SC/F19/WNPM/04

Results of Discriminant Analysis of Principal Component (DAPC) and Spatial Analysis of Principal Component (sPCA) and implications for the stock structure of western North Pacific common minke whale

Mioko Taguchi, Mutsuo Goto and Luis A. Pastene

INTERNATIONAL
WHALING COMMISSION

Results of Discriminant Analysis of Principal Component (DAPC) and Spatial Analysis of Principal Component (sPCA) and implications for the stock structure of western North Pacific common minke whale

Mioko Taguchi, Mutsuo Goto and Luis A. Pastene

Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo 104-0055, Japan
Contact e-mail: taguchi@cetacean.jp

ABSTRACT

This document shows the results of Discriminant Analysis of Principal Component (DAPC) and Spatial Analysis of Principal Component (sPCA) conducted using microsatellite data (16 loci), to investigate stock structure in western North Pacific common minke whale. In particular, the analyses were performed to assess the plusibility of the stocks proposed under Hypothesis C of the previous RMP *Implementation Review* for western North Pacific common minke whale. The DAPC was performed forcing *K* to different number of clusters that simulated putative stocks under Hypothesis C (O_W, O_E, J_W and J_E). The spatial distribution of clusters was compared with the geographical distribution of the putative stocks as specified in the mixing matrices of Hypothesis C. Under this rational, the DAPC analyses were performed forcing K = 2: assuming only O and J stocks, K = 3: assuming O_W, O_E and J stocks or O, J_W and J_E stocks, and K = 4: assuming O_W, O_E, J_W and J_E stocks. The DAPC analyses at K = 2 clearly showed two clusters with distribution corresponding to the known distribution of J and O stocks. The analysis at K = 3 subdivided the O stock cluster into two sub-clusters, and the analyses at K = 4 subdivided the O and J stock clusters into two sub-clusters each. The spatial distribution patterns for clusters under K = 3 and K = 4 was not consistent with the hypothesized distribution pattern of the putative stocks under Hypothesis C. Furthermore, the mtDNA conventional F_{ST} analysis showed no significant differences among the O stock sub-clusters and among the J stock sub-clusters suggesting that the additional clusters were an artifact. Additionally, the temporal distribution patterns of each sub-cluster were examined based on the idea that different stocks should show different frequency of occurrence reflecting independent population dynamics. This analysis suggested temporal differences only associated with the known pattern of distribution of the J and O stocks. Taking all results of DAPC into account, it is unlikely not only that the O_w or J_E stocks exist but also that multiple stocks exist with overlapping geographic range. Results of the sPCA analyses were consistent with those of the DAPC analyses. In conclusion, the present DAPC and sPCA study showed no evidence of the existence of additional stocks other than O and J stocks, and therefore these analyses provided no support for Hypothesis C of the previous RMP *Implementation Review* for western North Pacific common minke whale.

INTRODUCTION

Three stock structure hypotheses were used during the last *Implementation Review* (*IR*) of western North Pacific common minke whale. Hypothesis A involved two stocks (O and J stocks); Hypothesis B involved three stocks (O, J and Y stocks) and Hypothesis C involved five stocks (J_E, J_W, O_E, O_W and Y) (IWC, 2013) (Appendix 1). It was not possible for the IWC SC to agree on the plausibility of the three stock structure hypotheses and as a consequence, these hypotheses were 'no agreement' and were therefore treated as if they had assigned 'Medium' plausibility in the trials (IWC, 2013).

The main disagreement was on whether or not a coastal O stock (O_W) and a coastal J stock (J_E) as specified in Hypothesis C exist. In 2012, a group of geneticists summarized their interpretations of the relative support for and against the five hypothesized stocks. They concluded that the genetic evidence for a J_E stock was low while the evidence for an O_W stock was moderate (IWC, 2013).

Goto *et al*. (2018) provided the last review of the genetic and non-genetic information accumulated since the last *IR*. Most of the analyses followed specific recommendations from the JARPNII review workshops, endorsed by the IWC SC. They concluded that all different analyses and results pointed out to a single O stock distributed from the Japanese coast till approximately 170°E.

In providing 'moderate' support for a putative Ow stock, the group of geneticists in 2012 noted that the 'PCA results using J-purged O stock sample provided support for an additional stock in O_W compared to O_E'. However, Gaggiotti and Gascuel (2011), whose conducted the original PCA analyses, noted that the apparent sub-structuring in O stock samples 'is not correlated with body length or geographic position so for the moment it has no clear biological explanation'. They did not speculate on whether such sub-structuring was related to possible stock structure. Subsequently, Pastene *et al*. (2016b) conducted a DAPC analysis (Jombart *et al.*, 2010) on J and unassigned samples-purged O stock samples and found no

evidence of additional O stock structure. Waples and Hoelzel (2017) noted that some combinations of principal components (PCs) shown in Pastene *et al*. (2016b) showed two or more essentially separated clusters, and provided interpretations for such results as follows, i) additional structure, and ii) age structure, iii) temporal changes.

Another issue discussed at the IWC SC was the 'unassigned' samples in the STRUCTURE analysis, with some authors suggesting that the unassigned whales could represent a different stock (Wade and Baker, 2017). Pastene *et al.* (2016a) showed that the 'unassigned' samples were widely distributed geographically and through the mtDNA phylogenetic clades, and concluded that the unassigned samples were not related to the occurrence of a different stock. Taguchi *et al.* (2017) showed that the number of unassigned samples decreased with the increase in the number of microsatellite loci, and concluded that they reflect lack of power in the analyses. A workshop on western North Pacific common minke whale stock structure held in 2018 discussed the treatment of unassigned samples in stock structure analyses. The workshop noted that the unassigned samples could be explained by assignment error, the presence of an additional stock, and/or admixture between J and O stocks. The workshop agreed with the use of non-purged data in the stock structure analyses for the present *IR* (IWC, 2019).

In recent meetings, the IWC SC has recommended at least four genetic analyses that could assist in determining the plausibility of different stock structure hypotheses: (1) F_{ST}, F_{IS}, heterozygosity, haplotype diversity and related measures, (2) PCA analyses including partitioning based on multiple components and DAPC, (3) Spatially explicit analyses (BAPS, TESS, Geneland and sPCA), and (4) Update kinship analyses including most recent samples.

This document presents the results of the analyses based on DAPC and sPCA. Given the past arguments on the existence of coastal stocks in the Pacific side of Japan, all analyses in this study were performed in terms of testing the existence of O_W and J_E stocks under the Hypothesis C (Appendix 1).

MATERIALS AND METHODS

Dataset

The 16 microsatellite DNA data from a total of 4,707 North Pacific common minke whales collected during 1994-2016 was used to perform the DAPC and sPCA analyses (Table 1). A similar sample set (n = 4,706) was used for mtDNA control region sequence analyses.

Data Analyses

DAPC

The DAPC in the R package *'adegenet'* was conducted based on the k-means clustering method, meaning without a priori group assignments. Taking into account the hypothesized stocks involved in the stock structure hypotheses (Appendix 1), the number of genetic clustering of individuals was forced to K = 2: O and J stocks under the Hypotheses A; K = 3: O_W, O_E and J stocks or O, J_W and J_E stocks under the Hypothesis C; and K = 4: O_W, O_E, J_W and J_E stocks under the Hypothesis C. If the O_W and/or J_E stocks exist, geographic clusters corresponding to the respective stock range specified in the mixing matrices of Hypothesis C (Table 2) should be observed in the analyses at K = 3 and K = 4. In this sense, we examined the spatial differences in frequency of occurrence of each cluster produced by the DAPC analyses at K = 2-4, which was assessed by the chi-squared test of independence.

