SC/MAR17/RMP/01

DAPC analysis for Brydes whales in the North Pacific using microsatellite DNA data

Taguchi, M., Goto, M., Takahashi, M., Kitakado, T., Pastene, L.A.

INTERNATIONAL WHALING COMMISSION

DAPC analysis for Bryde's whales in the North Pacific using microsatellite DNA data

Mioko Taguchi[1], Mutsuo Goto[1], Megumi Takahashi[1, 2], Toshihide Kitakado[2] and Luis A. Pastene[1]

[1] *Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo 104-0055, Japan*
[2] *Tokyo University of Marine Science and Technology, 1-5-7 Konan, Minato-ku, Tokyo 108-0075, Japan*
Contact e-mail: taguchi@cetacean.jp

ABSTRACT

To examine the stock structure of the Bryde's whales in the North Pacific, a Discriminant Analysis of Principal Component (DAPC) was conducted using seventeen microsatellite DNA loci of a total of 1,019 Bryde's whales which had been collected in sub-areas 1W, 1E and 2 till 2014. The same dataset was used in the previous analyses based on heterogeneity test and STRUCTURE analysis. Microsatellite data from other oceanic basins were also applied to the DAPC analysis for comparative purposes: eastern South Pacific; off Peru (n=48), western South Pacific; off Fiji (n=25) and eastern Indian Ocean; off Java (n=50). The pairwise F_{ST} among the sampling areas were estimated to assist the interpretation of the stock structure derived from the DAPC analysis. The DAPC analyses revealed no structure within the North Pacific, however, showed that Bryde's whales from the North Pacific, eastern and western South Pacific and eastern Indian Ocean belong to four distinct stocks. The negative results of DAPC analysis for the North Pacific were explained by the low F_{ST} estimates among the three sub-areas (1W, 1E and 2), and these results were consistent with the previous STRUCTURE results that revealed no structure among the three sub-areas. Considering that the previous heterogeneity tests showing no differences within sub-area 1 but significant differences between sub-areas 1 and 2, for both mitochondrial and microsatellite DNA, the present observations would mean a low degree of genetic differentiation between those sub-areas. Taking all results of the three different analyses (heterogeneity test, STRUCTURE and DAPC) together, it is suggested that two weakly differentiated stocks occur in the research area; one distributed mainly in sub-area 1 and the other distributed mainly in sub-area 2. Future research should include investigation on possible boundaries (or areas of mixing) between the stocks as the longitude line separating sub-areas 1 and 2 at 180°, is arbitrary.

INTRODUCTION

The RMP *Implementation* for the western North Pacific Bryde's whale was completed by the IWC SC in 2007 (IWC, 2008). During the *Implementation*, two sub-areas (Figure 1; IWC, 2009) and four stock structure hypotheses (IWC, 2007a), were used. Hypothesis 1 postulates a single stock in sub-areas 1 and 2; Hypothesis 2 postulates two stocks, one in sub-area 1 and the other in sub-area 2; Hypothesis 3 postulates two stocks, one in sub-areas 1 and 2 and the other in sub-area 2; and Hypothesis 4 postulates three stocks, two in sub-area 1 and the other in sub-area 2. The IWC SC examined the plausibility of the four hypotheses based on genetics and non-genetics information which was available in 2006 (IWC, 2007b), and agreed on the following plausibility: Hypothesis 1: high; Hypothesis 2: high; Hypothesis 3: high and Hypothesis 4: medium.

It should be noted here that genetic data and other information used in the previous *Implementation* involved samples which had been collected till 2003, and that a substantial number of samples and analyses are available now for the period up to 2014, from different sources and surveys.

