Technical Report (not peer reviewed)

Recent progress in studies of the stock structure of western North Pacific common minke whales

Mioko TAGUCHI*

Institute of Cetacean Research, 4–5 Toyomi-cho, Chuo-ku, Tokyo 104–0055, Japan

**Contact e-mail: taguchi@cetacean.jp*

ABSTRACT

This paper outlines recent progress in studies of the stock structure of western North Pacific common minke whales (*Balaenoptera acutorostrata*), focusing on analyses based on the GENELAND framework. The motivation to focus on GENELAND is that results of this approach in 2019 suggested a new stock structure hypothesis different from those currently accepted by the IWC Scientific Committee (a J stock primarily west of Japan and an O stock primarily east of Japan, with spatial mixing along the Pacific coast of Japan and the southern coastal Okhotsk Sea) as having high plausibility. This new hypothesis included an additional coastal stock (P stock) in those coastal regions, which consists of two genetic clusters inferred by GENELAND. Results from recent analyses suggest that inferences from GENELAND under a spatial model are not very robust for ICR′s dataset. These results also indicate that the inference of the existence of a P stock likely resulted from over-splitting the mixing zone (i.e., sub-areas 2C, 7CS, 7CN, and 11), which is known for spatial mixture between two distinct stocks (the J and O stocks), into more coastal (and primarily J) and more offshore (and primarily O) components. These components would represent J-affiliated and O-affiliated genetic groups rather than a single distinct population.

INTRODUCTION

In the western North Pacific, at least two stocks of the common minke whale have been recognized: (1) the Okhotsk Sea–West Pacific (known as the O stock), and (2) the Sea of Japan–Yellow Sea–East China Sea (known as the J stock). They are distinguished morphologically and reproductively (e.g., Omura and Sakiura, 1956; Kato, 1992), and genetically (e.g., Wada and Numachi, 1991; Goto and Pastene, 1997). In the context of management, many genetic analyses have been conducted, as reviewed by Pastene *et al.* (2022). Figure 1 shows the International Whaling Commission Scientific Committee (IWC SC)′s management sub-areas for western North Pacific common minke whales.

In 2019, the IWC SC reviewed the available genetic and non-genetic information and proposed three stock structure hypotheses for western North Pacific common minke whales (IWC, 2020):

Hypothesis A: there is a single J stock distributed in sub-areas 1W, 1E, 2C, 5, 6W, 6E, 7CS, 7CN, 10W, 10E, 11 and 12SW, and a single O stock in sub-areas 2C, 2R, 3, 4, 7CS, 7CN, 7WR, 7E, 8, 9, 9N, 10E, 11, 12SW, 12NE and 13;
Hypothesis B: as for hypothesis A, but there is a third stock (Y) that resides in sub-area 1W, 5 and 6W and overlaps with J stock in the southern part of sub-area 6W; and
Hypothesis E: there are four stocks, referred to Y, J, P, and O, two of which (Y and J) occur to the west of Japan, and only three of which (J, P, and O) are found to the east of Japan and in the

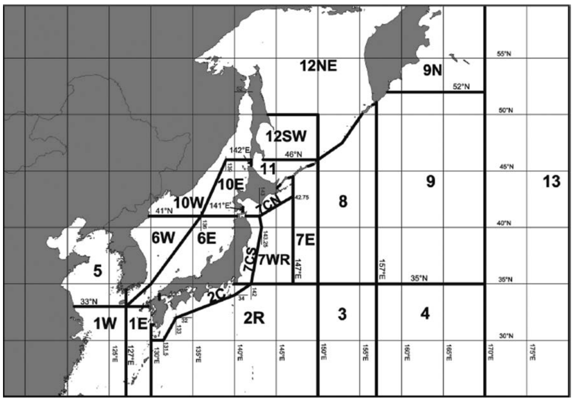

Figure 1. Management sub-area defined by IWC SC for western North Pacific common minke whales (IWC, 2020).

Okhotsk Sea. Stock P is a coastal stock.

