SC/59/BRG30

UPDATES OF STOCK STRUCTURE ANALYSES OF B-C-B STOCK OF BOWHEAD WHALES USING MICROSATELLITES

Toshihide Kitakado[1], Luis. A. Pastene[2], Mutsuo Goto[2] and Naohisa Kanda[2]

[1]*Faculty of Marine Science, Tokyo University of Marine Science and Technology, Konan 4-5-7, Minato-ku, Tokyo 108-8477, Japan*
[2]*The Institute of Cetacean Research, Toyomi-cho 4-5, Chuo-ku, Tokyo 104-0055, Japan*

ABSTRACT

We updated our genetic data analyses of the B-C-B stock of bowhead whales using the two reference microsatellite datasets. Samples of the whales were obtained from several different Alaskan villages engaged in aboriginal whaling, Canada and Okhotsk. The two reference datasets included a total of 22 or 33 loci, respectively. In order to describe spatial and temporal population structure of the B-C-B bowhead whales, the analyses were conducted using the computer program *Structure* (Pritchard *et al*. 2000, Falush *et al*. 2003) under two different models: no-admixture with independent allele frequencies and admixture with correlated allele frequencies among stocks. Results of the analysis suggested the possibility of multiple stocks. In particular, individuals passing by Barrow, at both spring and fall migration, probably came from at least two genetically distinct stocks of the bowhead whales. On the comparison between the no-admixture and admixture models, the former was better supported from estimates of the marginal likelihood and the parameter α. The mixing rate of stocks passing by Barrow was estimated based on the results of assignment probabilities under the assumption of no-admixture in each of the seasons, and the results indicated that multiple stocks were well-mixed during migration at that locality. Furthermore, the results showed a clear temporal migration pattern in fall season while no temporal pattern was observed in spring season.

INTRODUCTION

Stock structure of B-C-B stock of bowhead whales has been discussed in several recent meetings and workshops of the Scientific Committee. The most recent workshop agreed on four hypotheses on stock structure to be used in trials (IWC, 2007). Kitakado *et al*. (2007) analyzed 35 loci using the well-established software *Structure* (Pritchard *et al*. 2000, Falush *et al*. 2003) under several assumptions. In order to keep consistency of the data sets for the use of different country members of the IWC, US scientists refined datasets based on reliability of loci on genotyping as well as the number of missing loci in each individual, and finally two reference datasets were defined (22 or 33 loci). More detailed information on data is available in the workshop report (IWC 2007).

Here we updated our previous genetic analyses by *Structure* using the two reference microsatellite datasets, and discuss the implication of the results in the context of the stocks structure of the B-C-B bowhead whales.

MATERIAL AND METHOD

We employed the two reference datasets of microsatellites. We call the datasets with 22 and 33 loci as Ref 22 and Ref 33, respectively. Also we provide locality/season groups with labels, which are same as those of Givens *et al*. (2007). Sample sizes in the two reference sets are shown in Table 1.

Table 1: Number of samples used in the present microsatellite analysis by reference datasets, locality and season. A sample from Okhotsk wrongly included in Ref 33 was deleted.

Group labels	Area(Season)	Sample size	
		Ref 22	Ref 33
1	Barrow (spring)	107	97
2	Barrow (fall)	124	116
3	Chukotka (spring)	3	3
4	Chukotka (fall)	12	9
5	Gambell (spring)	5	5
6	Gambell (fall)	4	4
7	Kaktovik (fall)	15	12
8	Diomede (spring)	1	1
9	Nuiqsut (fall)	5	5
10	Pt. Hope (spring)	4	6
11	Savoonga (spring)	6	6
12	Savoonga (fall)	10	10
13	Wainwright (spring)	7	7
14	Igloolik , Canada	47	0
15	Okhotsk	62	0
	Total	412	281

Analyses were conducted using *Structure* 2.1 under two different models as follows:

1. No-admixture with independent allele frequencies among populations
2. Admixture with correlated allele frequencies under unknown values of hyper-parameters λ

The numbers of stocks (K) hypothesized in the analyses were set at K =1,2,…,5 for Ref 22 and K =1,2,…,4 for Ref 33. In each value of K in each assumption, three independent runs were conducted to take into account the variability of the marginal log-likelihood, which were approximately assessed through a Markov chain Monte Carlo (MCMC) method. Throughout the studies, the burn-in period was fixed at 50,000, and the number of further MCMC runs was set at 1,000,000 to estimate the parameters and marginal likelihood.

In addition to the number of stocks, we also compared the marginal likelihood across the two models, no-admixture and admixture. Furthermore, for this comparison we also assessed the value of α in the admixture model; if the value of α is close to 0, each individual is modeled as originating from a single population, and therefore the model tends the no-admixture model.

RESULTS AND DISCUSSION

The number of stocks

The results of the marginal log-likelihood for Ref 22 and Ref 33 were summarized in Tables 2 and 3, respectively. Also, plots of assignment probabilities obtained from the run with the median value of the logarithm of marginal likelihood in a total of three runs were shown in Figures 1 and 2

For Ref 22 the two different analyses supported hypotheses of more than three stocks in whole areas (see Table 2). It should be noted that the best among all the models is the no-admixture model with *K*=4. In fact, as shown in assignment probabilities, as a single stock occupies Okhotsk, at least two or three populations contribute to B-C-B stock of bowhead whales (see Figure 1-1).