Pastene *et al.* (2016a) updated the genetic analyses based on heterogeneity test of the North Pacific Bryde's whale using the extended mitochondrial and microsatellite datasets which had been collected from sub-area 1 through JARPNII, Japanese dedicated sighting and IWC/POWER surveys since the 2007 *Implementation*, as well as from sub-area 2 through the IWC/POWER surveys. The analyses in Pastene *et al.* (2016a) were conducted to examine further the plausibility of the four stock structure hypotheses (IWC, 2007a; 2007b) in consideration with the recommendations from the 2009 JARPNII review workshop (IWC, 2010) and the subsequent IWC SC Annual meetings (see Annex 5 of Tamura *et al.*, 2016). The study by Pastene *et al.*

(2016a) showed no significant genetic heterogeneity between whales in sub-areas 1W and 1E using both mitochondrial and microsatellite DNA markers, despite a considerable high statistical power. This finding suggested no stock structure within sub-area 1, which was consistent with the results from the previous mark-recapture (Kishiro, 1996) and satellite tagging studies (Murase *et al.*, 2016). Pastene *et al.* (2016a) also demonstrated significant differentiation of this species between sub-areas 1 and 2 for the first time, which suggested the possibility of the existence of different stocks in each sub-area. Regarding these results, the JARPNII final review workshop noted that the heterogeneity tests conducted by Pastene *et al.* (2016a) were aimed only at assessing spatial heterogeneity within sub-area 1 as well as between sub-areas 1 and 2 (IWC, 2017), and recommended that 'the presence of multiple stocks within sample partitions should be assessed employing, e.g. STRUCTURE and DAPC' (IWC, 2017). In response to this recommendation, Pastene *et al.* (2016b) examined the presence of additional stocks of Bryde's whales in the North Pacific using the Bayesian STRUCTURE analysis, and this study supported the hypothesis of a single stock of the Bryde's whale in sub-areas 1 and 2 in contrast to the inference from the heterogeneity tests by Pastene *et al.* (2016a).

The purpose of this study was to conduct a DAPC analysis using the same dataset as the previous two studies (Pastene *et al.*, 2016a; 2016b) which was the other recommendation by the JARPNII review workshop. Interpretation of the results from different approaches will assist the refinement of the stock structure hypotheses of the North Pacific Bryde's whale, which is one of the tasks for the upcoming RMP *Implementation* of this species.

MATERIALS AND METHODS

Dataset

The 17 microsatellite DNA data from a total of 1,019 Bryde's whales (Table 1), which was used in the work by Pastene *et al.* (2016a; 2016b), was used to examine the genetic structure of the North Pacific Bryde's whale by DAPC. The microsatellite DNA datasets of this species from the eastern South Pacific off Peru (n=48), western South Pacific off Fiji (n=25) and eastern Indian Ocean off Java (n=50), used by Kanda *et al.* (2007), were also used for comparative purposes.

Data analysis

The DAPC (Jombart *et al.*, 2010) in the R package *adegenet* was conducted for each of the two types of dataset, i.e., the North Pacific only and the entire Pacific and eastern Indian Oceans, to identify and describe clusters of genetically related individuals using the most likely number of genetic clusters (K). This analysis relies on data transformation using PCA as a prior step to DA which ensures that variables submitted to DA are perfectly uncorrelated, and that their number is less than that of individuals analyzed. Along with the assignment of individuals to clusters, DAPC provides a visual assessment of between-population genetic structures, permitting to infer complex patterns such as hierarchical clustering (Jombart *et al.*, 2010). The most likely number of genetic clustering of individuals was determined without prior information on population groupings using the function '*find.clusters*', which runs successive k-means clustering with increasing number of clusters (K) to achieve the optimal number of groups based on the minimum value of the Bayesian information criterion (BIC) (Jombart *et al.*, 2010). The DAPC using a priori geographical group assignments based on the sampling area was also conducted at K=3, i.e., sub-areas 1W, 1E and 2, for the North Pacific, as well as at K=6, i.e., sub-areas 1W, 1E, 2, Peru, Fiji, and Java.

The pairwise F_{ST} (Weir and Cockerham, 1984) with 95% confidence intervals were estimated using the *diveRsity* package in R to show the level of genetic differentiation between sampling areas, assuming that the level of genetic differentiation between sampling areas influences on the results of DAPC.