The IWC SC assigned high plausibility to Hypotheses A and B. However, it agreed that the plausibility of Hypothesis E could not be evaluated until the results of the conditioning process (fitting the associated population model to the time series of abundance and related demographic data available) became available and simultaneously recommended several further analyses before plausibility could be assigned to this hypothesis (IWC, 2020).

Hypothesis E was based on preliminary genetic analyses (de Jong and Hoelzel, 2019) using the spatially explicit Bayesian clustering program GENELAND (Guillot *et al.*, 2005), which divides the study area into K geographically coherent clusters and assigns each individual whale to one of them, based on multilocus genotypes and spatial coordinates. This analysis identified four genetic clusters in western north Pacific common minke whale (i.e., green, red, blue, and orange), and subsequent discussions led to the integration of the red and blue clusters into a single group termed purple (currently referred to as the P stock) (IWC, 2020). However, as noted above, the plausibility of Hypothesis E, which assumes the existence of the P stock, has not yet been evaluated, and further analyses were recommended (IWC, 2020).

In this report, I present an outline of recent progress in stock structure analyses of western North Pacific common minke whales, focusing on the GENELAND framework. The analysis presented here uses the same dataset as de Jong and Hoelzel (2019), which comprises microsatellite genotypes at 16 loci from 4,656 whales collected between 1994 and 2016. Hereafter referred to as the 'ICR's data set'.

ANALYTICAL PROCEDURES AND RESULTS

GENELAND results varied according to parameter settings

To explore the performance of the spatial model in GENELAND, which incorporates geographic coordinates into the clustering process, when applied to the ICR's dataset, one spatial baseline and several spatial and non-spatial variant analyses were performed, each with 5 or 10 replicates. The non-spatial model has been tested here for the first time. Table 1 summarizes the parameter settings and outcomes of each analysis. The spatial baseline run converged in only three out of ten replicates at K=4, where K denotes the number of clusters. Spatial variants (SpVar and SpFix series) showed similarly unstable con-

Table 1. Summary of GENELAND parameter settings and outcomes: 'npopmin'=lower bound of the number-of-clusters (K) search range; 'npopinit'=initial K value at the start of MCMC iterations; 'npopmax'=upper bound of the K search range; 'spatial'=whether spatial coordinates were used (TRUE) or not (FALSE); 'varnpop'=parameter specifying whether the K value is fixed (FALSE) or estimated by GENELAND (TRUE); 'seed'=random number seed used for each run; 'rep'=number of replicate runs; 'conv.'=number of converged runs; 'bestK'=most frequently supported K. The baseline parameter settings followed de Jong and Hoelzel (2019). Series highlighted were used in the subsequent analyses shown in Figures 2, 3, 5 and 6, and Figure 4, respectively: the blue series to examine the reproducibility and robustness of GENELAND results of de Jong and Hoelzel (2019), and the orange series to evaluate the influence of spatial information on the inference.

Series	npopmin	npopinit	npopmax	spatial model	varnpop	seed	rep	conv.	bestK
Baseline	1	7	7	TRUE	TRUE	101-110	10	3	4
SpVar1-5	1	1	7	TRUE	TRUE	101-105	5	2	2
SpVar3-5	1	3	7	TRUE	TRUE	101-105	5	5	3
SpVar5-5	1	5	7	TRUE	TRUE	101-105	5	2	4
SpVar7-5	1	7	7	TRUE	TRUE	101-105	5	2	4
SpFix2-10	2	2	2	TRUE	FALSE	101-110	10	4	—
SpFix3-10	3	3	3	TRUE	FALSE	101-110	10	7	—
SpFix4-10	4	4	4	TRUE	FALSE	101-110	10	5	—
NoSpVar7-10	1	7	7	FALSE	TRUE	101-110	10	10	7
NoSpVar2-10	1	2	7	FALSE	TRUE	101-110	10	10	2
NoSpVar3-10	1	3	7	FALSE	TRUE	101-110	10	10	3
NoSpFix2-10	2	2	2	FALSE	FALSE	101-110	10	10	—
NoSpFix3-10	3	3	3	FALSE	FALSE	101-110	10	10	—
NoSpFix4-10	4	4	4	FALSE	FALSE	101-110	10	10	—

vergence, and when GENELAND estimated K (vernpop= TRUE), the inferred values were inconsistent among runs and replicates, ranging from 2 to 4. In contrast, the non-spatial variants (NoSpVar and NoSpFix series) consistently converged in all runs and replicates, but the K search often remained fixed at the initial value (npopinit). These results indicate that the GENELAND analyses are sensitive to parameter settings, particularly under the spatial model, for the ICR's dataset.

