposal several other data for each individual. These included the year and month of capture, the geographic coordinates, distance to the coast, body length, sex, and probability of belonging to the O stock (obtained from microsatellite data).

Analyses

We used Principal Component Analyses based on individual microsatellite genotypes (Patterson et al 2006). We applied this methodology separately to the bycatch samples from areas 2 and 6 and to bycatch and scientific whaling samples from areas 7, 8, and 9. In this latter case, we followed the approach of Gaggiotti and Durand (2010) to explore the effect of different degrees of purging the data set from J-stock individuals. However, in these new analyses we also explored the effect of purging the data set from O-stock individuals. Thus, we considered five scenarios:

Scenario 1: all individuals were included regardless of their potential origin.

Scenario 2: only individuals identified as belonging to the J stock with a probability larger than 90% using the microsatellite criterion (those labelled "J" in the data base) were excluded.

Scenario 3: only individuals belonging to the O stock with a probability larger than 90% using the microsatellite criterion (those labelled "O" in the data base) were included in the analysis.

Scenario 4: individuals identified as belonging to the O stock with a probability larger than 70% using the microsatellite criterion (those labeled "O" and "O?" in the data base) were excluded.

Scenario 5: only individuals identified as belonging to the O stock with a probability larger than 90% using the microsatellite criterion (those labeled "O" in the data base) were excluded.

Scenario 1 considers no purging, scenario 2 represents mild purging and scenario 3 represents aggressive purging. The inclusion of scenario 4 and 5 allows us to verify if the substructuring observed under mild and aggressive purging (scenarios 2 and 3) disappears when O-stock individual are removed from the analysis.

In order to explain the genetic structuring observed we carried out separate regression analyses for the first and second PC axes against body length and month of capture. We also included longitude, latitude, and probability of belonging to the O stock.

All analyses were carried out using packages adegenet, prcomp, fields and glm in R (http://cran.r-project.org/).

RESULTS

The number of variables included in the PCAs is very large (up to 199, depending on the scenario considered) so the number of significant PC axes (p-values > 0.01) is substantial (up to 14; see Tables 1 and 2). The two first axes explain between 7.6% and 9.7% of observed variation in microsatellite data depending on the scenario considered. Below we present separately the results for areas 7-9 and for areas 2 and 6.

Areas 7, 8, and 9

The results indicate that there is significant genetic structuring under all the scenarios considered (Table 1). The purging of J-stock individuals does not lead to important changes in this regard. However eliminating O-stock individuals greatly decreases both the number of significant PC axes and their p-values. Nevertheless, this could be due in part to the large decrease in sample size that occurs when O-stock individuals are removed.

Figure 1 shows several representations of the results obtained for areas 7, 8 and 9. The PCA maps show clear spatial heterogeneity under Scenario 1 (all individuals). PC1 distinguishes two areas, East and West of 147°E. PC2 shows further heterogeneity to the East of 147°. Purging of J-stock individuals (Scenarios 2 and 3), leads to the appearance of a cluster centered at 145°E. The histograms and biplots also change with the purging of J-stock individuals. Interestingly, under scenario 1, there is no apparent clustering but under scenarios 2 and 3, it is possible to distinguish three clusters. These results suggest that minke whales exhibit a hierarchical genetic structuring with a first level of structuring corresponding to stocks J and O and then further substructuring within the O stock. Note that when O-stock individuals are removed (last two scenarios) the clusters disappear. In principle, this would confirm the substructuring of O-stock individuals but it could also be explained, at least in part, by a very large decrease in sample size when these individuals are excluded.

Table 1: Variance explained and p-value of first three PCs for microsatellite data set covering areas 7, 8, and 9. Last column gives the total number of significant PCs at p < 0.01.

Scenario	PC1	PC2	PC3	Nr of significant PCs
1) All	0.251859	0.199199	0.189695	12
	1.71655e-15	1.81092e-05	3.00075e-05	
2) No J	0.223631	0.200006	0.193030	14
q > 0.9	8.43293e-09	4.44713e-06	1.31837e-06	
3) Only O	0.227086	0.205223	0.193647	13
q > 0.9	1.60106e-09	2.37175e-07	6.99096e-07	
4) No O	0.257954	0.251304	0.226198	4
q > 0.7	0.00401116	0.000530303	0.00973906	
5) No 0	0.259992	0.242160	0.209856	2
q > 0.9	9.35567e-06	2.22248e-05	0.0131243	

In order to further verify that the clustering revealed using the PCA methodology is not an artifact, we simulated a panmictic population of 10000 individuals with equal sex ratios using Easypop (Balloux et al. 2001). We generated a sample of 1946 genotypes (the same number present in the microsatellite data set) scored for 16 loci with the same average number of alleles as that observed in the minke whale dataset. The geographic locations of the real samples were then randomly assigned to the simulated samples. The results (Figure 2) show that although some heterogeneity is observed in the maps of PC1 and PC2, histograms and biplots do not exhibit any clustering. Moreover, only the first PC is significant (p-value = 5.66313e-06) confirming that there is no genetic structuring.