The temporal difference in the frequency of occurrence of each cluster was also examined by the chi-square test of independence, based on the idea of 'two or more 'O'-like stocks that had strongly overlapping geographic ranges' (Waples and Hoelzel, 2017). If the idea of geographical overlapping of O-like stocks is correct, then the frequency should be different among seasons and/or years reflecting the difference in population dynamics among stocks with different biological properties, *i.e.*, migration pattern and reproductive season.

To facilitate our understanding of the results by DAPC, the composition of DAPC clusters in each STRUCTURE (Pritchard *et al.,* 2000) assignment was also examined. In order to review if unassigned samples are a signal of additional stocks, we also examined the correspondence between clusters and unassigned samples, and this was assessed by the test of goodness of fit.

Mitochondrial DNA variations

To verify an inference by the present DAPC analyses, the haplotype diversity (h) was estimated in each cluster produced by DAPC and conventional pairwise F_{ST} estimates were calculated between the clusters using the mtDNA sequence data. If the DAPC analysis appropriately clustered individuals, h in each cluster should be characteristic and increase when sub-clusters in the DAPC analyses at K = 3-4 are combined. Furthermore, if this is the case, pairwise F_{ST} estimates should show significant differentiation not only between main clusters but also between sub-clusters at K = 3-4. These analyses were performed using the Arlequin ver. 3.01 (Excoffier *et al.*, 2007).

sPCA

A spatially explicit multivariate method, sPCA (Jombart *et al.*, 2008) was performed to investigate the spatial pattern of genetic variability using microsatellite allele frequency data of individuals in the R package *'adegenet'*. This method finds synthetic variables that optimize the product of allele frequency variance and the spatial autocorrelation as measured by Moran's *I* (Moran 1950), which offers advantages in that it does not rely on assumptions of Hardy–Weinberg equilibrium with linkage equilibrium between loci, and can reveal genetic clines as well as discrete populations (Jombart *et al.*, 2008).

The eigenvalues provided by the sPCA are highly positive (global) when the synthetic variables have a large variance and exhibit positive autocorrelation; and conversely, sPCA eigenvalues are largely negative (local) when the spatial PCs have a high variance and display negative autocorrelation. The lagged scores for each of the first PCs produced by the sPCA were plotted to the geographic space to identify a spatial genetic structure. In this analysis, the latitude and longitude coordinates were converted to the UTM coordinates, and 150 m of jitter was added to the UTM coordinates since multiple samples were at times collected from the same location, to produce a Delaunay triangulation network. A Monte Carlo simulation-based test (Jombart *et al.*, 2008) was also performed to reinforce the presence of global (neighboring individuals are more similar than expected) and/or local (neighboring individuals are more dissimilar than expected) structures with 10,000 permutations when the structures were inferred in the sPCA. In order to improve our understanding of the result by sPCA, the composition of sPCA clusters in each STRUCTURE assignment was examined.

RESULTS

DAPC

Spatial analysis

At $K = 2$, the DAPC analysis clearly divided the samples into two clusters (Cl-1 and Cl-2) (Figures 1a) and these clusters corresponded well with J and O stocks assigned by the program STRUCTURE (Figure 2a). The geographical distribution of the clusters corresponded well with the stock range specified in the mixing matrices under the Hypotheses A. The frequency of occurrence of each cluster was significantly different among sub-areas (Figure 3a and Table 3), which would suggest geographical differences in occurrence of J and O stock whales.

At $K = 3$, the DAPC analysis subdivided the O stock cluster (Cl-2) into two sub-clusters (Cl-2a, Cl-2b) (Figures 1b and 2b), which were observed across all sub-areas (Figure 3b) in equal proportion (Table 3), even in putative pure stock region, *i.e.*, sub-areas 7E, 8 and 9. These observations suggested not only that the geographical distribution of the O stock sub-clusters contradict the O_W and O_E ranges in the mixing matrices under the Hypotheses C (Table 2), but also that the observed O stock sub-clusters did not make sense biologically.

At $K = 4$, the DAPC analysis subdivided the J stock cluster (Cl-1) into two sub-clusters (Cl-1a, Cl-1b) (Figures 1c and 2c), and the O stock cluster (Cl-2) into two sub-clusters (Cl-2a, Cl-2b) (Figures 1c and 2c). All four sub-clusters were observed across all sub-areas (Figure 3c). The J stock sub-clusters distributed geographically in equal proportion as well as O stock sub-clusters (Table 3), even in putative pure stock region, *i.e.*, sub-areas 6E, and 7E, 8 and 9, respectively. These observations suggested not only that the geographical distribution of the each of O and J stock sub-clusters contradict the O_W and O_E ranges and J_W and J_E ranges in the mixing matrices under the Hypotheses C (Table 2), but also that the observed sub-clusters did not make sense biologically.

Temporal analysis

The temporal heterogeneity tests showed annual (Table 3 and Figures 4a, 4c and 4e) and monthly (Table 3 and Figures 4b, 4d and 4f) differences in occurrence of main clusters (Cl-1 and Cl-2) consisting mainly of J and O stocks, respectively but not between sub-clusters (between Cl-2a and Cl-2b and between Cl-1a and Cl-1b) (Table 3 and Figures 4c-4f). As summarized in Table 3, these results can be interpreted only as temporal differences in appearance of J and O stocks.

Unassigned samples

The unassigned samples were allocated to the DAPC clusters in equal proportion at $K = 2$ (Figure 2a, $P = 0.846$), $K = 3$ (Figure 2b, $P = 0.298$) and $K = 4$ (Figure 2c, $P = 0.127$).

Mitochondrial DNA analysis

The *h* in the clusters consisting mainly of J stock whales (Cl-1, Cl-1a, Cl-1b) ranged from 0.874 to 0.889 (Figure 5), which were comparable to the *h* of J stock (0.87, Park *et al.*, 2006). The *h* in the clusters consisting mainly of O stock whales (Cl-2, Cl-2a, Cl-2b) ranged from 0.941 to 0.944 (Figure 5), which were comparable to the *h* of O stock (0.95, Pastene *et al.*, 2016b). In addition, *h* did not increase even when the sub-clusters observed in each of the main two clusters were combined: Cl-2a plus Cl-2b at $K = 3$ (0.943); Cl-2a plus Cl-2b at $K = 4$ (0.943); and Cl-1a plus Cl-1b at $K = 4$ (0.886) (Figure 5).

These results did not show any signals of differentiation between sub-clusters in each of the main clusters produced by the DAPC analyses at K = 3-4.

Table 4 shows the conventional F_{ST} estimates between DAPC clusters at each value of K. In the analyses at K = 2, significant genetic differentiation was observed between Cl-1 *vs.* Cl-2 (Table 4a). In the analyses at K = 3, significant differentiation was observed between Cl-2a *vs.* Cl-1 and between Cl-2b *vs.* Cl-1. No significant differentiation was found between Cl-2a *vs.* Cl-2b (Table 4b). In the analyses at K = 4, significant genetic differentiation was observed between Cl-1a *vs.* Cl-2a; Cl-1a *vs.* Cl-2b; Cl-1b *vs.* Cl-2a; and Cl-1b *vs.* Cl-2b (Table 4c). No significant differentiation was found between Cl-1a *vs.* Cl-1b; and Cl-2a *vs.* Cl-2b (Table 4c).

These results suggested that the sub-clusters in each of the main DAPC clusters at K = 3 (Cl-2a/Cl-2b) and K = 4 (Cl-1a/Cl-1b, Cl-2a/Cl-2b) were artifact clusters.

sPCA

The sPCA revealed that the first global score (λ_1) was the largest eigenvalue in terms of variance and of spatial autocorrelation and could be clearly distinguished from all the other eigenvalues (Figure 6). This result suggested that only the global structure, associated to λ_1, should be interpreted. The Monte-Carlo test confirmed the existence of at least one global pattern ($P > 0.001$).