Partial reproduction of the four GENELAND clusters by the baseline

The three converged spatial baseline replicates reproduced the four GENELAND clusters (Figure 2: red, green,

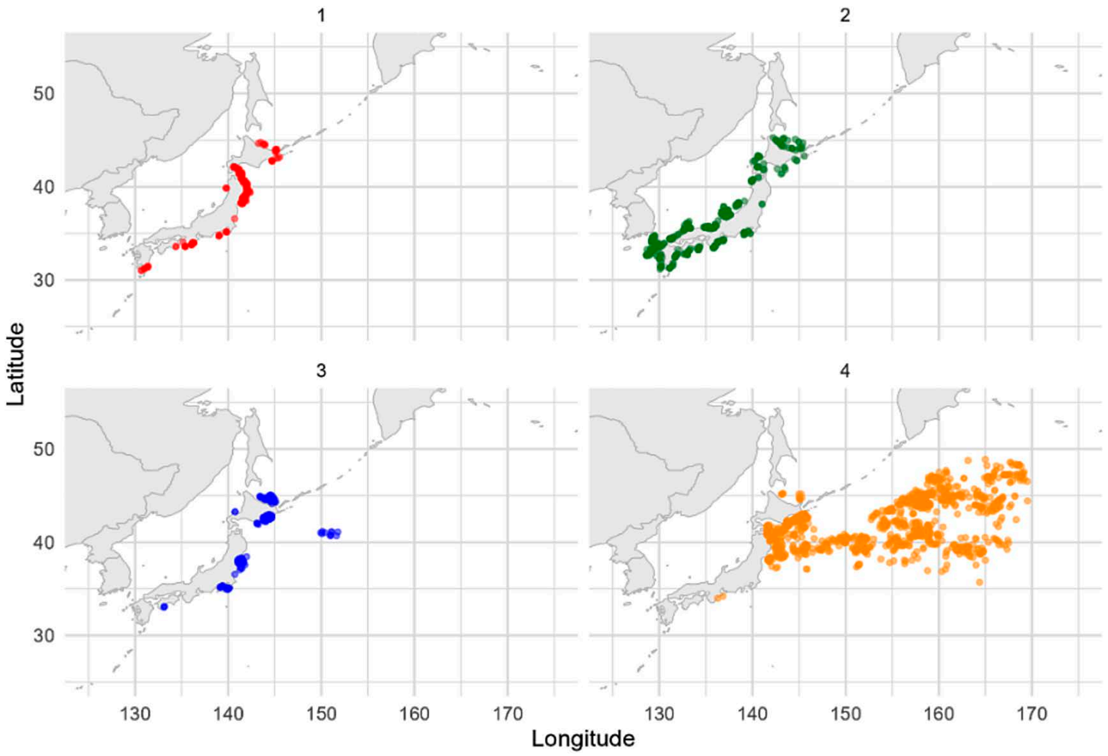

Figure 2. Geographic distributions of GENELAND clusters derived from the consensus *q* (average posterior assignment probability), defined as the average posterior assignment probability, across the three converged spatial baseline replicates.