RESULTS

DAPC for the North Pacific

Figures 2 and 3 show the results of DAPC analysis for North Pacific samples only, without and with a priori group assignments, respectively. The DAPC based on the k-means clustering method, meaning without a priori group assignments, did not discriminate among clusters (K=10, Figure 2a), and samples from different geographical origin were widely distributed among the clusters (Figure 2b). When grouping the populations into three geographic groups, i.e., sub-areas 1W, 1E and 2, although the analysis showed that

a few separation was evident along the first and second discriminant function axes, substantial overlap among geographical clusters was remained (Figure 3).

DAPC for the entire Pacific and eastern Indian Oceans

Figures 4 and 5 show the results of DAPC analysis without and with a priori group assignments for the entire Pacific and eastern Indian Ocean, respectively. However the DAPC based on the k-means clustering method did not completely discriminate among clusters (K=10, Figure 4a), the majority of individuals in Java of the eastern Indian Ocean were assigned to cluster 3 with a lesser contribution to other clusters (Figure 4b). Furthermore, a number of individuals from the Pacific sampling areas belonged to each of the clusters (Figure 4b). When using a geographic group (K=6; sub-areas 1W, 1E, 2, Fiji, Java and Peru), DAPC separated not only the Javanese but also the Fijian samples from the North Pacific and Peruvian ones along the first and second principal component axes, respectively (Figure 5). Peruvian samples were also plotted distant from the North Pacific clusters, although some overlapping among those clusters was remained on the ordination plot indicating a low degree of genetic differentiation (Figure 5).

Genetic distances between sampling areas

The pairwise F_{ST} estimates between the sampling areas within the North Pacific were not significantly different from zero (Figure 6), meaning extremely weak or no genetic differentiation between sub-areas in the North Pacific. In contrast, the distinct genetic differentiation was observed between sampling areas from different oceans (Figure 6), which also suggested that a genetic differentiation from North Pacific was stronger of the Javanese and Fijian Bryde's whales than the Peruvian.

DISCUSSION

The present DAPC analyses did not show evidence of genetic structure of the North Pacific Bryde's whales in sub-areas 1 and 2, which was consistent with the previous STRUCTURE analyses conducted by Pastene *et al.* (2016b). Considering, however, that the heterogeneity tests conducted by Pastene *et al.* (2016a) showed significant genetic differentiation between these two sub-areas for both mitochondrial and microsatellite DNA markers, it is highly likely that two weakly differentiated stocks of the North Pacific Bryde's whales occur in sub-areas 1 and 2, respectively.

The DAPC analysis for Bryde's whales in the entire Pacific and the eastern Indian Oceans demonstrated distinct three genetic groups, i.e., Fiji, Java, and the North Pacific and Peru. Although the DAPC analysis also showed that Peruvian Bryde's whales are differentiated from the North Pacific ones, a finer genetic differentiation within the North Pacific was not observed. These observations were consistent with the results of STRUCTURE analyses carried out by Kanda *et al.* (2007). Given that the pairwise F_{ST} estimates between sub-areas in the North Pacific did not significantly differ from zero, it is suggested that the DAPC analysis could not detect the genetic structure suggested by the previous heterogeneity test due to the weak genetic differentiation, as it was the case of the STRUCTURE analysis. A parsimonious interpretation of the different results is that two weakly differentiated stocks occur in sub-areas 1 and 2. This inference, coupled with the shallow divergence between mitochondrial haplotypes of the North Pacific Bryde's whale (Pastene *et al.*, 2016a), would also mean that the Bryde's whales in sub-areas 1 and 2 are ecologically restricted, but they have been closely related on an evolutionary time scale. Future research should include investigation on possible boundaries (or areas of mixing) between the stocks as the longitude line separating sub-areas 1 and 2 at 180°, is arbitrary.