The lagged scores of the first PC was thus plotted to the geographical coordinates, which showed that the majority of samples in sub-areas 7E, 8 and 9 were clearly differentiated from the samples in sub-areas 6E (Figures 7a and 8). This suggests that the first global score separated clearly the J and O stocks but the analysis provided no evidence of any additional stocks. In sub-areas 7CN, 7CS and 11, both clusters were observed (Figures 7a-c), and many of them were less differentiated (Figure 8).

Percentage of matching of clusters between STRUCTURE and DAPC and between STRUCTURE and sPCA was shown in Table 5. The STRUCTURE/sPCA matching proportion (80.19% for J-stock and 84.75 % for O-stock) was lower than the STRUCTURE/DAPC matching (97.39% for J-stock and 95.94 % for O-stock).

DISCUSSION

Implications for the stock structure of the DAPC analyses

Biological validity of clusters

The only result of the DAPC with biological validity was that corresponding to K = 2. These two clusters corresponded clearly to the J and O stocks, and were consistent with the pattern of distribution of these two stocks in previous studies (Hatanaka and Miyashita, 1997), and with the spatial distribution pattern specified in the mixing matrices for Hypothesis A. When the program was forced to K = 3 and K = 4, simulating the scenario of additional stocks under Hypothesis C, J stock sub-clusters and O stock sub-clusters behave in a similar manner as the main clusters of the J and O stocks, respectively. Statistical analyses using an independent genetic marker (mtDNA) confirmed these patterns: clearly differentiation between O and J stock clusters but no differentiation among J stock sub-clusters or O stock sub-clusters. This general pattern was also confirmed by mtDNA diversities. Therefore, the extra clusters representing the putative stocks of Hypothesis C have no biological validity.

Clustering methods can produce artificial clusters. This could be the case of the results of Gaggiotti and Gascuel (2011) and Pastene *et al.* (2016b), whose found some clusters within the O stock samples in the PCA analyses. Waples and Hoelzel (2017) suggested that such clusters could be explained by additional stock structure e.g. two or more 'O'-like stocks that had strongly overlapping geographic ranges, age structure within a panmictic population, temporal differentiation or sampling bias. For the first explanation for additional stocks with mixing, it is expected that temporal heterogeneity tests will show differences in frequency of the occurrence between sub-clusters, owing to differences in population dynamics among stocks. The heterogeneity test in the present study did not show any temporal differences between sub-clusters, although the temporal differences in the occurrence between the two main clusters (J and O stocks) were suggested. Therefore, the clustering pattern of the previous PCA analyses could be explained by other reasons different from additional stock structure.

Interpretation of unassigned samples

Wade and Baker (2017) suggested the possibility that the unassigned samples represent an additional stock. The present DAPC analysis was performed forcing not only K = 2 but also K = 3-4, with non-purged data. If the unassigned samples are suggestive of any additional stocks as Wade and Baker (2017) pointed out, they are more likely to belong to particular clusters produced by the analyses at K = 3-4. However, the unassigned samples were assigned to all DAPC clusters in equal proportion at any of K, which means that unassigned samples are not a signal of additional stocks.

Implications for the stock structure of the sPCA analyses

The sPCA analysis clearly separated J and O stocks but provided no evidence for additional stocks. This analysis also indicated that many samples in sub-areas 7CN, 7CS and 11 were less differentiated. Considering the sPCA uses a spatial autocorrelation to infer a spatial structure, it is highly possible that geographical components dilute a resolution of stock assignment by sPCA in the mixing areas. This could be also explained with the following observations: (1) STRUCTURE/sPCA matching proportion was lower than the STRUCTURE/DAPC matching despite the sPCA inferred only J and O stock differences as in the case with STRUCTURE and DAPC, and (2) the mismatches in the STRUCTURE/sPCA mainly occurred in the mixing sub-areas. If this is the case, temporal changes in the mixing proportion of J and O stock whales would reduce further the assignment resolution.

CONCLUSIONS

The present DAPC analyses supported the occurrence of the J and O stocks but provided no evidence for additional structure as suggested by Hypothesis C. This inference was supported by the mtDNA analyses of DAPC clusters. The sPCA also separated clearly J and O stocks and provided no evidence for additional structure. However, this approach has difficulty in assigning sample to stocks in areas with strong spatial and temporal pattern of mixing.

REFERENCES

Excoffier, L., Laval, G. and Schneider, S. 2007. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47-50.

Gaggiotti, O. and Gascuel, F. 2011. Stock structure of North Pacific minke whales as revealed by Principal Component Analyses of microsatellite data. Paper SC/63/RMP23 presented to the IWC Scientific Committee, June 2011 (unpublished). 7pp.

Goto, M., Taguchi, M. and Pastene, L.A. 2018. Recent genetic analyses following recommendations from the IWC Scientific Committee confirm no sub-structure of the 'O stock' common minke whale in the western North Pacific. Paper SC/F18/MI/01 presented to the Workshop on WNP common minke whale stock structure in preparation for the start of the *Implementation Review* in April 2018, February 2018 (unpublished). 16pp.

International Whaling Commission. 2013. Report of the Scientific Committee, Annex D1. Report of the Working Group on the *Implementation Review* for Western North Pacific Common Minke Whales. *J. Cetacean Res. Manage. 14 (suppl.)*:118-136.

International Whaling Commission. 2014. Report of the Scientific Committee, Annex D1. Report of the Working Group on the *Implementation Review* for Western North Pacific Common Minke Whales. Appendix 2. North Pacific minke whale Implementation Simulation Trial Specifications. Adjunct 2. Mixing matrices. *J. Cetacean Res. Manage. 15 (suppl.)*:159-166.

International Whaling Commission. 2019. Report of the Scientific Committee, Annex xx. *J. Cetacean Res. Manage. 19 (suppl.)*:xxx-xxx.

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.

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

Jombart, T., Devillard, S., Dufour, A.B. and Pontier, D. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101:92-103.

Moran, P.A.P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.

Park, J.Y., Goto, M., Kanda, N., Sohn, H., Kim, Z.G. and Pastene, L.A. 2006. Preliminary mitochondrial DNA control region sequencing analysis of by-caught J-stock minke whale from Korea and Japan. Paper SC/58/NPM3 presented to the IWC Scientific Committee, May-June 2006 (unpublished). 7pp.

Pastene, L.A., Goto, M., Taguchi, M. and Kitakado, T. 2016a. Temporal and spatial distribution of the 'J' and 'O' stocks of common minke whale in waters around Japan based on microsatellite DNA. Paper SC/F16/JR38 presented to the JARPNII special permit expert panel review workshop, February 2016 (unpublished). 14pp.

Pastene, L.A., Goto, M., Taguchi, M. and Kitakado, T. 2016b. Updated genetic analyses based on mitochondrial and microsatellite DNA indicated no sub-structure of the 'O' stock common minke whale in the western North Pacific. Paper SC/F16/JR40 presented to the JARPNII special permit expert panel review workshop, February 2016 (unpublished). 19pp.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

Taguchi, M., Goto, M. and Pastene, L.A. 2017. A synthesis of the work conducted on stock structure of western North Pacific common minke whale in response to recommendations from the IWC Scientific Committee. Paper SC/67a/SDDNA05 presented to the IWC Scientific Committee, May 2017 (unpublished). 5pp.

Wade, P. and Baker, S.C. 2017. Response to SC/F16/JR40 regarding new information on North Pacific minke whale population structure. Appendix 4. *J. Cetacean Res. Manage*. 18 (*suppl.*): 273-274.
Waples, R. and Hoelzel, R. 2017. North Pacific minke whale stock structure redux. Appendix 5. *J. Cetacean Res. Manage*. 18 (*suppl.*): 275-276.