Table 2. Summary of results for the reference dataset Ref 22

K	Run	Log-marginal likelihood (logML)	Variance of logML	Difference from case of *K*=1	
1	1	-38607	163		
	2	-38607	163		
	3	-38606	162		
	Median	**-38607**		**0**	
Model 1: No-admixture					
2	1	-37889	334		
	2	-37890	334		
	3	-37893	341		
	Median	**-37890**		**717**	
3	1	-37588	609		
	2	-37584	603		
	3	-37584	602		
	Median	**-37584**		**1022**	
4	1	-37360	818		
	2	-37360	819		
	3	-37355	811		
	Median	**-37360**		**1246**	best
5	1	-37578	1395		
	2	-37302	1032		
	3	-37897	1944		
	Median	**-37578**		**1028**	
Model 2: Admixture with unknown lambda					
2		-37877	380		
		-37876	386		
		-37868	372		
	Median	**-37876**		**730**	
3		-37789	977		
		-37800	998		
		-37794	985		
	Median	**-37794**		**813**	
4		-37781	1538		
		-37757	1496		
		-37738	1461		
	Median	**-37757**		**850**	
5		-37743	1925		
		-37746	1921		
		-37732	1900		
	Median	**-37743**		**863**	

For Ref 33, which does not include Canadian and Okhotsk samples, the analyses under the no-admixture model indicated that B-C-B stock of bowhead whales is comprised of more than 1 stock (see Table 3). On the other hand, under the admixture model, the marginal likelihood showed the hypothesis of K=1 was better than that of K=2 although the difference between those values in the marginal likelihood was relatively small.

Table 3. Summary of results for the reference dataset Ref 33

K	Run	Log-marginal likelihood (logML)	Variance of logML	Difference from case of *K*=1	
1	1	-35448	196		
	2	-35449	197		
	3	-35449	197		
	Median	**-35449**		**0**	
Model 1: No-admixture					
2	1	-35183	475		
	2	-35186	481		
	3	-35185	478		
	Median	**-35185**		**264**	best
3	1	-38259	6773		
	2	-35293	1001		
	3	-35277	973		
	Median	**-35293**		**156**	
4	1	-35295	1003		
	2	-35446	1299		
	3	-35287	994		
	Median	**-35295**		**153**	
Model 2: Admixture with unknown lambda					
2		-35484	888		
		-35566	1056		
		-35456	822		
	Median	**-35484**		**-35**	
3		-36290	3029		
		-35640	1794		
		-35848	2159		
	Median	**-35848**		**-399**	
4		-37055	5016		
		-36012	3018		
		-35674	2367		
	Median	**-36012**		**-563**	

As described in Pritchard *et al*. (2000), the values of the marginal likelihood function should be carefully used. However, the large difference between that value for K=4 and the values for others under the no-admixture model strongly support K=4 for Ref 22. For Ref 33, which does not include sample from Canadian and Okhotsk localities, a two-stock scenario is the most likely hypothesis under the assumption of no-admixture. Although there was relatively moderate difference in the marginal likelihood under the admixture, especially in Ref 33, the evidence of the single-stock hypothesis for B-C-B stock is never strong.

Comparison of models

As easily identified in Tables 2 and 3, the log-likelihood for K=4 under the no-admixture model marked the best among all the models for Ref 22, and the log-likelihood for K=2 under no-admixture did the best for

Ref 33. The performance of such evaluation method should be investigated further, but these scenarios are considered the most plausible scenario based on *Structure* at this stage.

In principle, the meanings of outputs of assignment probabilities are different between the two cases of no-admixture and admixture. In the former case, it is really “an assignment probability of an individual to each putative population origin” while in the latter case it means “a proportion of an individual’s genome originated from each population”. Under the assumption of multiple stocks with currently peculiar discrete breeding grounds, the assumption of no-admixture may be promising.

In fact, the posterior distribution of α shown in Figure 3 suggested that the value of α was so small, and therefore it may also imply that the no-admixture model is promising.

Mixing proportions

The estimates of mixing proportions under the best scenario were shown in Table 4. Unfortunately, there were not consistent pattern of mixing proportions between the cases of Ref 22 and Ref 33, but the results indicated that multiple populations are well-mixed in the migratory corridor of B-C-B stock.

Table 4. Group-wise mixing proportions under the best scenario for the two reference datasets

			Mixing proportion			
Dataset	Group labels	Area(Season)	Stock 1	Stock 2	Stock 3	Stock 4
Ref 22	1	Barrow (spring)	0.362	0.352	0.285	0.002
	2	Barrow (fall)	0.286	0.385	0.327	0.002
	3	Chukotka (spring)	0.577	0.174	0.233	0.016
	4	Chukotka (fall)	0.316	0.438	0.246	0.000
	5	Gambell (spring)	0.537	0.421	0.042	0.000
	6	Gambell (fall)	0.550	0.192	0.255	0.003
	7	Kaktovik (fall)	0.382	0.275	0.343	0.000
	8	Diomede (spring)	0.295	0.000	0.705	0.000
	9	Nuiqsut (fall)	0.572	0.042	0.386	0.000
	10	Pt. Hope (spring)	0.576	0.412	0.013	0.000
	11	Savoonga (spring)	0.160	0.547	0.292	0.000
	12	Savoonga (fall)	0.213	0.350	0.437	0.000
	13	Wainwright (spring)	0.423	0.245	0.332	0.000
	14	Igloolik , Canada	0.489	0.355	0.155	0.000
	15	Okhotsk	0.001	0.001	0.018	0.980
Ref 33	1	Barrow (spring)	0.547	0.453		
	2	Barrow (fall)	0.545	0.455		
	3	Chukotka (spring)	0.718	0.282		
	4	Chukotka (fall)	0.527	0.473		
	5	Gambell (spring)	0.776	0.224		
	6	Gambell (fall)	0.334	0.666		
	7	Kaktovik (fall)	0.522	0.478		
	8	Diomede (spring)	0.187	0.813		
	9	Nuiqsut (fall)	0.337	0.663		
	10	Pt. Hope (spring)	0.702	0.298		
	11	Savoonga (spring)	0.513	0.487		
	12	Savoonga (fall)	0.496	0.504		
	13	Wainwright (spring)	0.525	0.475		
	14	Igloolik , Canada	-	-		
	15	Okhotsk	-	-		

Next, we investigated temporal stock structure at Barrow using the assignment probabilities estimated under the non-admixture model with *K*=2 for Ref 33. The probabilities to a stock were split into two seasons, spring and fall. These assignment probabilities were plotted against the days passed since April 1st. The smooth functions were estimated with GAM.