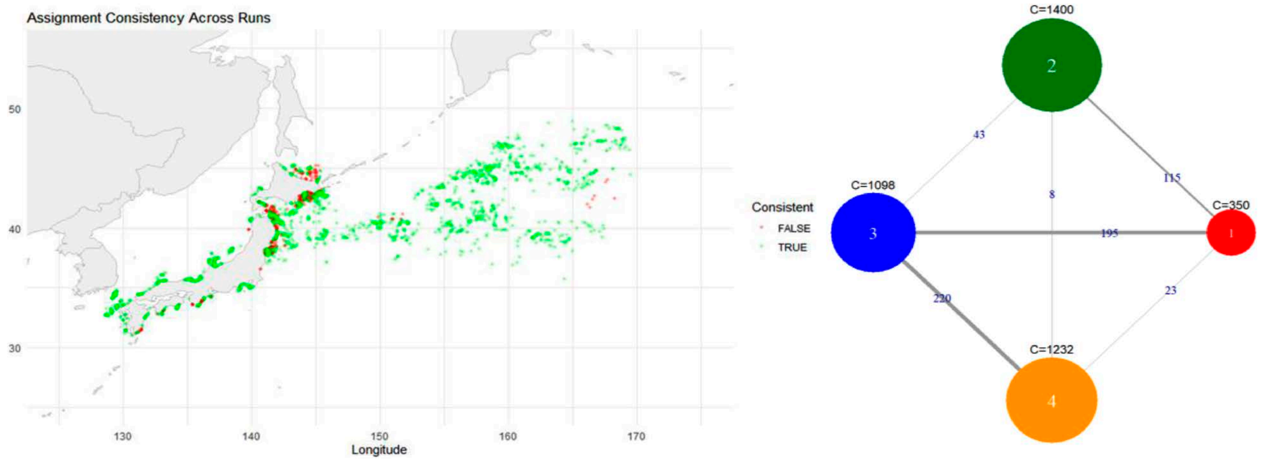

Figure 3. Cluster assignment agreement among the three converged spatial baseline replicates. In the network diagram, node indicates GENELAND clusters, and node size represents the number of consistently assigned to the respective cluster (C) across the replicates, and edge width and labels indicate the number of cross-assignments between clusters. Note that orange and green correspond O and J stocks, respectively.

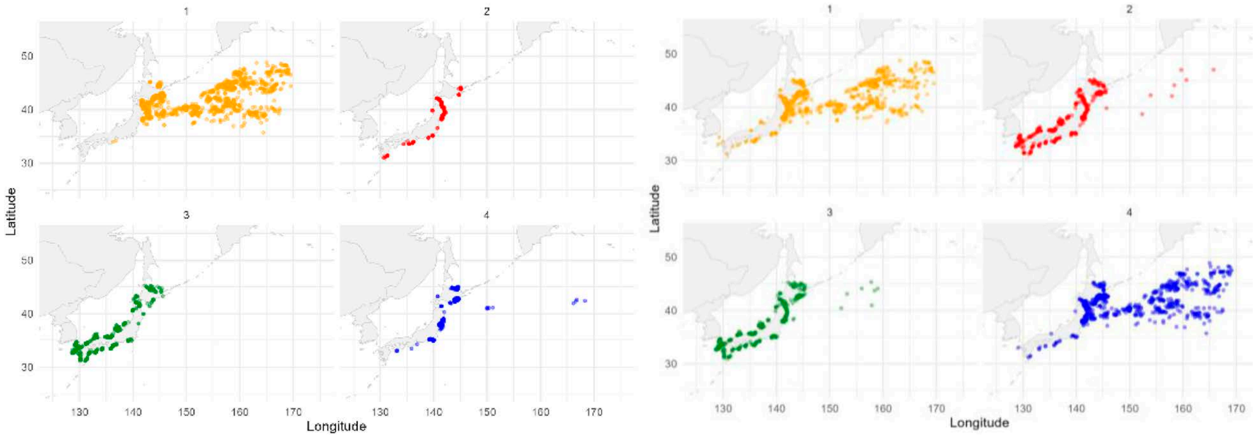

Figure 4. Geographic distributions of GENELAND clusters derived from the consensus *q* (average posterior assignment probability) across the converged spatial (SpFix4-10; left) and non-spatial (NoSpFix4-10; right) replicates, with K fixed at 4.

blue, orange) reported by de Jong and Hoelzel (2019), with the green and orange clusters corresponding to the J and O stocks, respectively.