Table 1. Monthly (upper) and annual (lower) sample size of microsatellite DNA data by sub-area, used in this study.

Month	1E	2C	6E	7CN	7CS	7WR	7E	8	9	10E	11	Total
Jan	7	73	144	5	29							258
Feb	6	45	75	1	7							134
Mar	14	30	103	6	14							167
Apr	12	23	192	23	246							496
May	9	23	131	74	444	51	36	42	40		9	859
Jun	13	12	47	143	104	33	11	102	89	1	11	566
Jul	6	14	42	77	36	4	2	95	194		52	522
Aug	2	8	46	30	11	1		12	209	1	30	350
Sep		1	23	570	5			1	9		4	613
Oct	2	11	23	252	6					3	10	307
Nov	1	36	61	15	28					10	13	164
Dec	4	96	115	20	33					3		271
Total	76	372	1002	1216	963	89	49	252	541	18	129	4707

Year	1E	2C	6E	7CN	7CS	7WR	7E	8	9	10E	11	Total
1994									20			20
1995									99			99
1996				30		1		16			30	77
1997						1	1	31	66			99
1998						25	29	45				99
1999				46	2	2					50	100
2000				19	5				16			40
2001	1	10	25	13	20	21	7	21	28	4	3	153
2002	7	19	45	122	17	1		8	32	3	5	259
2003	5	17	61	15	80	5	7	38	38		7	273
2004	4	19	66	82	14				83		3	271
2005	4	33	54	95	82	2		14	49	3	6	342
2006	3	28	75	61	95	12	2	38	24		3	341
2007	7	41	69	96	115	6		15	6		6	361
2008	9	22	66	61	75			5	54	2	3	297
2009	3	16	68	67	85	8	3	18	7		1	276
2010	3	17	73	67	62				14		4	240
2011	6	22	48	109	29				2		1	217
2012	4	25	55	99	97	5		3			4	292
2013	5	20	53	66	49				3	2		198
2014	3	21	74	67	53					1	2	221
2015	5	28	84	63	45						1	226
2016	7	34	86	38	38					3		206
Total	76	372	1002	1216	963	89	49	252	541	18	129	4707

Table 2. Summary of geographic clusters that should be observed in DAPC analyses if the O_W and/or J_E stocks exist, which was postulated from the putative range of each stock represented in the mixing matrices under the Hypothesis C (IWC, 2014).

K	Stock	Putative distribution area
2	J	1E/2C/6E/7CS/7CN/10E/11
	O	2C/7CS/7CN/7WR/7E/8/9/11
3	J_W	1E/6E/10E/11
	J_E	2C/7CS/7CN
	O	2C/7CS/7CN/7WR/7E/8/9/11
4	J	1E/2C/6E/7CS/7CN/10E/11
	O_W	2C/7CS/7CN/7WR/11
	O_E	7WR/7E/8/9/11
	J_W	1E/6E/10E/11
	J_E	2C/7CS/7CN
	O_W	2C/7CS/7CN/7WR/11
	O_E	7WR/7E/8/9/11

Table 3. Summary of results of heterogeneity tests for DAPC analyses.

	Heterogeneity test	*P*-value	Significance	Interpretation
Spatial (Sub-area)	*K* = 2 between Cl-1 and Cl-2	1.00E-04	***	Spatial differences in occurrence of J and O
	K = 3 among all clusters	1.00E-04	***	Spatial differences in occurrence of J and O
	K = 3 between Cl-2a and Cl-2b	6.15E-01	ns	No differences between clusters within O
	K = 4 among all clusters	1.00E-04	***	Spatial differences in occurrence of J and O
	K = 4 between Cl-2a and Cl-2b	2.29E-01	ns	No differences between clusters within O
	K = 4 between Cl-1a and Cl-1b	8.29E-01	ns	No differences between clusters within J
Temporal (Year)	*K* = 2 between Cl-1 and Cl-2	1.81E-57	***	Annual differences in occurrence of J and O
	K = 3 among all clusters	2.99E-60	***	Annual differences in occurrence of J and O
	K = 3 between Cl-2a and Cl-2b	9.85E-01	ns	No differences between clusters within O
	K = 4 among all clusters	1.38E-55	***	Annual differences in occurrence of J and O
	K = 4 between Cl-2a and Cl-2b	2.96E-01	ns	No differences between clusters within O
	K = 4 between Cl-1a and Cl-1b	3.45E-01	ns	No differences between clusters within J
Temporal (Month)	*K* = 2 between Cl-1 and Cl-2	2.36E-169	***	Seasonal differences in occurrence of J and O
	K = 3 among all clusters	2.43E-177	***	Seasonal differences in occurrence of J and O
	K = 3 between Cl-2a and Cl-2b	4.56E-01	ns	No differences between clusters within O
	K = 4 among all clusters	2.93E-160	***	Seasonal differences in occurrence of J and O
	K = 4 between Cl-2a and Cl-2b	7.01E-01	ns	No differences between clusters within O
	K = 4 between Cl-1a and Cl-1b	8.84E-01	ns	No differences between clusters within J

Table 4. Conventional pairwise F_{ST} estimates based on mtDNA between clusters estimated by DAPC: (a) $K = 2$, (b) $K = 3$ and (c) $K = 4$. Bold text indicates statistical significance after FDR correction.

(a) K = 2	Cl−1	Cl−2
Cl−1		
Cl−2	**0.0735**	

(b) K = 3	Cl−2a	Cl−2b	Cl−1	Cl−2a+ Cl−2b
Cl−2a				
Cl−2b	−0.0001			
Cl−1	**0.0828**	**0.0811**		
Cl−2a+ Cl−2b	−0.0004	−0.0004	**0.0811**	

(c) K = 4	Cl−2a	Cl−2b	Cl−1a	Cl−1b	Cl−2a+ Cl−2b	Cl−1a+ Cl−1b
Cl−2a						
Cl−2b	−0.0001					
Cl−1a	**0.0688**	**0.0736**				
Cl−1b	**0.0704**	**0.0753**	−0.0004			
Cl−2a+ Cl−2b	−0.0004	−0.0004	**0.0706**	**0.0722**		
Cl−1a+ Cl−1b	**0.0704**	**0.0753**	−0.0006	−0.0006	**0.0722**	

Table 5. Percentage of matching of clusters between approaches, *i.e.*, STRUCTURE/DAPC and STRUCTURE/sPCA.

		STRUCTURE		
		J	O	Unasg
DAPC	Cl-1	97.39	4.06	50.47
(K=2)	Cl-2	2.61	95.94	49.53
sPCA	Positive	80.19	15.25	42.89
	Negative	19.81	84.75	57.11

