An additional genetic analysis to investigate the plausibility of different stock scenarios in North Pacific common minke whale IST

Mutsuo Goto, Naohisa Kanda and Luis A. Pastene

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

ABSTRACT

The Workshop on western North Pacific common minke whale *Implementation Simulation Trials* (IST) defined three new baseline stock scenarios for the western North Pacific. These were based exclusively on mtDNA data derived from JARPN and JARPN II surveys. Using AIC (Akaike's Information Criterion) we evaluated the plausibility of these different scenarios. To conduct such evaluation we used up to four mtDNA haplotypes that presented the first four highest frequencies in sub-areas 7, 8 and 9 as independent parameters for the AIC. Different combinations of these haplotypes were made to evaluate the baseline scenarios and some variants. In addition we used a phylogenetic analysis of all haplotypes to define five groups of haplotypes which were used in the AIC-based evaluation. Results of the AIC for the different combinations of haplotypes were consistent, providing the best evaluation for baseline scenario A which suggest that sub-areas 7, 8 and 9 are composed of a single stock (O) with sporadic occurrence of individuals from a different stock (W) in part of sub-area 9. Although the AIC revealed that baseline scenario A was the most consistent with the available mtDNA data in these areas, a more comprehensive evaluation of stock structure in the western North Pacific is still needed. Such comprehensive evaluation should consider the results of several approaches, both genetics and non genetics.

INTRODUCTION

The Workshop on North Pacific common minke whale *Implementation Simulation Trials* (IST) was conducted in January 2002 in Seattle (Anon, 2002). Four terms of references (TOR) had been defined for the Workshop (IWC, 2002). The TOR 1) read as follow: 'review results of further analyses of genetic data for sub-areas 7, 8 and 9 to decide what trials modifications may be necessary in consequence, and re-estimation of mixing proportion between 'J' and 'O' stocks in sub-areas 2, 7 and 11'. Consequent with this TOR, the Workshop discussed new genetic information in these sub-areas, which was available basically from data collected by the JARPN (1994-1999) and JARPN II (2000-2001).

Both results of analyses based on mitochondrial DNA (mtDNA) and nuclear DNA using microsatellite were available to the Workshop. Only the information derived from the mtDNA analysis was discussed and interpreted in some detail and the subsequent scenarios of stock structure in the western North Pacific were based on the results and interpretations of this genetic marker. Thus during the Workshop there was not any comprehensive discussion on the stock definition of the minke whale based on different and independent approaches. To our view there is the need for further discussion on the term 'stock' used in the trials of minke whale (see also Kanda *et al.* 2002).

The Workshop, based on mtDNA information, adopted three new baseline stock scenarios in developing the IST for North Pacific minke whale: scenarios A, C and D (Anon., 2002). It should be noted that scenario B represent the case of no W stock in sub-areas 7, 8 and 9 and it had been considered previously in the trials.

Baseline A: three-stock scenario ('J', 'O', 'W') with the 'W' stock found only in part of sub-area 9 and only sporadically.

Baseline C : four-stock scenario overall, with 'O_W , 'O_E and 'W' to the east of Japan. Boundaries are fixed at 147°E and 157°E and there is no mixing between the stocks.

Baseline D : three-stock scenario ('J', 'O', 'W'), with 'O' and 'W' stocks to the east of Japan mixing across the area from 147°E and 162°E.

Plausibility of these different scenarios was not discussed at the Workshop and this task was assigned to the Scientific Committee meeting (Anon., 2002). Here we present the results of an AIC (Akaike's Information Criterion)-based evaluation of the stock scenarios defined by the Workshop. We were inspired in a preliminary AIC analysis conducted during the Workshop to explain the frequency of haplotype 9 (Goto *et al.*, 2002) in various sub-areas. We extended that analysis to include additional haplotypes.

We emphasize that this evaluation is conducted to investigate which of the different stock scenarios is more consistent with frequencies of mtDNA haplotypes observed in sub-areas 7, 8 and 9. The definition of stock in the western North Pacific minke whale still requiring further consideration and a more comprehensive examination based on results and interpretation of different genetic markers and non-genetic approaches is needed.