Spatially structured disagreement in GENELAND cluster assignments

Figure 3 shows disagreements in cluster assignment across the three converged baseline replicates, which were concentrated in sub-areas 2C, 7CS, 7CN, and 11, where spatial mixtures of the J and O stocks are known to occur (hereafter referred to as the mixing zone). The network diagram showed that inconsistent assignments occurred mainly between specific clusters (i.e., between orange and blue, blue and red, or green and red). These results indicate that the observed disagreements are not random misassignments but rather reflect high uncertainty in the red and blue clusters.

Effect of spatial prior information on GENELAND clustering

Figure 4 shows geographic distribution patterns of clusters derived from spatial and non-spatial GENELAND variants when K was fixed at 4. Only the spatial model recovered red and blue clusters in the mixing zone, whereas the non-spatial model merged them into the J/O structure. Therefore, red and blue clusters appear to arise from spatial modeling assumptions rather than any strong genetic signal.

Over-partitioning of the J–O genetic axis

To examine whether the GENELAND K=4 clusters can be explained solely by the J–O genetic axis, Discriminant Analysis of Principal Components (DAPC; Jombart *et al.*, 2010) was performed (Figure 5). The red and blue clusters showed bimodal distributions along the first discriminant axis (DA1), each having a main peak overlapping with the J-like green and O-like orange clusters, respectively. This pattern likely resulted from over-partitioning of the continuous J/O mosaic structure, separating it into more coastal J-dominated (red) and more offshore O-dominated (blue) mixing zones.

No additional genetic structure beyond the J–O axis

To explore potential axes beyond the J–O genetic axis, a Factorial Correspondence Analysis (FCA) was conducted (Figure 6). Scree and scatter plots showed that Axis 1 explained the largest proportion of genetic variance and clearly separated the J-like green and O-like orange components. Red and blue clusters completely overlapped with the J-like green and O-like orange clusters, indicating no evidence of additional genetic structure beyond the main J–O axis.

NON-GENETIC ANALYSES

Various non-genetic data (e.g., body length, flipper coloration, conception date) were also analyzed to evaluate the four GENELAND clusters (Taguchi *et al.*, 2019). However, none of these analyses supported the inference that the red and blue clusters represent a resident stock in the mixing zone. Among them, body length distribution provided particularly important implications for stock assessment. Additionally, the trend in the number of sightings per unit effort in part of the mixing zone under the P stock scenario is being analyzed, and preliminary results indicate no drastic decline, which would be expected under a scenario of a coastal isolated stock.

Lack of mature females in the P stock

Figure 7 shows body length distribution among sub-area

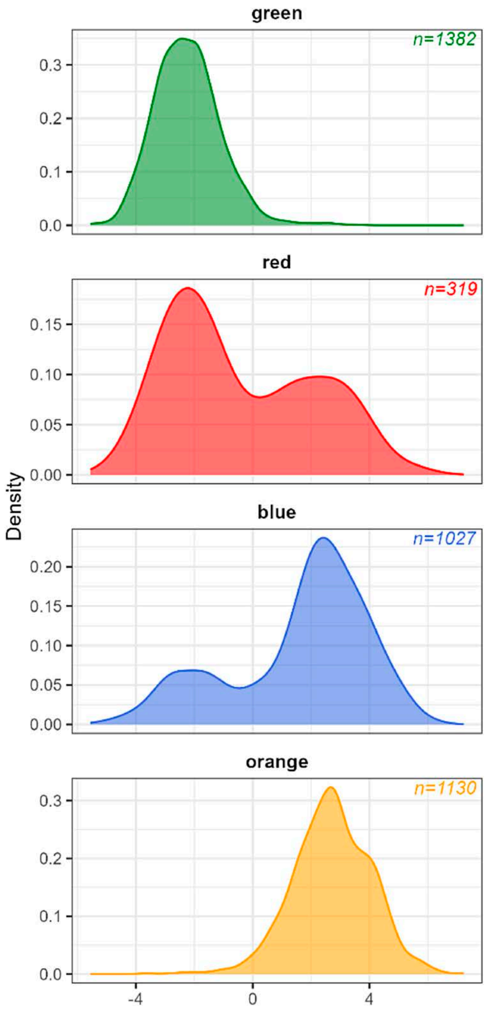

Figure 5. Density plot along the first discriminant axis (DA1) from DAPC. DAPC was trained using individuals from pure J (sub-areas 6E and 10E) and O (sub-areas 8 and 9) zones with consensus *q* (average posterior assignment probability)>0.9 across five STRUCTURE (K=2; Pritchard *et al.*, 2000) replicates. All individuals were then projected onto the STRUCTURE (K=2)-trained DA1 and colored according to the four GENELAND clusters, using assignments with consensus *q*>0.9 only.