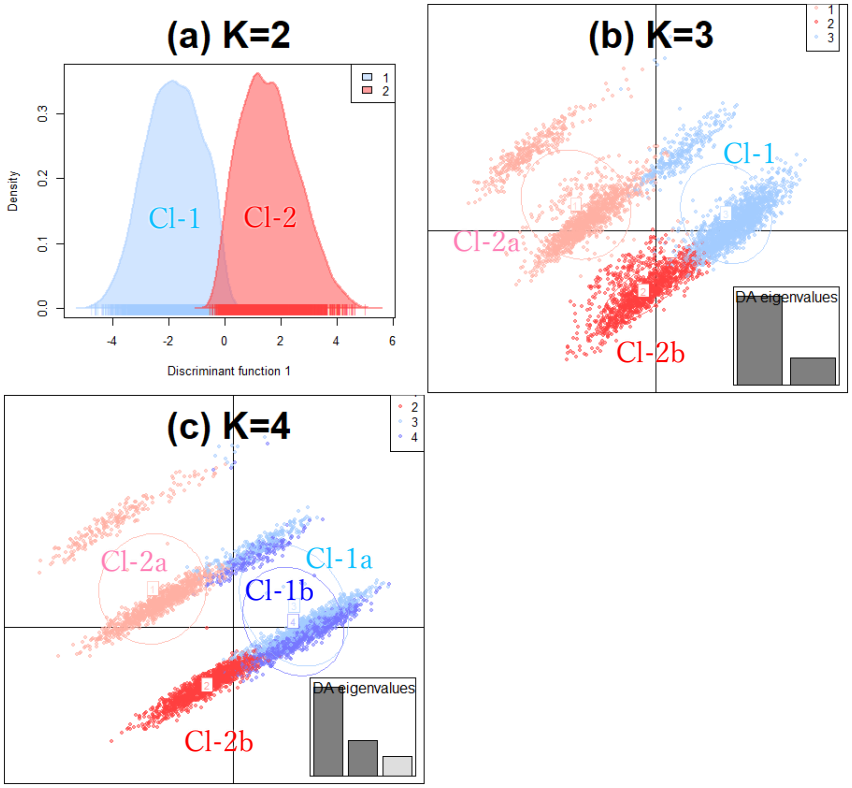


Figure 1. Density plot showing the first PC of the DAPC at K=2 (a), and scatter plots showing the first and second PCs of the DAPC at K=3 (b) and K = 4 (c).

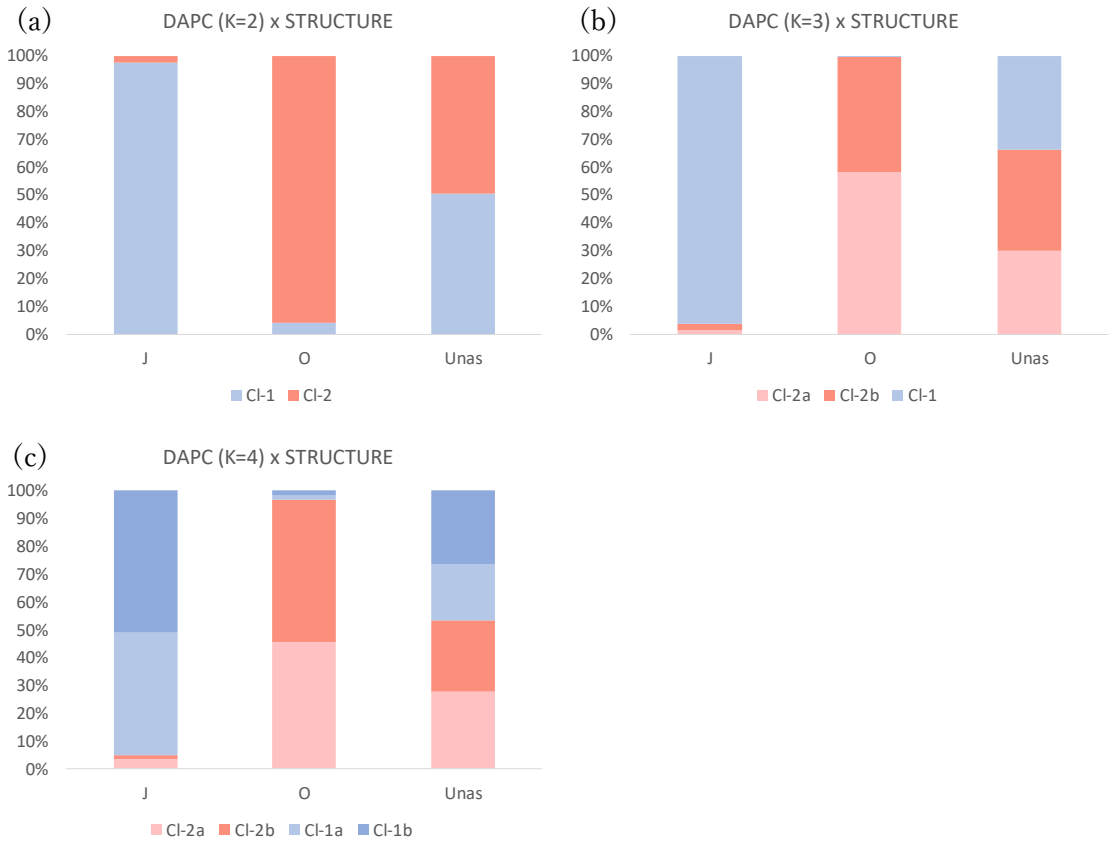


Figure 2. Percentage of composition of DAPC clusters in each STRUCTURE assignment: (a) K = 2, (b) K = 3 and (c) K = 4.

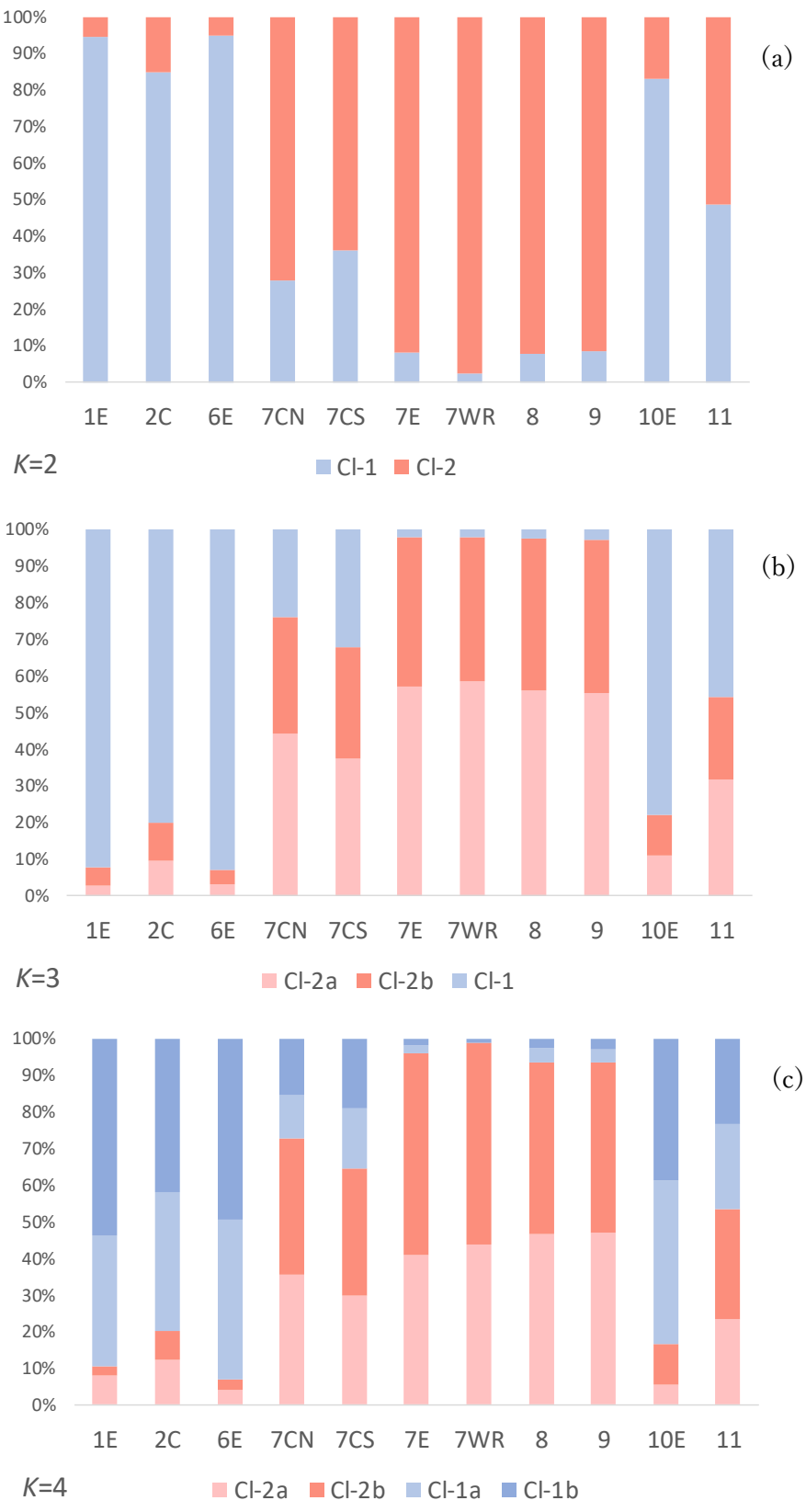


Figure 3. Spatial change in composition of DAPC clusters: (a) $K = 2$, (b) $K = 3$ and (c) $K = 4$.