As shown in Figure 4, the results indicated a clear temporal migration pattern in fall season while no temporal pattern was observed in spring season.

The values of Fst for the each of the loci were shown in Figure 5. Although most of the loci showed low Fst values, a few of them definitely had the values quite informative for drawing population differentiation.

In summary, results of the present microsatellite analysis, based on two reference datasets, are not inconsistent with a multiple stock scenario for B-C-B bowhead whales. These results are better explained by a situation in which two or more stocks migrate during migration season (e.g. Stock Scenario D defined by the intersessional workshop, IWC, 2007).

ACKNOWLEDGEMENT

We are grateful to John W. Bickham and Ryan Huebinger for providing microsatellite genotypes of the bowhead whale and to the IWC-Data Availability Group (DAG) for facilitating the process of data access.

REFERENCES

Falush, D., M. Stephens and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies, *Genetics*, 164: 1567-1587.

Givens, G.H., R. Huebinger, J.W. Bickham, C. George and R. Suydam. 2007. Patterns of genetic differentiation in bowhead whales (Balaena mysticetus) from the Western Arctic, Paper SC/J07/AWMP2 presented to the Second Intersessional AWMP Workshop to prepare for the 2007 Bowhead Implementation Review, January, 2007.

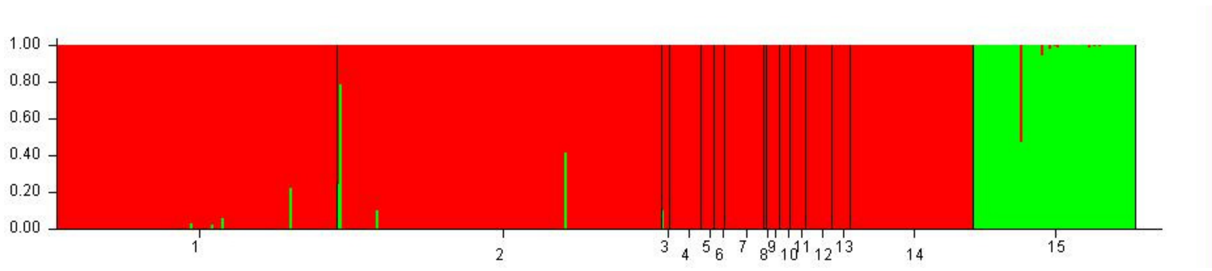
International Whaling Commission. 2006. Report of the First Intersessional AWMP Workshop for the 2007 Bowhead Implementation Review.

International Whaling Commission. 2007. Report of the Second Intersessional AWMP Workshop to prepare for the 2007 Bowhead Implementation Review.

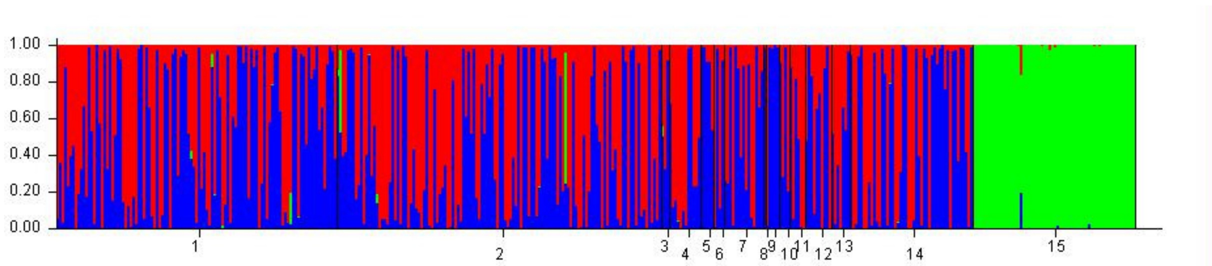
Kitakado, T., L.A. Pastene and M. Goto. 2007. Stock structure analyses of B-C-B stock of bowhead whales using microsatellites, Paper SC/J07/AWMP3 presented to the Second Intersessional AWMP Workshop to prepare for the 2007 Bowhead Implementation Review, January, 2007.