by GENELAND cluster and sex. The red and blue clusters showed that most females did not reach the body length at sexual maturity. The absence of mature females in these clusters is biologically and demographically difficult to explain if the P-stock is assumed to be a resident stock. The J-like green cluster also consists mostly of immature whales; however, J stock whales are not considered resident in the study area. The apparent lack of immature males in the O-like orange cluster likely reflects misassignment of immature O stock whales distributed in the mixing area (Hatanaka and Miyashita, 1997) to the blue cluster.

CONCLUDING REMARKS

Evidence supporting the existence of the P stock under hypothesis E has been put forward from the GENELAND analysis only. However, the limited robustness of this inference indicates that even the GENELAND analysis provides only limited support for the P stock as a single, distinct stock. The software is based on a spatial Bayesian clustering model that assumes each individual belongs to one of K discrete populations and integrates genetic and geographic information to identify spatially coherent clusters. In the ICR's dataset, the genetic composition changes gradually along the spatial gradient from J-dominated to O-dominated waters (e.g., JJJJJ [pure J zone: sub-areas 1E, 6E and 10E] – JJJJO [coastal part of mixing zone] – JJOOO [offshore part of mixing zone] – OOOOO [pure O zone: sub-areas 8 and 9]). The spatial model in GENELAND tends to impose discrete population borders even across gradual genetic transitions, leading to over-partitioning within not only the mixing zones but also between the mixing and adjacent pure zones. This behavior likely explains the artificial genetic clusters (i.e., red and blue). In summary, these analyses do not support the existence of an additional coastal P stock along the

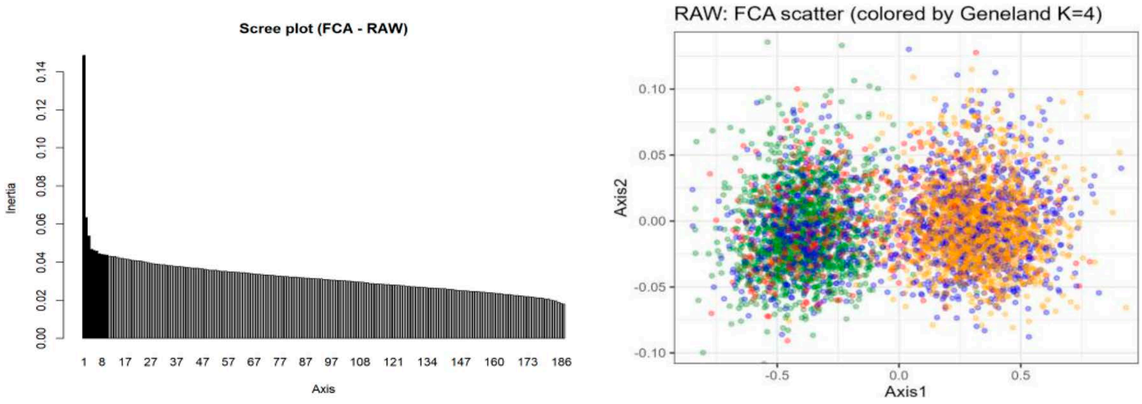

Figure 6. Scree plot and scatter plot of FCA analysis with GENELAND K=4 cluster colors.