MATERIAL AND METHOD

AIC model

Plausibility of scenarios were evaluated computing the AIC values assuming binomial likelihood as follows (Akaike, 1973; 1985, Kishino *et al.*, 1991):

AIC=-2 × (binomial likelihood of the model)
+2 × (number of the free parameter of the model)

Binomial likelihood of the model is expressed as:

$$-\ln L = -\sum_{i}^{m}\sum_{j}^{a}\left(n_{ij}\ln p_{ij}\right) \qquad (1)$$

where n is sample size of ith haplotype at jth group and p is the expected proportion of ith haplotype at jth group.

Number of the free parameter of the model is $a \times (m-1)$, where a (j= 1, 2, 3, 4) is the number of groups and m (i= 1, 2, 3, 4, 5) is number of distinct haplotypes.

The relative weight (*wt*) of the model was assessed by estimating the likelihood of each model from exp(-1/2△AIC), where △AIC is the difference in AIC between an alternative model and the best AIC model (Burnham and Anderson, 1998).

Grouping of sub-areas

Each sub-area in the eastern side of Japan was divided into western (W) and eastern (E) sectors. Sub-area 7 was divided at 147°E, sub-area 8 at 153°E and sub-area 9 at 162°E (see Fig. 1 in Anon., 2002). AIC evaluation was made on different models related to each baseline stock scenario and some of their variants. Ten models evaluated are shown in the first column in Table 2. For example the notation $a=b\neq c=d$ implies a model where samples a and b are assumed to have the same haplotype proportion, similarly c and d have the same proportion, but these proportions differ between (a, b) and (c, d). In this example, the number of groups (*a*) is two.

Haplotype used for AIC

By taking into consideration of a recommendation from the Workshop, we considered other haplotypes than addition to haplotype 9. These haplotypes were chosen because they present the highest frequencies in the total sample of sub-areas 7, 8 and 9 (Table 1). Haplotype '9' had the highest frequency (13.4%). Haplotypes 7 (9.3%), 17 (6.1%) and 16 (5.7%) were next highest frequencies. The combination of haplotypes used in the AIC analysis is as follow:

1) Hap 9 and Other (m=2)
2) Hap 9, Hap 7 and Other (m=3)
3) Hap 9, Hap 7, Hap 17 and Other (m=4)
4) Hap 9, Hap 7, Hap 17, Hap 16 and Other (m=5)

We also used an additional combination of haplotypes based on phylogenetic analysis of haplotypes. In this case the evolutionary distance between two nucleotide sequences (haplotypes) was calculated according to Kimura's two parameters method (Kimura, 1980). Phylogenetic reconstruction of unique haplotypes was made using the neighbor-joining (NJ) method (Saitou and Nei, 1987) as implemented in the PHYLIP package (Felsenstein, 1993). To evaluate the confidence intervals, we used the bootstrap method (Felsenstein, 1985) and 500 replications were made. From this phylogenetic tree, we tentatively defined five groups of haplotypes based on their phylogenetic relationships (Fig. 1). In Table 2 we call this grouping as 'cluster'.

The specifications for scenarios A, B and C in Table 2 are simple. However those for scenario D are a little more complicated.

The case of Baseline D

Baseline D assumes two-stocks ('O', 'W') in the east of Japan which mix across the area 147°E and 162°E. According to this scenario, the highest densities of 'O' stock animals occur in sub-area 7W and decline to the east, while the highest densities of 'W' stock animals occur in sub-area 9E and decline to the west. We basically extend the idea of equation (1) to cover the Baseline D of a steady trend (i. e. linear gradient) in proportion of each haplotype from 7W to 9E. For example for the case of the combination of haplotype 9 and other (Table 2-1), the specifications for this scenario are as follow:

$p_{7W\,hap9} = p_A$
$p_{7E\,hap9} = 4/5\,p_A + 1/5\,p_B$
$p_{8W\,hap9} = 3/5\,p_A + 2/5\,p_B$
$p_{8E\,hap9} = 2/5\,p_A + 3/5\,p_B$
$p_{9W\,hap9} = 1/5\,p_A + 4/5\,p_B$
$p_{9E\,hap9} = p_B$

and then proportion of 'Other' are

$p_{7Wother} = 1 - p_A$
$p_{7E\,other} = 1 - (4/5\,p_A + 1/5\,p_B)$
$p_{8W\,other} = 1 - (3/5\,p_A + 2/5\,p_B)$
$p_{8E\,other} = 1 - (2/5\,p_A + 3/5\,p_B)$
$p_{9W\,other} = 1 - (1/5\,p_A + 4/5\,p_B)$
$p_{9E\,other} = 1 - p_B$

We assume the values of p_A and p_B by maximizing the equation (1) using non-linear maximization method. In the case of using more than one haplotype, the analysis is extended in the same way.

RESULTS AND DISCUSSION

The AIC values for different models related to stock scenarios and for different combination of haplotypes are shown in Table 2. Table 2-1 shows the results for the combination of haplotype 9 and other; Table 2-2 for haplotype 9, 7 and other; Table 2-3 for haplotype 9, 7, 17 and other; Table 2-4 for the case haplotype 9, 7, 17, 16 and other and Table 2-5 shows the results for the case of haplotypes defined by the phylogenetic analysis. Models 1, 6, 7 and 10 in Table 2 correspond to the cases of baseline scenarios A, B, C and D, respectively. Models 1, 2, 3, 4 and 5 are the same ones tested by the AIC during the Workshop. Models 8 and 9 correspond to variants of baseline scenario C (Figs. 3b and 3d in Anon., 2002).

In all cases, apart from Table 2-5, model 1 is the best evaluated by the AIC. This model corresponds to the baseline scenario A and then this scenario is the more plausible. The second best evaluated model in all cases (except Table 2-5) is model 3, which allow for sub-area 9W animals to be different from the rest of the sub-areas. This model is approximately 2/3 as likely as the most plausible scenario A (model 1). Models 6, 7 and 10 corresponding to baseline scenarios B, C and D have little support.

In the case of Table 2-5, which shows the results of AIC for the phylogenetic approach, the best model is model 3, which had the second best evaluation in the other AIC examinations. A possible explanation for this shift from model 1 to model 3 as the best-evaluated model is that the definition of the five groups by the phylogenetic analysis could not be completely correct. Clades in the tree of Fig. 1 are not supported by 50% majority rule in 500 simulation trees.

These results strongly supported baseline scenario A, which postulate that sub-areas 7, 8 and 9 are composed by a single stock (O stock) with sporadic occurrence of individuals from a different stocks in part of sub-area 9. It should be emphasized that all the scenarios and models on stock structure evaluated in this exercise were based on mtDNA haplotype frequency data. Thus it can be concluded that baseline scenario A is the most consistent with these mtDNA data. As mentioned earlier a more comprehensive evaluation of the stock structure in the western North Pacific should take into consideration results of the analyses using other genetic markers as well those from other non-genetic analyses.

ACKNOWLEDGMENTS

We thank H. Hatanaka, D.S. Butterworth, H. Okamura and T. Hakamada for helpful comments and suggestions on this manuscript.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle, *In*: B.N. Petrov and F. Csaki (eds.) *2nd International Symposium on Information Theory*. Akademiai Kiado, Budapest. 267-281.

Akaike, H. 1985. Prediction and entropy, *In*: A.C. Atkinson and S.E. Fienberg (eds.) *A Celebration of Statistics, the ISI Centenary Volume*. Springer, New York. 1-24.

Anonymous, 2002. Report of the Workshop on North Pacific common minke whale (*Balaenoptera acutorostrata*) *Implementation Simulation Trials*. Seattle, 19-22 January 2002.

Burnham, K.P. and Anderson, D.R. 1998. *Model selection and inference: a practical information-theoretic approach. Springer-Verlag New York.*

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 44: 1097-1108.