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

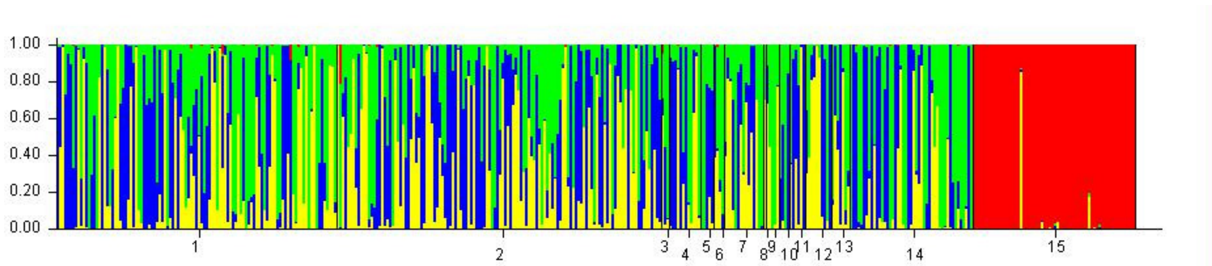
K=2



K=3



K=4



K=5

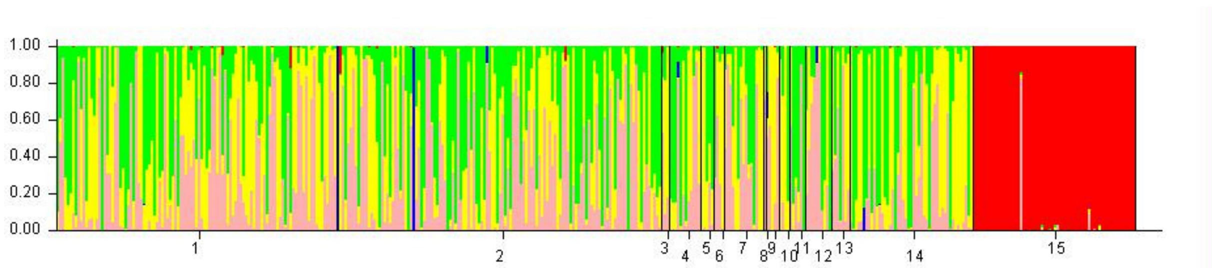
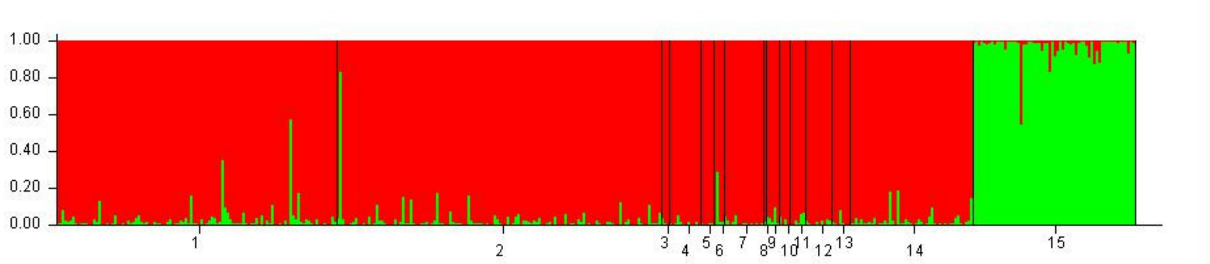
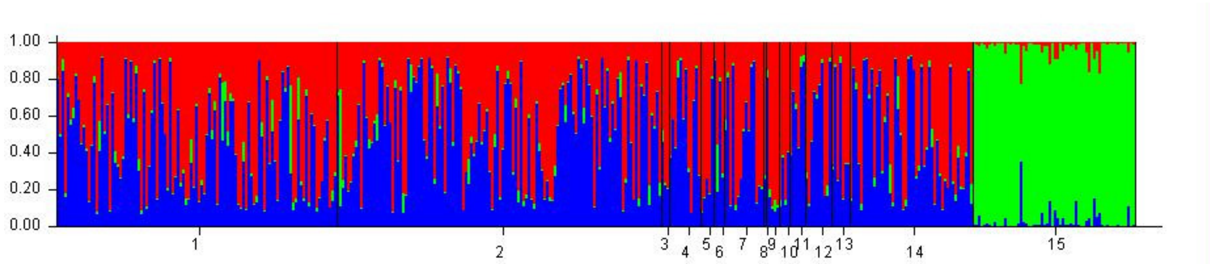


Figure 1-1. Assignment probabilities for individuals to populations based on the reference dataset Ref 22 under Model 1 (no-admixture). The numbers on the horizontal line mean the group labels. (Group labels: 1=Barrow (spring), 2=Barrow (fall), 3=Chukotka (spring), 4=Chukotka (fall), 5=Gambell (spring), 6=Gambell (fall), 7=Kaktovik (fall), 8=Diomede (spring), 9=Nuiqsut, 10= Pt. Hope , 11=Savoonga (spring), 12=Savoonga (fall), 13=Wainwright (spring), 14=Igloolik, Canada, 15=Okhotsk)