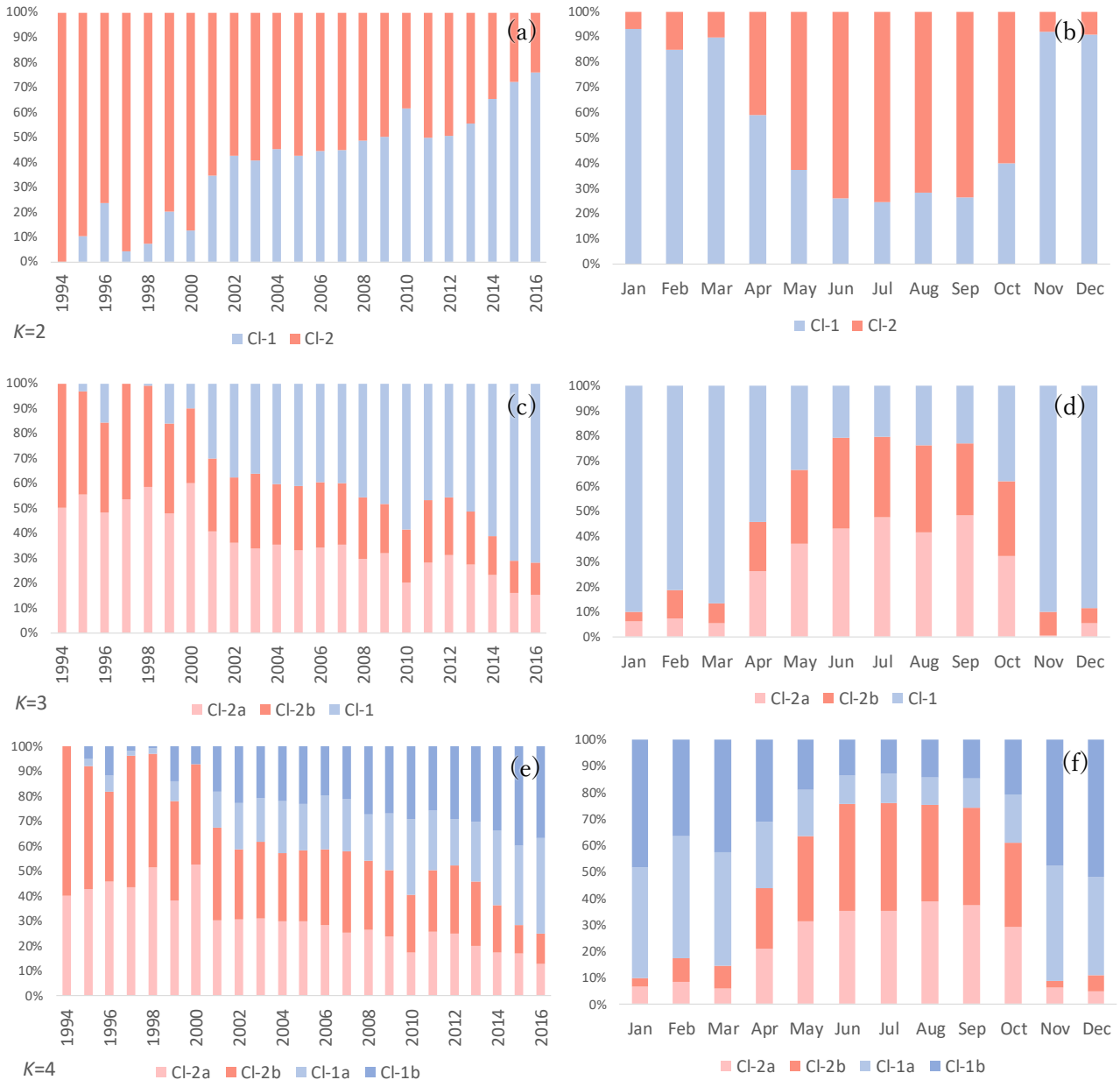


Figure 4. Annual (left panels) and monthly (right panels) changes in composition of DAPC clusters: (a)-(b) $K = 2$, (c)-(d) $K = 3$, and (e)-(f) $K = 4$.

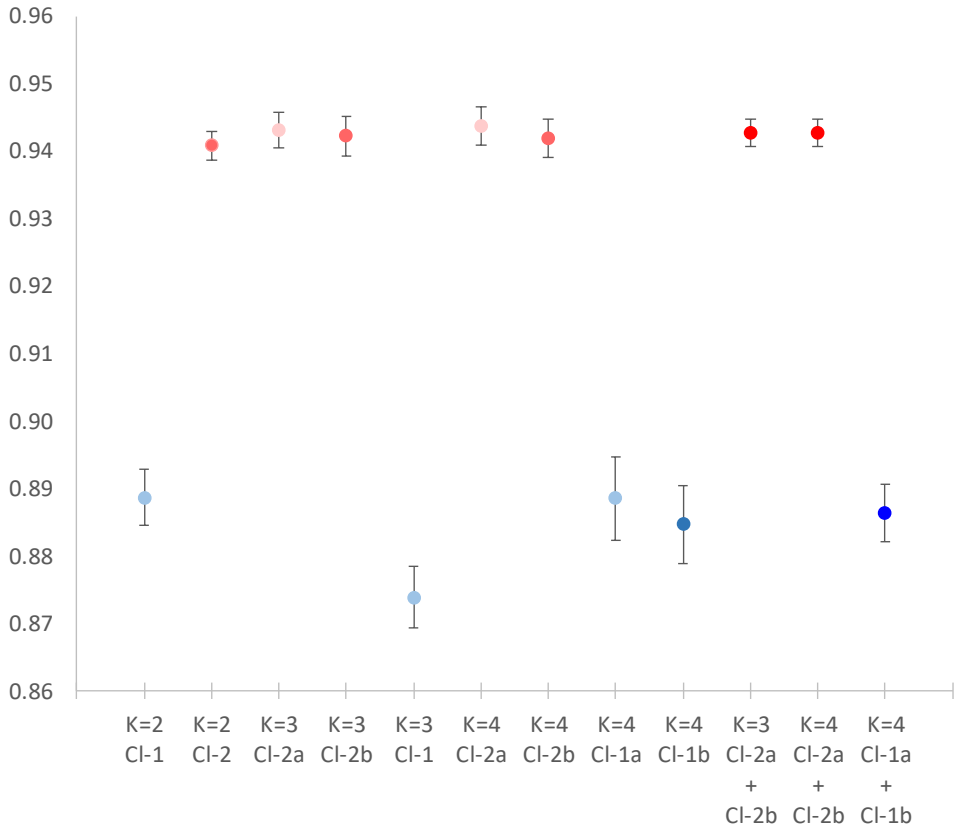


Figure 5. Haplotype diversity in each DAPC cluster. Error bar indicates sample standard deviations.