Felsenstein, J. 1993. PHYLIP 3.5c (Manual and program available from Joe Felsenstein, University of Washington, WA, Seattle)

Goto, M., Kanda, N and Pastene, L.A. 2002. Further mtDNA analysis on North Pacific minke whales includeing JARPN and JARPN II samples from 1994 to 2001. Paper SC/J02/NP10 presented to the IWC Scientific Committee Workshop on North Pacific comm..on minke whale (*Balaenoptera acutorostrata*) Implementation Simulation Trials, Seattle 19-22 January 2002 (unpublished). 9pp.

International Whaling Commission. 2002. Report of the Scientific Committee. Annex D. Report of the sub-committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 4:00–00.

Kanda, N., Goto, M. and Pastene, L.A. 2002. Considerations on stock structure of minke whales in western North Pacific. Paper 54/SC/SD3 was presented to this meeting.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.

Kishino, H., Kato, H., Kasamatsu, F. and Fujise, Y. 1991. Detection of heterogeneity and estimation of population characteristics from the field survey data: 1987/88 Japanese feasibility study of the Southern Hemisphere minke whales. *Ann. Inst. Statist. Math.* 43 (3): 435-453.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic tree. *Mol. Biol. Evol.* 4: 406-425.

Table 1. Haplotype frequencies used in the AIC evaluation, by sub-area.

	SA7		SA8		SA9		Total	Proportion
Hap.	7W	7E	8W	8E	9W	9E		
9	17	5	5	5	35	8	75	0.1344
7	19	1	3	9	13	7	52	0.0932
17	8	1	2	4	9	10	34	0.0609
16	14	2	3	1	8	4	32	0.0573
Other	117	29	33	47	84	55	365	0.6541
Total	175	38	46	66	149	84	558	

Table 2. The AIC values for different models related to stock scenarios and for different combination of haplotypes. Relative weight of these values is also shown.

(2-1): Hap 9 and Other

	Models	No. of group	Hap 9 and Other	ΔAIC	wt	Remark
1	7W=7E=8W=8E=9E(94+97)=9W(94+97)≠9W(95+00+01)	2	427.89	0	1	Baseline A
2	7W=7E=8W=8E≠9W=9E	2	435.95	8.06	0.018	
3	7W=7E=8W=8E=9E≠9W	2	428.39	0.5	0.779	
4	7W=7E=8W=8E=9E(94+97)≠9W(94+97)≠9W(95+00+01)	3	429.15	1.26	0.533	
5	7W=7E=8W=8E≠9E(94+97)≠9W(94+97)≠9W(95+00+01	4	431.14	3.25	0.197	
6	7W=7E=8W=8E=9W=9E	1	442.47	14.58	7E-04	Baseline B
7	7W≠7E=8W=8E≠9W=9E	3	437.94	10.05	0.007	Baseline C
8	7W=7E=8W≠8E≠9W=9E	3	437.44	9.55	0.008	
9	7W≠7E=8W=8E=9E≠9W	3	430.39	2.5	0.287	
10	7W≠9E(7E, 8W, 8E, 9W: gradient)	2	440.66	12.77	0.002	Baseline D

(2-2): Hap 9, 7 and Other

	Models	No. of group	Hap 9, 7 and Other	ΔAIC	wt	Remark
1	7W=7E=8W=8E=9E(94+97)=9W(94+97)≠9W(95+00+01)	2	761.87	0	1	Baseline A
2	7W=7E=8W=8E≠9W=9E	2	769.91	8.04	0.018	
3	7W=7E=8W=8E=9E≠9W	2	762.31	0.44	0.803	
4	7W=7E=8W=8E=9E(94+97)≠9W(94+97)≠9W(95+00+01)	3	764.94	3.07	0.215	
5	7W=7E=8W=8E≠9E(94+97)≠9W(94+97)≠9W(95+00+01	4	768.74	6.87	0.032	
6	7W=7E=8W=8E=9W=9E	1	774.45	12.58	0.002	Baseline B
7	7W≠7E=8W=8E≠9W=9E	3	773.47	11.6	0.003	Baseline C
8	7W=7E=8W≠8E≠9W=9E	3	772.31	10.44	0.005	
9	7W≠7E=8W=8E=9E≠9W	3	765.69	3.82	0.148	
10	7W≠9E(7E, 8W, 8E, 9W: gradient)	2	774.64	12.77	0.002	Baseline D