In summary, the present and the two previous studies (Pastene *et al.*, 2016a; 2016b) demonstrated a single stock of Bryde's whale in sub-area 1, and a weakly subdivision of this species between sub-areas 1 and 2. Genetic results suggesting a single stock in sub-area 1 are consistent with the pattern of movement of whales in this sub-area shown by mark-recapture (Kishiro, 1996) and satellite tagging studies (Murase *et al.*, 2016).

ACKNOWLEDGEMENTS

We gratefully acknowledge the researchers and crew members that participated in JARPN/JARPNII, IWC/POWER and Japanese dedicated sighting surveys for the collection of genetic samples. We thank Mr. Tomonari Iida (ICR) who helped the laboratory works. We also thank Dr. Naohisa Kanda (JANUS-Japan) for preparing the microsatellite dataset used in this study.

REFERENCES

International Whaling Commission. 2007a. Report of the Scientific Committee. *J. Cetacean Res. Manage. 9* (*Suppl.*): 1-73.

International Whaling Commission. 2007b. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure (RMP). *J. Cetacean Res. Manage. 9* (*Suppl.*): 88-128.

International Whaling Commission. 2008. Report of the Scientific Committee. *J. Cetacean Res. Manage. 10 (Suppl.)*: 1-74.

International Whaling Commission. 2009. Report of the Scientific Committee. *J. Cetacean Res. Manage. 11* (*Suppl.*): 1-74.

International Whaling Commission. 2010. Report of the Expert Workshop to Review the Ongoing JARPN II Programme. *J. Cetacean Res. Manage. 11 (suppl. 2)*: 405-449.

International Whaling Commission. 2017. Report of the Expert Panel of the final review on the western North Pacific Japanese Special Permit programme (JARPN II). *J. Cetacean Res. Manage.* 18: xxx-xxx.

Kanda, N., Goto, M., Kato, H., McPhee, M.V. and Pastene, L.A. 2007. Population genetic structure of Bryde's whales (*Balaenoptera brydei*) at the inter-oceanic and trans-equatorial levels. *Conserv. Genet.* 8: 853-864.

Kishiro, T. 1996. Movements of Marked Bryde's Whales in the Western North Pacific. *Rep. Int. Whal. Commn.* 46: 421-428.

Murase, H., Tamura, T., Otani, S. and Nishiwaki, S. 2016. Satellite tracking of Bryde's whales *Balaenoptera edeni* in the offshore western North Pacific in summer 2006 and 2008. *Fish. Sci.* 82: 34-45.

Pastene, L.A., Goto, M., Taguchi, M. and Kitakado, T. 2016a. Updated genetic analyses based on mtDNA and microsatellite DNA suggest possible stock differentiation of Bryde' whales between management sub-areas1 and 2 in the North Pacific. Paper SC/F16/JR44 presented to the JARPNII special permit expert panel review workshop, Tokyo, February 2016 (unpublished). 17pp.

Pastene, L.A., Goto, M. and Taguchi, M. 2016b. Additional genetic analyses on stock structure in North Pacific Bryde's and sei whales. Paper SC/66b/SD01 presented to the IWC Scientific Committee, Bled, June 2016 (unpublished). 11pp.

Tamura, T., Kishiro, T., Yasunaga, G., Murase, H., Kitakado, T and Pastene, L.A. 2016. The Japanese Whale Research Program under Special Permit in the western North Pacific Phase-II (JARPN II): results and conclusions in the context of the three main objectives, and scientific considerations for future research. Paper SC/F16/JR1 presented to the JARPNII special permit expert panel review workshop, Tokyo, February 2016 (unpublished). 68pp.

Jombart, T., Devillard, S. and Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11: 94.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evol.* 38: 1358-1370.

Table 1. Sample sizes of Bryde's whales used in the present study, by sampling area and source of samples.