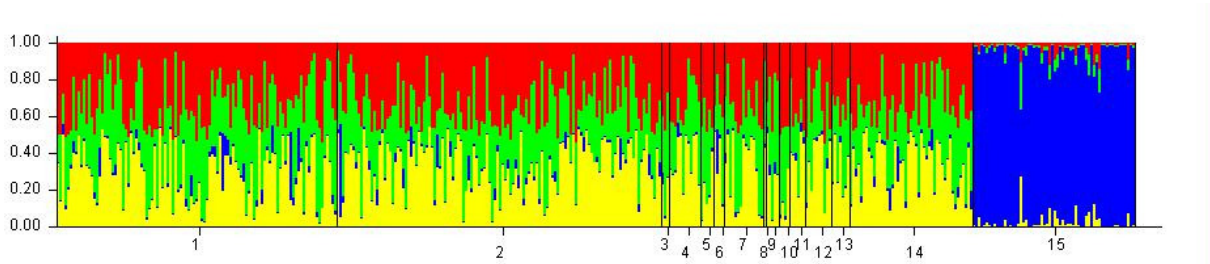
K=2



K=3



K=4



K=5

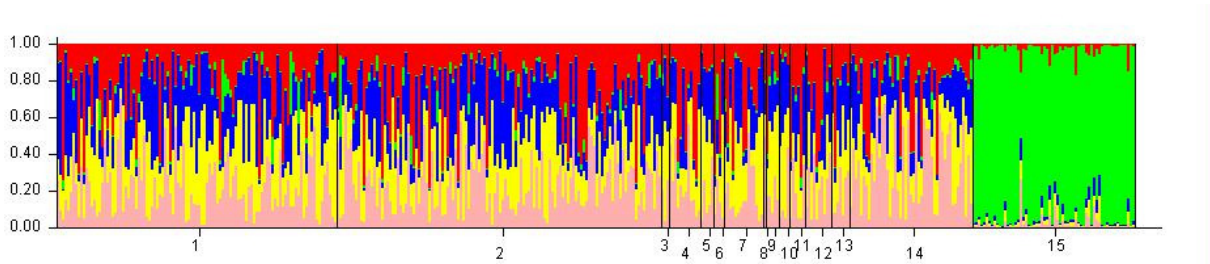
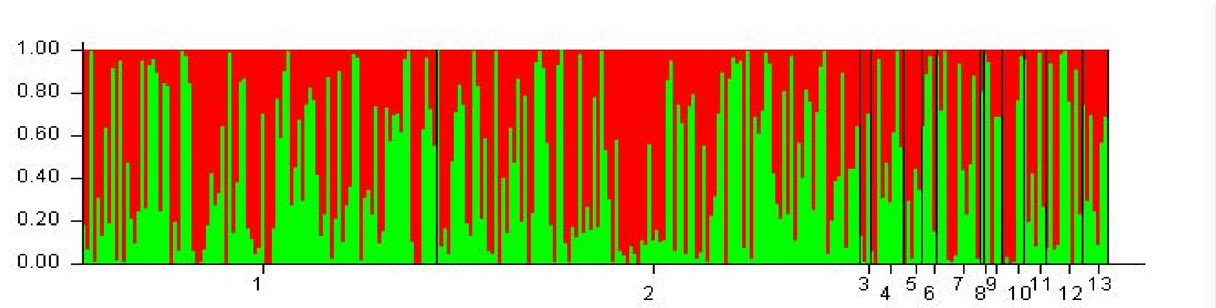
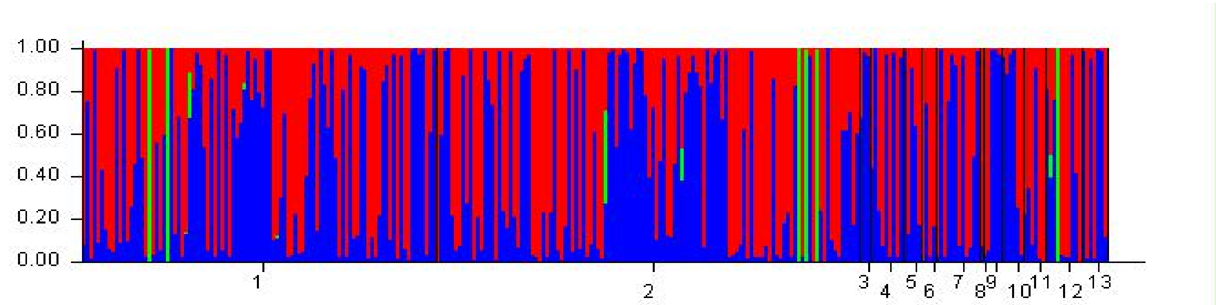


Figure 1-2. Assignment probabilities for Ref 22 under Model 2 (admixture).
(Group labels: 1=Barrow (spring), 2=Barrow (fall), 3=Chukotka (spring), 4=Chukotka (fall), 5=Gambell (spring), 6=Gambell (fall), 7=Kaktovik (fall), 8=Diomede (spring), 9=Nuiqsut, 10= Pt. Hope , 11=Savoonga (spring), 12=Savoonga (fall), 13=Wainwright (spring), 14=Igloolik, Canada, 15=Okhotsk)