(2-3): Hap 9, 7, 17 and Other

	Models	No. of group	Hap 9, 7, 17 and Other	ΔAIC	wt	Remark
1	7W=7E=8W=8E=9E(94+97)=9W(94+97)≠9W(95+00+01)	2	1003.36	0	1	Baseline A
2	7W=7E=8W=8E≠9W=9E	2	1007.79	4.43	0.109	
3	7W=7E=8W=8E=9E≠9W	2	1004.07	0.71	0.701	
4	7W=7E=8W=8E=9E(94+97)≠9W(94+97)≠9W(95+00+01)	3	1008.27	4.91	0.086	
5	7W=7E=8W=8E≠9E(94+97)≠9W(94+97)≠9W(95+00+01	4	1008.92	5.56	0.062	
6	7W=7E=8W=8E=9W=9E	1	1014.39	11.03	0.004	Baseline B
7	7W≠7E=8W=8E≠9W=9E	3	1013.35	9.99	0.007	Baseline C
8	7W=7E=8W≠8E≠9W=9E	3	1011.77	8.41	0.015	
9	7W≠7E=8W=8E=9E≠9W	3	1008.29	4.93	0.085	
10	7W≠9E(7E, 8W, 8E, 9W: gradient)	2	1016.53	13.17	0.001	Baseline D

(2-4): Hap 9, 7, 17, 16 and Other

	Models	No. of group	Hap 9, 7, 17, 16 and Other	ΔAIC	wt	Remark
1	7W=7E=8W=8E=9E(94+97)=9W(94+97)≠9W(95+00+01)	2	1229.66	0	1	Baseline A
2	7W=7E=8W=8E≠9W=9E	2	1234.30	4.64	0.098	
3	7W=7E=8W=8E=9E≠9W	2	1230.52	0.86	0.651	
4	7W=7E=8W=8E=9E(94+97)≠9W(94+97)≠9W(95+00+01)	3	1235.18	5.52	0.063	
5	7W=7E=8W=8E≠9E(94+97)≠9W(94+97)≠9W(95+00+01	4	1237.71	8.05	0.018	
6	7W=7E=8W=8E=9W=9E	1	1238.91	9.25	0.01	Baseline B
7	7W≠7E=8W=8E≠9W=9E	3	1239.33	9.67	0.008	Baseline C
8	7W=7E=8W≠8E≠9W=9E	3	1236.47	6.81	0.033	
9	7W≠7E=8W=8E=9E≠9W	3	1234.25	4.59	0.101	
10	7W≠9E(7E, 8W, 8E, 9W: gradient)	2	1237.99	8.33	0.016	Baseline D

(2-5): Phylogenetic analysis

	Models	No. of group	Cluster	ΔAIC	wt	Remark
1	7W=7E=8W=8E=9E(94+97)=9W(94+97)≠9W(95+00+01)	2	1560.71	5.19	0.075	Baseline A
2	7W=7E=8W=8E≠9W=9E	2	1562.77	7.25	0.027	
3	7W=7E=8W=8E=9E≠9W	2	1555.52	0.00	1	
4	7W=7E=8W=8E=9E(94+97)≠9W(94+97)≠9W(95+00+01)	3	1560.02	4.50	0.105	
5	7W=7E=8W=8E≠9E(94+97)≠9W(94+97)≠9W(95+00+01	4	1565.70	10.18	0.006	
6	7W=7E=8W=8E=9W=9E	1	1567.87	12.35	0.002	Baseline B
7	7W≠7E=8W=8E≠9W=9E	3	1569.26	13.74	0.001	Baseline C
8	7W=7E=8W≠8E≠9W=9E	3	1567.85	12.33	0.002	
9	7W≠7E=8W=8E=9E≠9W	3	1561.60	6.08	0.048	
10	7W≠9E(7E, 8W, 8E, 9W: gradient)	2	1569.73	14.21	8E-04	Baseline D