In order to explain the observed genetic structuring we carried out regression analyses using the PC1 and PC2 scores as dependent variable and the probability of belonging to the O-stock, body length, date of sampling, latitude and longitude as explicative variables. For these analyses we only used microsatellite data from scientific whaling (JARPN, JARPNII). The effect of the probability of assignment to stock O is significant under all scenarios for PC1 while its effect on PC2 is not significant under scenario 3 (harsh purging of J-stock individuals). This indicates that there is a good concordance between results obtained with STRUCTURE and PCA methods. It also suggests that the substructuring of the O-stock would be revealed by a STRUCTURE analysis focused only on this stock. The effect of body length is significant only for PC1 and under scenario 1 (all individuals included). Sampling date and latitude have no effect at all while longitude has a significant effect only under scenario 1.

Table 3. Relationship between explicative variables and PC scores. Shown are the p-values for the correlation between PCs and explicative variables.

Scenario	Axis	Pr Assignment	Body Length	Sampling date	Latitude	Longitude
1) All	PC1	2e-16	7.94e-4	-	-	0.0463
	PC2	2.09e-8	-	-	-	0.0396
2) No J q > 0.9	PC1	4.19e-8	-	-	-	-
	PC2	-	-	-	-	-
3) Only O q > 0.9	PC1	8.13e-14	-	-	-	-
	PC2	-	-	-	-	-

The significant correlations with body length and longitude when both J and O individuals are considered were expected from previous knowledge. The absence of a correlation with body length when only O-stock

individuals are considered indicates that the substructuring at the genetic level is subtle and does not translate in morphological differences. Finally, the absence of an association with geographical location is due to the non-linear relationship between genetic structure and geographical coordinates.

Areas 2 and 6

We carried a more limited analysis of the bycatch samples from subareas 2 and 6 in order to evaluate the possibility that they comprise two different stocks.

Here we considered only three scenarios: 1) all individuals included, 2) only J-stock individuals with a probability of assignment of at least 0.7 were included, and 3) only J-stock individuals with a probability of assignment of at least 0.9 were included. The results show that there is much less structuring among these individuals. Only two PC axes are significant for scenarios 1 and 2 and only 1 for scenario 3 (Table 2). Figure 3 shows that although some spatial structuring can be observed in the maps of PC1 and PC2, the histograms and biplots are very similar to those obtained from a panmictic population (c.f. Figure 2). We conclude, therefore, that there is no structuring among J-stock individuals East and West of Japan.

Table 2: Variance explained and p-value of first three PCs for microsatellite data set covering areas 2 and 6. Last column gives the total number of significant PCs at p < 0.01.

Scenario	PC1	PC2	PC3	Nr of significant PCs
All	0.249932	0.235616	0.205672	2
	0.00111074	0.00143518	0.125995	
Only J	0.252308	0.237991	0.212290	2
q > 0.7	0.00734025	0.00854219	0.137285	
Only J	0.256388	0.239290	0.216331	1
q > 0.9	0.00590444	0.0114818	0.0991089	

CONCLUSIONS

These results provide support for the existence of a hierarchical structure among minke whales from subareas 7-9. More precisely, there seems to be two main genetic clusters corresponding to the putative J and O stocks identified by STRUCTURE analyses (Kanda et al. 2009) and further substructuring within the O-stock. This substructuring is subtle and uncorrelated with body length. These results are concordant with those presented at the intersessional workshop held in Busan.

Besides further exploring the substructuring within the O-stock, we also explored the possibility of substructuring within the J stock. In this latter case, however, our results do not support the existence of two different stocks East and West of Japan.

REFERENCES

Balloux, F.. 2001. EASYPOP (version 1.7), A computer program for the simulation of population genetics. J. Heredity 92: 301-302.

Gaggiotti OE and E Durand. 2010. Update analyses of the genetic structure of Western North Pacific minke whales. SC62/NPM30.

Goto, M, N Kanda, T Kishiro, H Yoshida, H Kato and LA Pastene. 2009. Further mitochondrial DNA analysis on stock structure in the western North Pacific common minke whales. IWC Scentific Committee Document SC/61/JR7.

Kanda N, M Goto, T Kishiro, H Yoshida, H Kato and LA Pastene. 2009. Update of the analyses of individual identification and mixing of J and O stocks of common minke whale around Japanese waters examined by microsatellite analysis. IWC Scientific Committee Document SC/61/JR5.