K=2



K=3



K=4

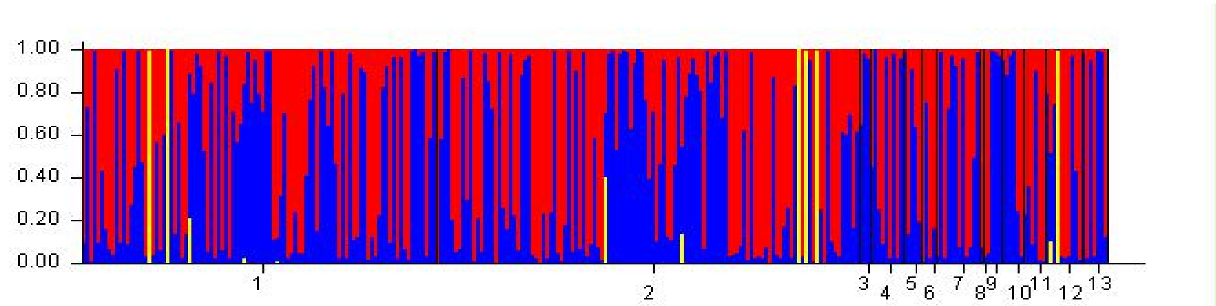
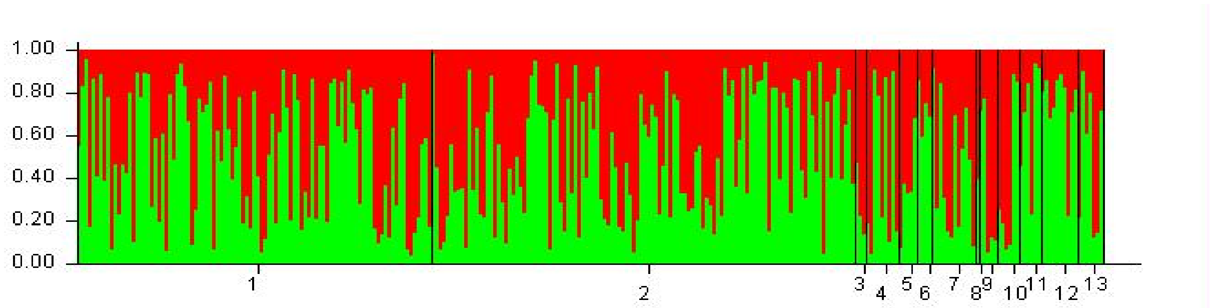
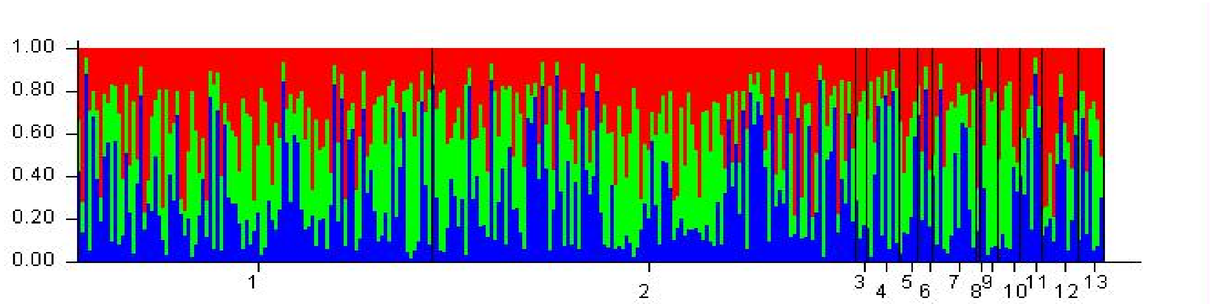


Figure 2-1. Assignment probabilities for individuals to populations based on the reference dataset Ref 33 under Model 1 (no-admixture).
(Group labels: 1=Barrow (spring), 2=Barrow (fall), 3=Chukotka (spring), 4=Chukotka (fall), 5=Gambell (spring), 6=Gambell (fall), 7=Kaktovik (fall), 8=Diomede (spring), 9=Nuiqsut, 10= Pt. Hope , 11=Savoonga (spring), 12=Savoonga (fall), 13=Wainwright (spring))

K=2



K=3



K=4

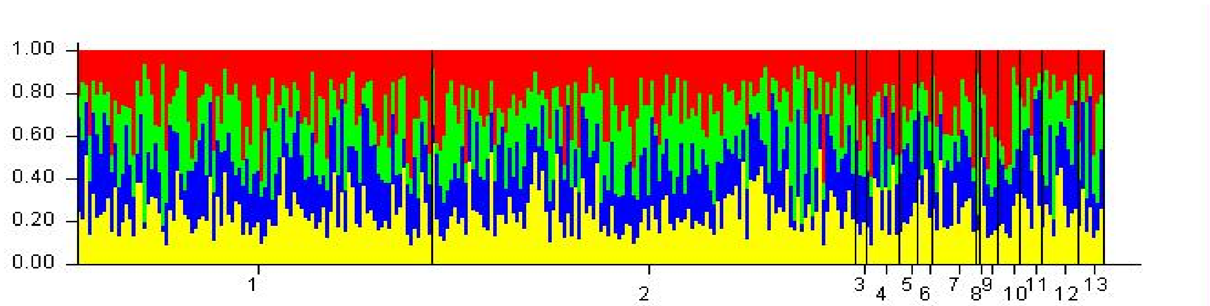
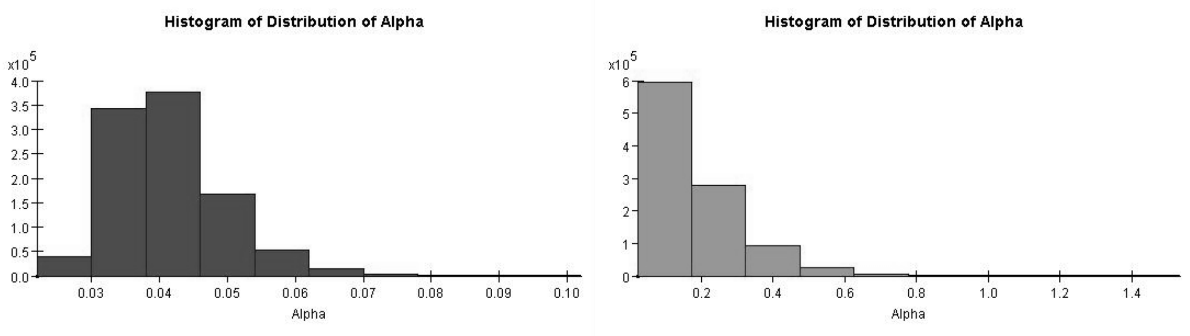
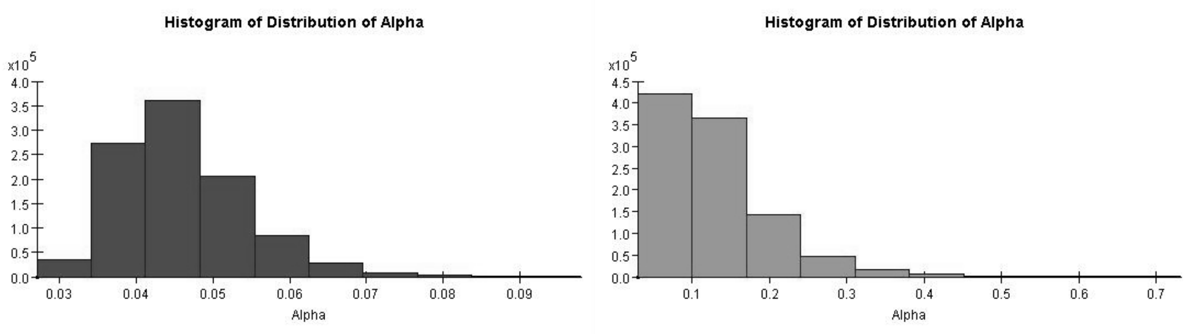