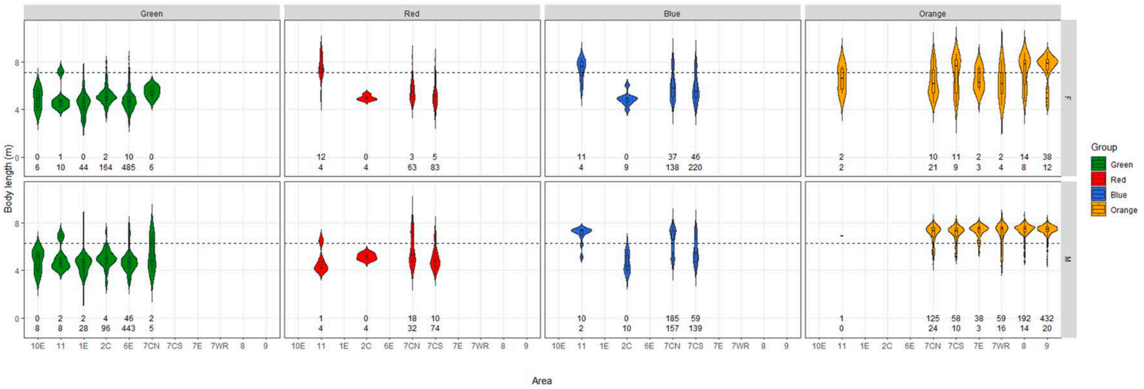

Figure 7. Body length distribution across sub-areas, grouped by sex and the four GENELAND clusters, using only individuals with consensus *q* (average posterior assignment probability) > 0.9. The dashed lines indicate the estimated body length at sexual maturity for western North Pacific common minke whales (7.1 m for females and 6.3 m for males; Kato, 1992). The values shown above the x-axis represent the number of mature (top) and immature (bottom) individuals, classified based on bodylength criteria.

Pacific coast of Japan and the southern coastal Okhotsk Sea (and not beyond those waters).

REFERENCES

Goto, M. and Pastene, L.A. 1997. Population structure of the western North Pacific minke whale based on an RFLP analysis of the mtDNA control region. *Rep. int. Whal. Commn* 47: 531–537.

Guillot, G., Mortier, F. and Estoup, A. 2005. GENELAND: A computer package for landscape genetics. *Mol. Ecol. Notes* 5(3): 712–715.

Hatanaka, H. and Miyashita, T. 1997. On the feeding migration of the Okhotsk Sea-West Pacific stock of minke whales, estimates based on length composition data. *Rep. int. Whal. Commn* 47: 557–564.

International Whaling Commission. 2020. Annex D: Report of the Sub-Committee on *Implementation Reviews* and *Simulation Trials*. *J. Cetacean Res. Manage.* 21 (*Suppl*.): 77–106.

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(1): 1–15.

de Jong, M. and Hoelzel, R.A. 2019. Collaborative analysis of WNP minke whale stock structure using Japanese microsatellite DNA database and spatially explicit population structure analyses. Paper SC/F19/WNPM/02 presented to the First Intersessional Workshop on the Implementation Review for western North Pacific minke whales, February 2019 (unpublished). 25 pp. [Available from the IWC Secretariat].

Kato, H. 1992. Body length, reproduction and stock separation of minke whales off northern Japan. *Rep. int. Whal. Commn* 42: 443–453.

Omura, H. and Sakiura, H. 1956. Studies on the little piked whale from the coast of Japan. *Sci. Reports Whales Res. Inst.* 11: 1–37.

Pastene, L.A., Goto, M., Taguchi, M. and Fujise, Y. 2022. Phylogeny and Population Genetic Structure of Minke Whales Worldwide: A Review of Recent Studies. *In: Marine Mammals*. IntechOpen. Available at: http://dx.doi.org/10.5772/intechopen.102675.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155(2): 945.

Taguchi, M., Goto, M. and Pastene, L.A. 2019. Genetic and non-genetic evidences suggest a low plausibility for western North Pacific common minke whale stock structure Hypothesis E. Paper SC/68A/SDDNA/02 presented to IWC Scientific Committee, May 2019 (unpublished). 22 pp. [Available from the IWC Secretariat].

Wada, S. and Numachi, K. 1991. Allozyme analyses of genetic differentiation among the popultions and species of the *Balaenoptera*. *Rep. int. Whal. Commn* 13(*Special Issue*): 125–154.