Patterson, N, AL Price, D Reich. 2006. Population Structure and Eigenanalysis. PLoS Genetics 2(12): e190
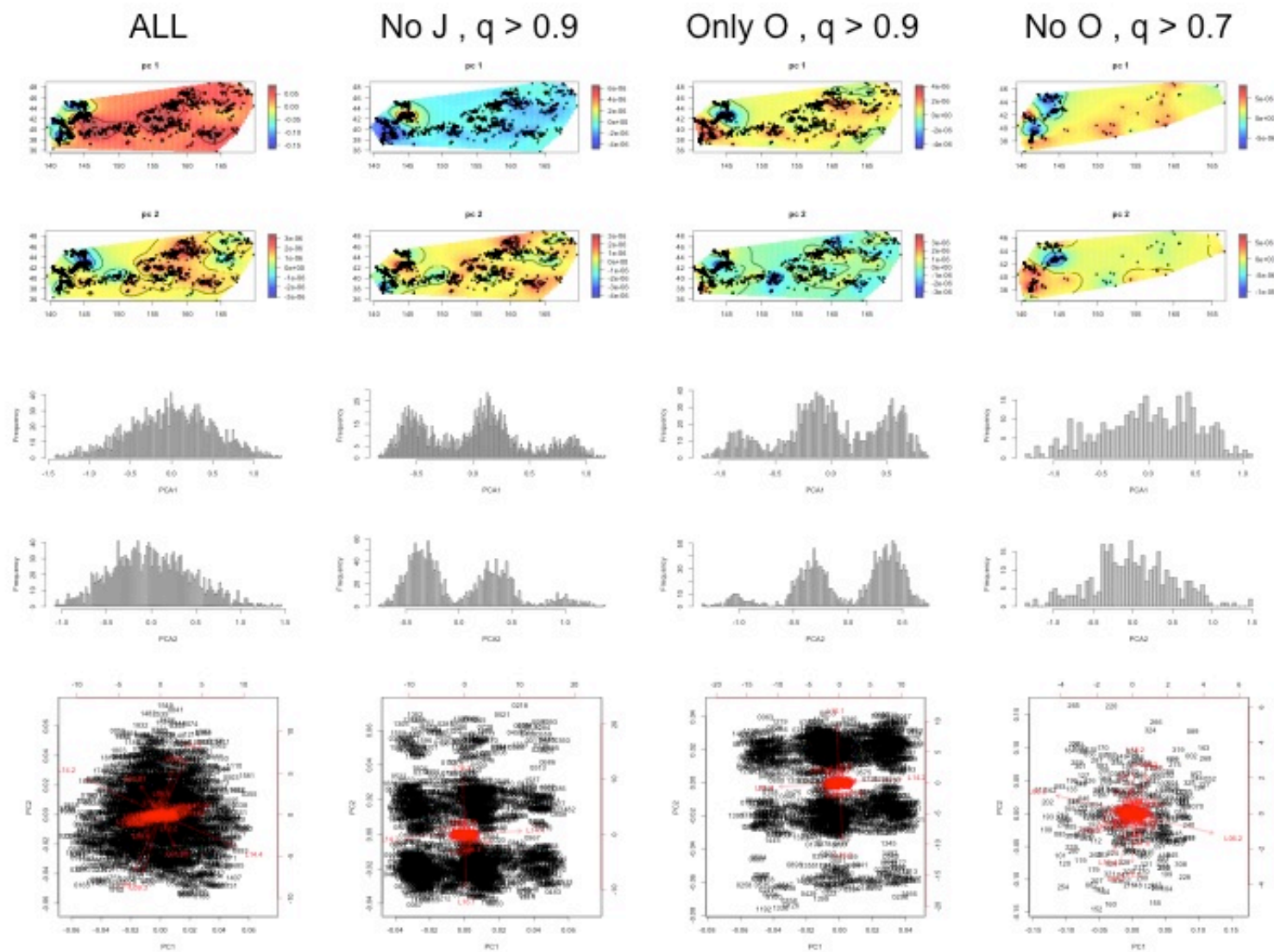
Figure 1: PCA results for areas 7-9 under four different scenarios (see text)