Figure 2-2. Assignment probabilities for Ref 33 under Model 2 (admixture).
(Group labels: 1=Barrow (spring), 2=Barrow (fall), 3=Chukotka (spring), 4=Chukotka (fall), 5=Gambell (spring), 6=Gambell (fall), 7=Kaktovik (fall), 8=Diomede (spring), 9=Nuiqsut, 10= Pt. Hope , 11=Savoonga (spring), 12=Savoonga (fall), 13=Wainwright (spring))

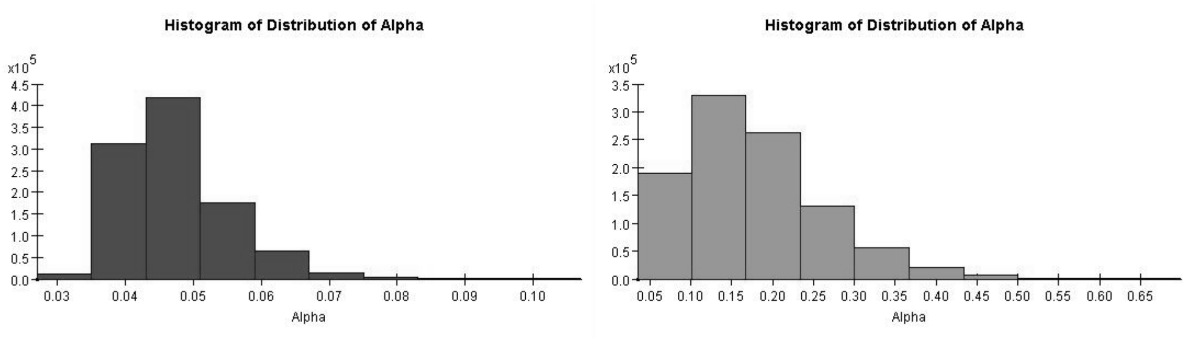
***K* =2**



***K* =3**



***K* =4**



***K*=5**

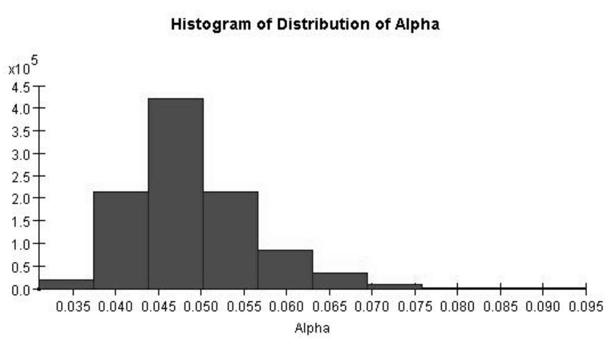
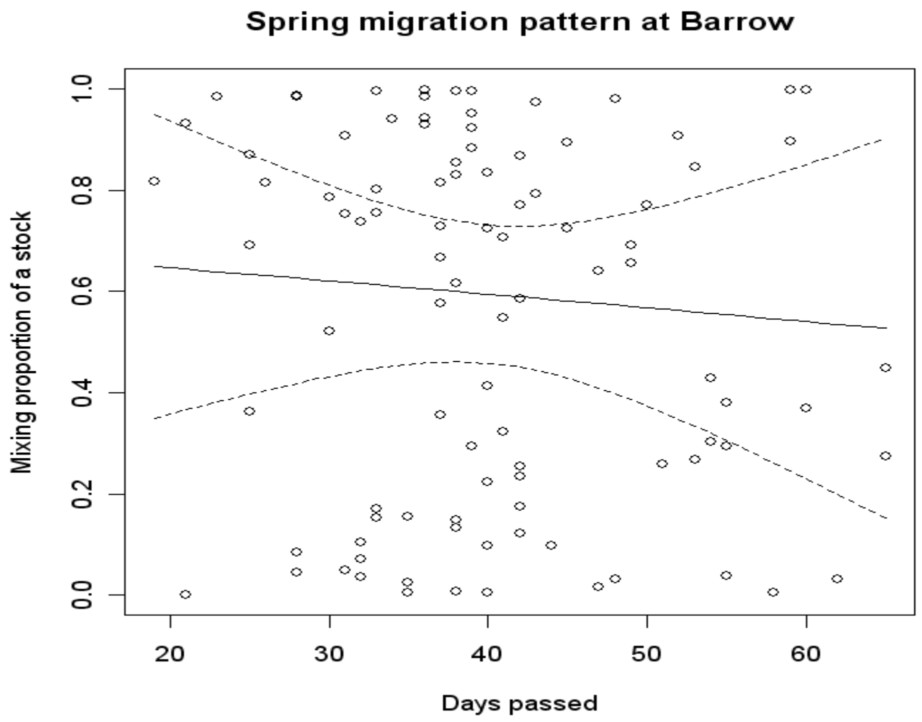


Figure 3. Posterior distribution of α under Model 2 (admixture) for Ref 22 (left) and Ref 33 (right).