Sampling area	JARPNII	Commerc.	POWER	Jap. dedic.	Spec. permit	Total
Sub-area 1W	620	171		56		847
Sub-area 1E	59	28	30	2		119
Sub-area 2		1	52			53
Java					50	50
Fiji					25	25
Peru		48				48
Total	679	248	82	58	75	1142

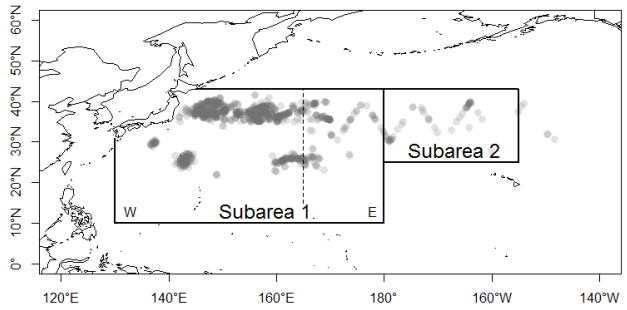

Figure 1. Map of the locations where Bryde's whale specimens were collected.

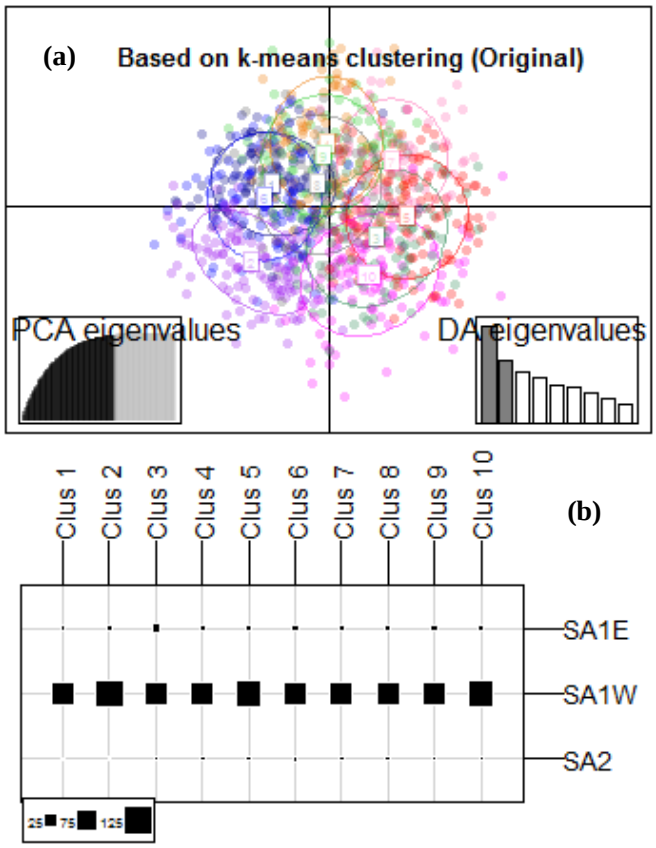

Figure 2. Results of DAPC analysis based on k-means clustering method for the North Pacific Bryde's whale; (a) DAPC scatter plot, (b) geographical distribution of the North Pacific Bryde's whales which were grouped into each cluster by k-means method

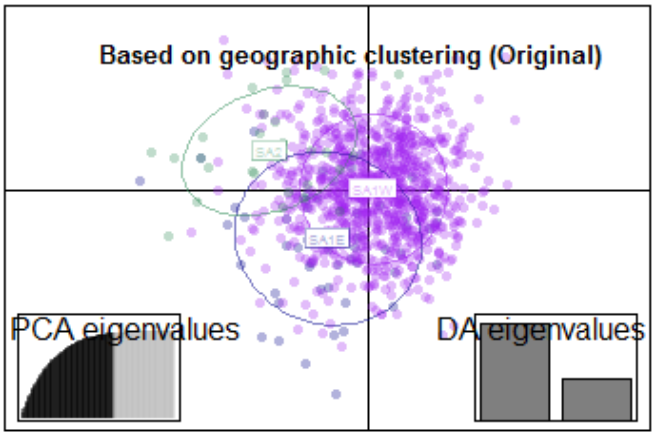

Figure 3. Results of DAPC analysis based on geographic clustering for the North Pacific Bryde's whale