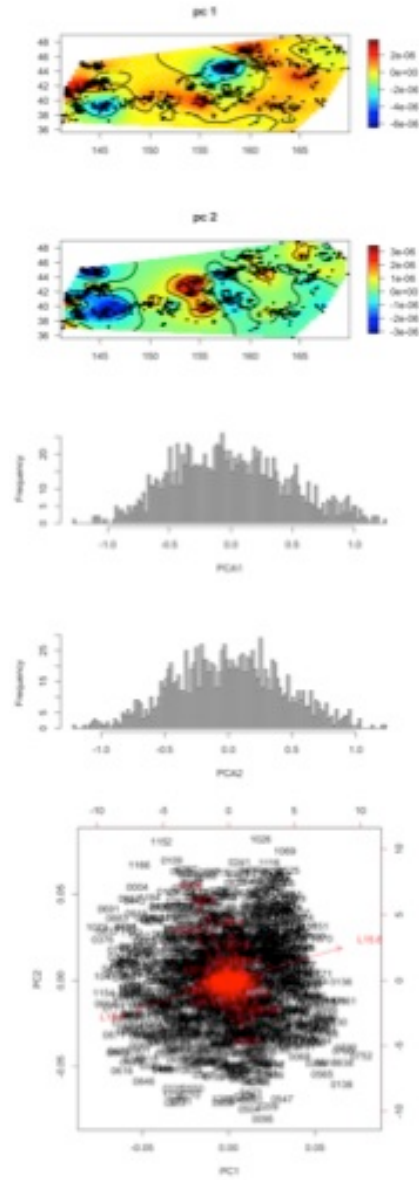
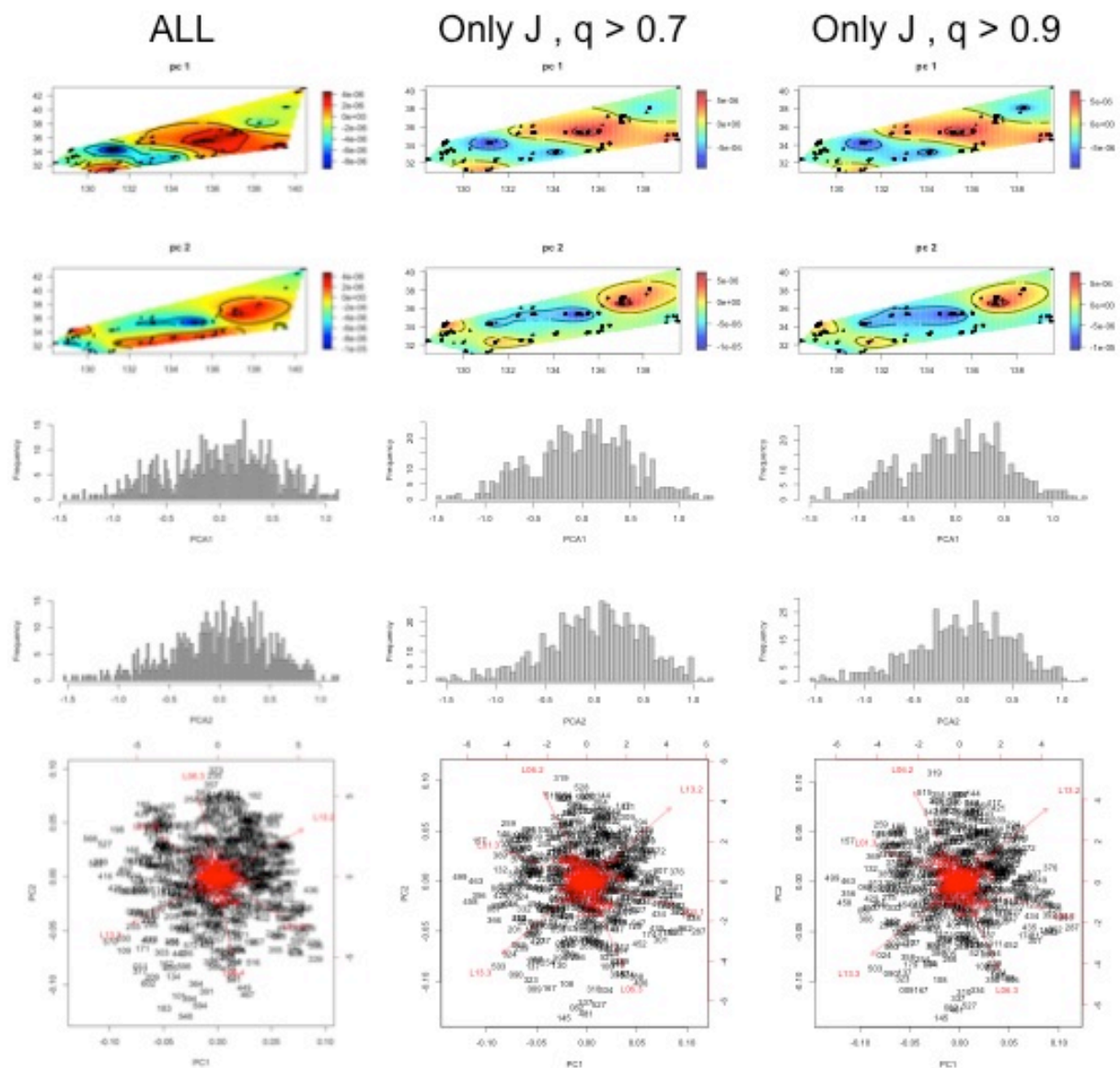Figure 2: PCA results for the simulated data of a single panmictic population

Figure 3: PCA results for minke whales from subareas 2 and 6.