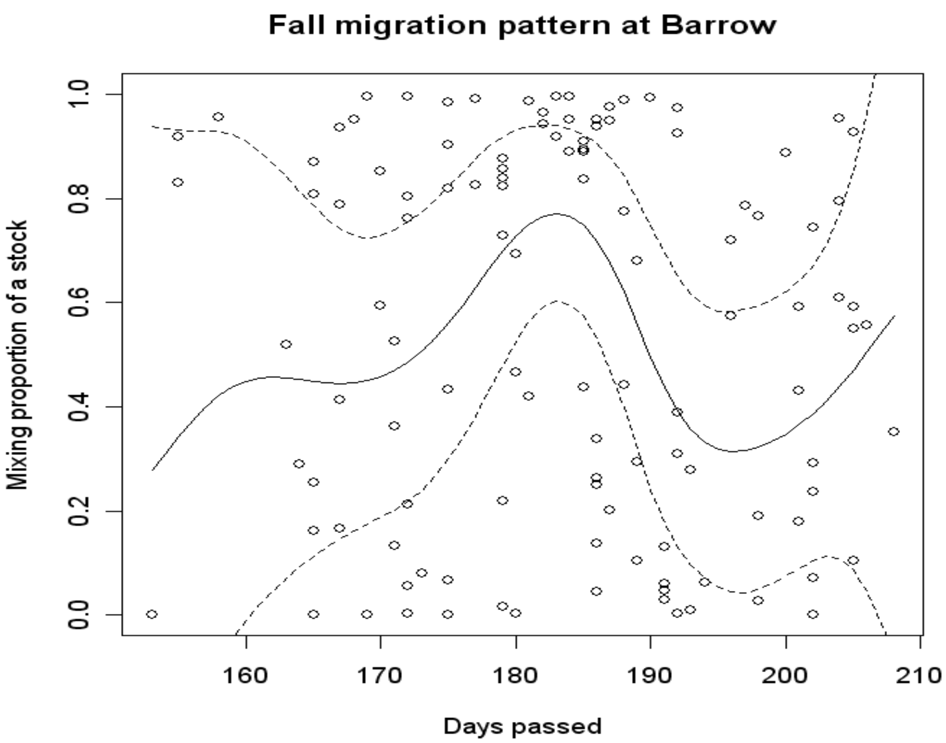


Figure 4. Spring and fall migration patterns at Barrow based on the estimates of the mixing proportions under the no-admixture model with *K*=2 for Ref 33. The smooth functions were estimated with GAM. The dotted lines show 95% condifidence bounds. The horizotal axes mean days passed since April 1st.

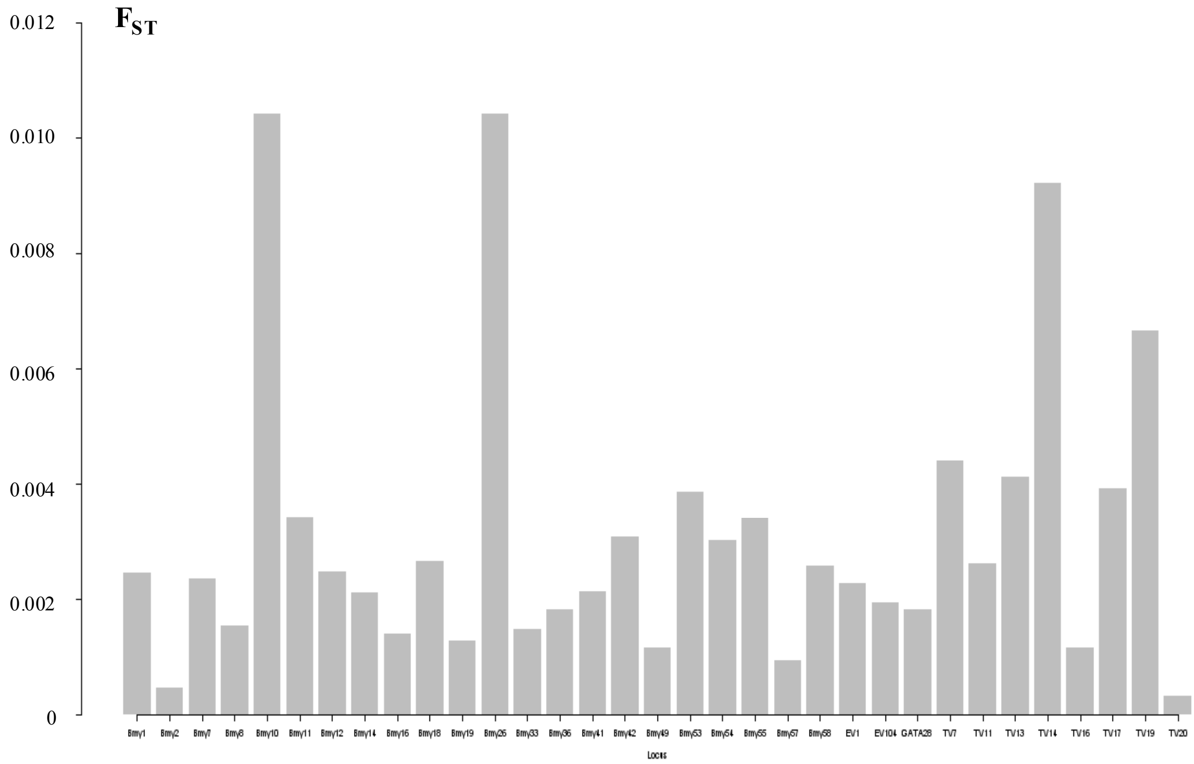


Figure 5. The values of Fst among loci under the no-admixture model with *K*=2 for Ref 33.