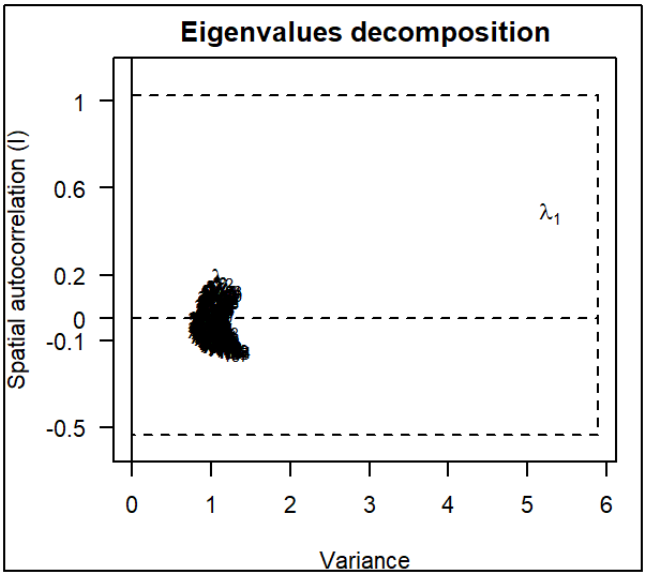


Figure 6. sPCA eigenvalues with genetic variance on the x-axis and spatial autocorrelation (Moran's *I*) on the y-axis. The positive and negative scores on the y-axis are referred to as global and negative scores in the text, respectively. The first global score (λ_1) having the highest value for variance and spatial autocorrelation can be easily distinguished from the other scores. The λ_1 can be interpreted here as a distinct population structure, while the local ones do not show any obvious feature.

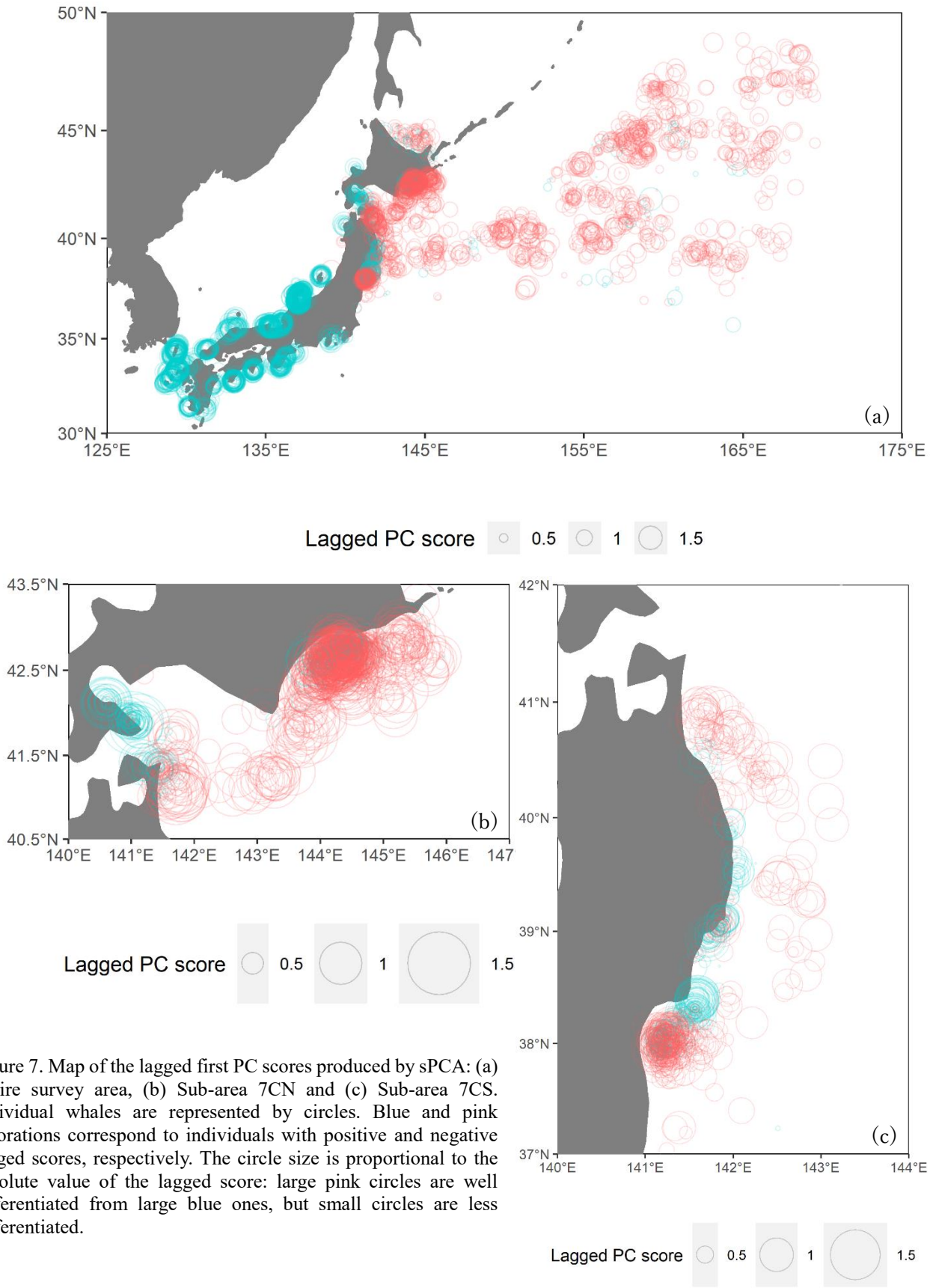


Figure 7. Map of the lagged first PC scores produced by sPCA: (a) Entire survey area, (b) Sub-area 7CN and (c) Sub-area 7CS. Individual whales are represented by circles. Blue and pink colorations correspond to individuals with positive and negative lagged scores, respectively. The circle size is proportional to the absolute value of the lagged score: large pink circles are well differentiated from large blue ones, but small circles are less differentiated.