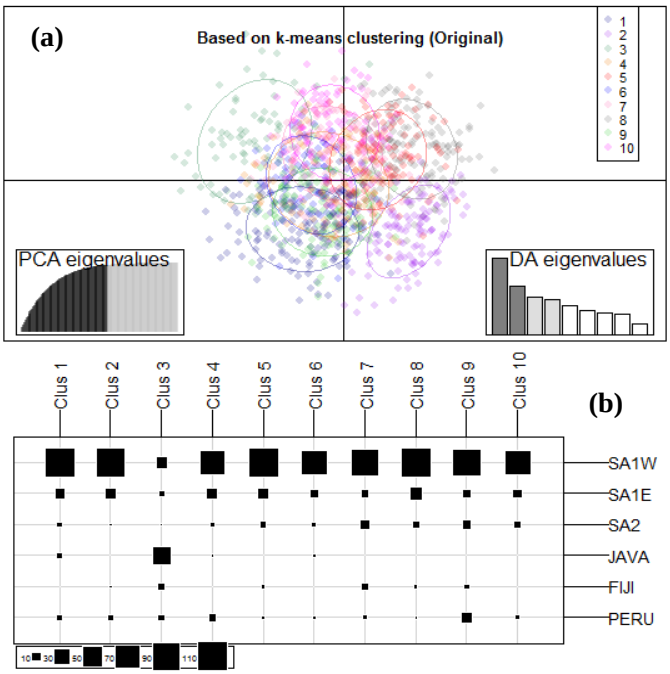

Figure 4. Results of DAPC analysis based on k-means clustering method for the Bryde's whale in the entire Pacific and eastern Indian Oceans; (a) DAPC scatter plot, (b) geographical distribution of the Pacific Bryde's whales which were grouped into each cluster by k-means method

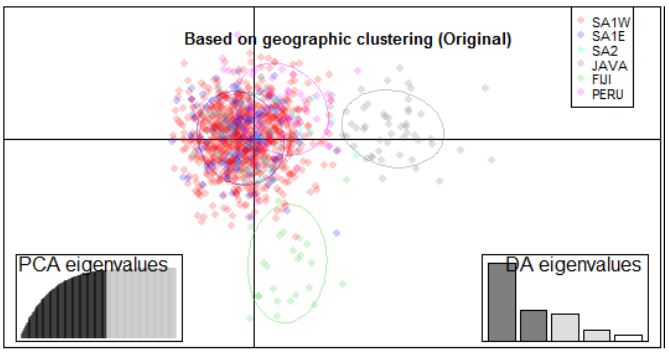

Figure 5. Results of DAPC analysis based on geographic clustering for the Bryde's whale in the entire Pacific and eastern Indian Oceans

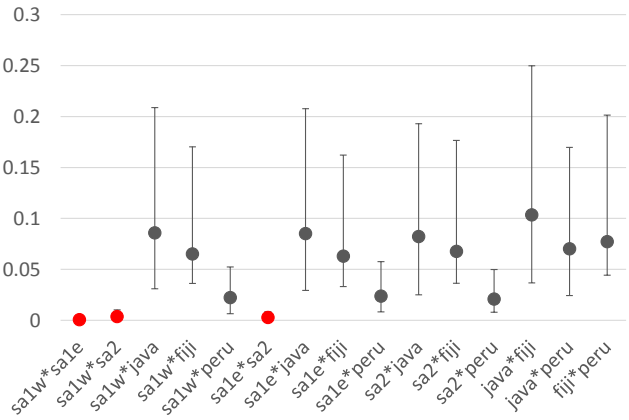

Figure 6. Genetic distance (pairwise F_{ST} estimates) between sampling areas; red symbols indicate pairwise F_{ST} estimates between sub-areas in the North Pacific