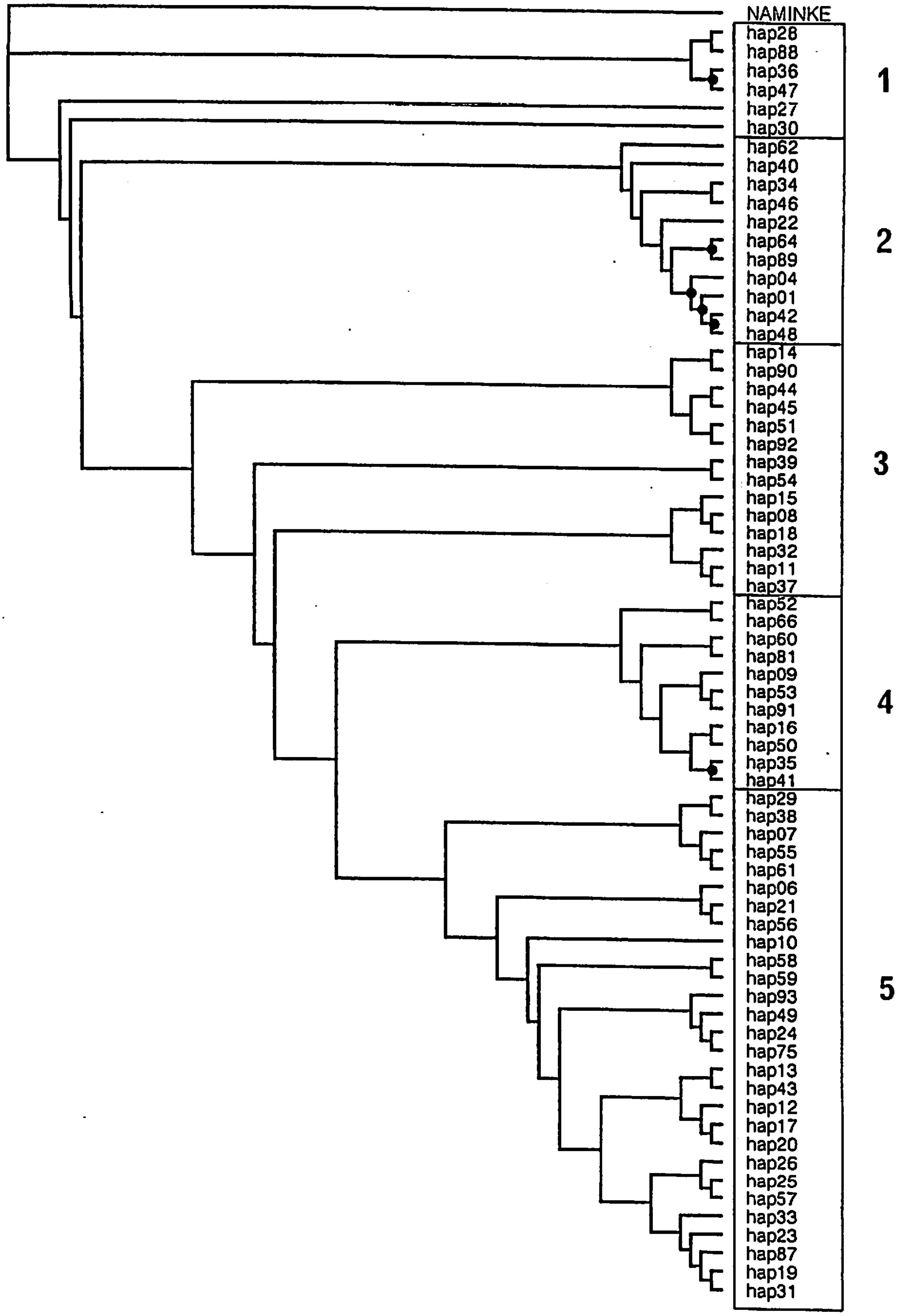

Fig. 1. Phylogenetic tree of North Pacific minke whales haplotypes constructed by Neighbor-Joining method. The dotted clades are supported by 50% majority rule of 500 bootstrap replications. Five groups were defined for AIC-based evaluation (see Table 2-5).