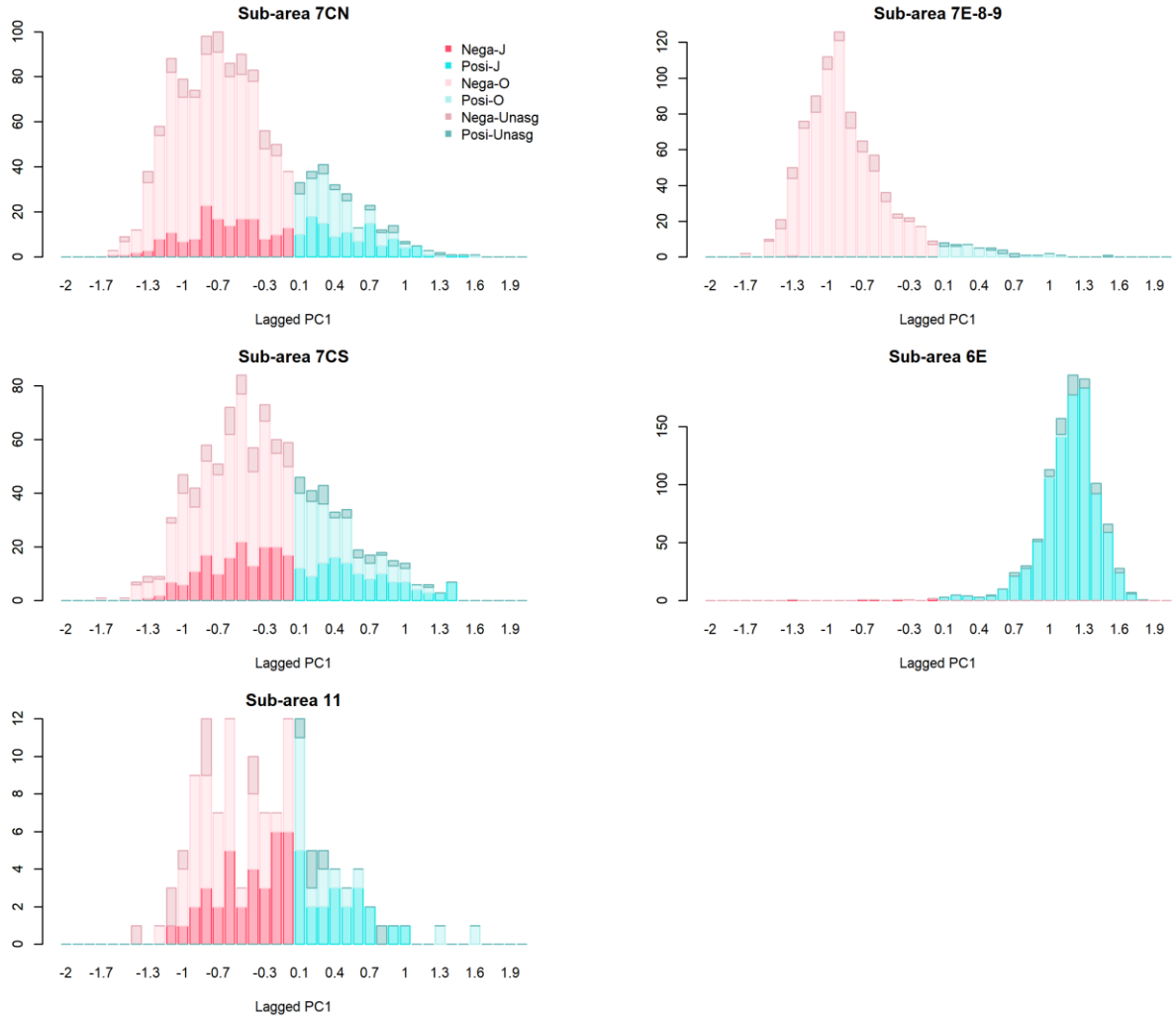


Figure 8. Frequency of the lagged first PC score produced by sPCA in each of mixing (left panels) and pure (right panels) stock areas. Pink and blue colorations indicate negative and positive clusters by sPCA, respectively. Shade of the color refers STRUCTURE assignment (J, O and Unassigned).

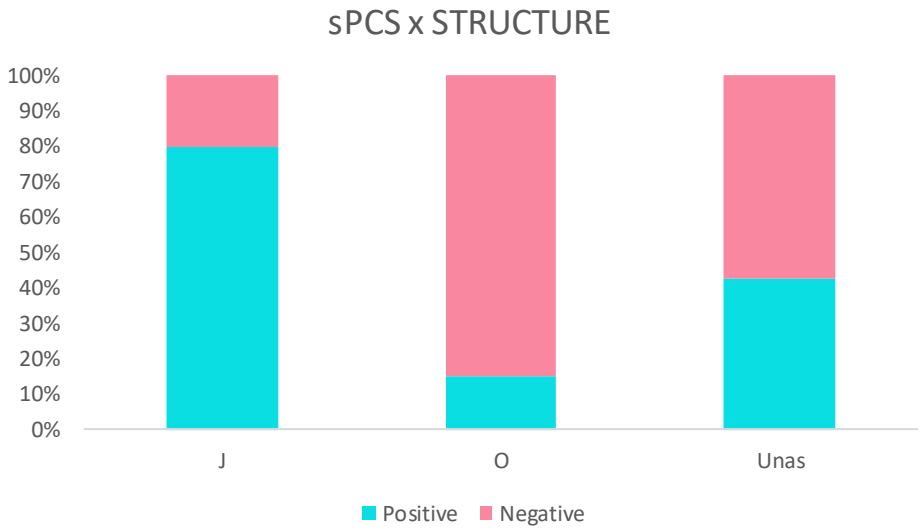
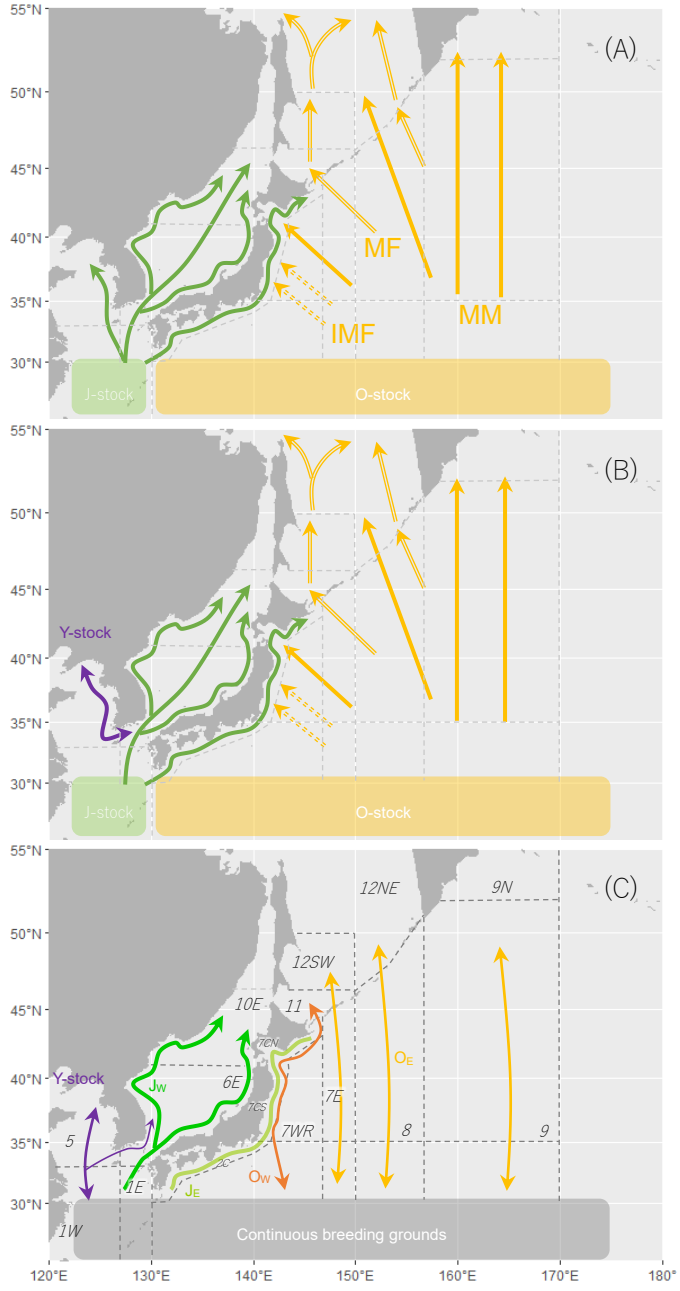


Figure 9. Percentage of composition of sPCA clusters in each STRUCTURE assignment.

Appendix 1

Stock structure hypotheses for western North Pacific common minke whale used in the RMP *Implementation Review* in 2013.



Hypothesis A: 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. The O stock migrates in summer mainly to the Okhotsk Sea (sub-areas 12SW and 12NE). Both J and O stocks overlap temporally along the Pacific coast (sub-areas 7CS and 7CN) and the southern part of the Okhotsk Sea (sub-areas 11 and 12SW); Hypothesis B: same as for hypothesis A, but a different stock (Y stock) which resides in the Yellow Sea and overlaps with J stock in the southern part of sub-area 6; and Hypothesis C: five stocks, referred to Y, J_W, J_E, O_W, and O_E, two of which (Y and J_W) occur in the Sea of Japan, and three of which (J_E, O_W and O_E) are found east of Japan. The O_W and O_E stocks are separated by a hard boundary.