Ann. Zool. Fennici 41: 465–475 ISSN 0003-455X
Helsinki 22 June 2004

Geographical structure and postglacial history of mtDNA haplotype variation in Atlantic salmon (*Salmo salar* L.) among rivers of the White and Barents Sea basins

Tina Asplund[1], Alexei Veselov[2], Craig R. Primmer[3], Igor Bakhmet[2], Alexander Potutkin[4], Sergey Titov[5], Alexander Zubchenko[4], Igor Studenov[6], Svjatoslav Kaluzchin[7] & Jaakko Lumme[8]*

[1] *Environmental Center of Western Finland, P.O. Box 77, FIN-67101 Kokkola, Finland*
[2] *Institute of Biology, Karelian Research Center, Pushkinskaya 11, 185610 Petrozavodsk, Russia*
[3] *Department of Biological and Environmental Sciences, P.O. Box 65, FIN-00014 University of Helsinki, Finland*
[4] *Knipovich Polar Research Institute of Marine Fisheries and Oceanography (PINRO), Murmansk, Russia*
[5] *State Research Institute on Lake and River Fisheries (GosNIORKh), St. Petersburg, Russia*
[6] *Polar Research Institute of Marine Fisheries and Oceanography, Northern Branch (SevPINRO), Arkhangelsk, Russia*
[7] *Varzuga Scientific Research Centre of Ecology of Polar Ecosystem, Murmansk region, Russia*
[8] *Department of Biology, University of Oulu, P.O. Box 3000, FIN-90014 Oulu, Finland (*corresponding author's e-mail: jaakko.lumme@oulu.fi)*

Received 2 Dec. 2003, revised version received 12 Feb. 2004, accepted 17 Feb. 2004

Asplund, T., Veselov, A., Primmer, C. R., Bakhmet, I., Potutkin, A., Titov, S., Zubchenko, A., Studenov, I., Kaluzchin, S. & Lumme, J. 2004: Geographical structure and postglacial history of mtDNA haplotype variation in Atlantic salmon (*Salmo salar* L.) among rivers of the White and Barents Sea basins. — *Ann. Zool. Fennici* 41: 465–475.

Atlantic salmon (*Salmo salar* L.) populations from 30 spawning rivers along the northeastern European coast were analyzed for restriction site variation on the standard 2000 base pair fragment of mitochondrial DNA. Overall genetic diversity was higher than in the North Sea or Baltic Sea populations. Linear partitioning of the populations into three geographical areas contributed 27% to the total molecular variance. The Atlantic haplotypes dominate in the westernmost populations along the Barents Sea up to Svyatoi Nos. A haplotype belonging to the North American phylogenetic branch was also frequent along the western Barents coast. The Kola Peninsula rivers draining to the White Sea are dominated by haplotypes hypothesized to originate from the Komi Ice Lake, diluted by Atlantic immigrants. On the Karelian and Arkhangelsk coast of the White Sea, only two haplotypes were present, which are suggested to be derived directly from the freshwater refugium.

Introduction

The global and local population structure of Atlantic salmon (*Salmo salar* L.) is a particularly interesting topic to be analyzed by molecular methods. The species is economically valuable, but vulnerable, and merits conservation measures. The spawning and feeding locations of anadromous populations are geographically separated, creating the need for stock identification in the open sea catchment areas. The distribution range is wide and covers previously glaciated areas in the north. In this paper, we elucidate the genetic structure of the northernmost populations, in order to localize the glacial refugia and trace the postglacial recolonization routes. Glacial isolation in ecologically diverse refugia may lead to cryptic adaptations, which are of importance for delineating the management units of stocks. The tolerance and co-occurrence of Baltic salmon with the parasite *Gyrodactylus salaris* is such an adaptation, suggested to have evolved during the latest glacial maximum in populations isolated in a freshwater refugium, while the salmon stocks derived from Atlantic coasts remained susceptible to this pest (Meinilä *et al*. 2004).

Ståhl's (1987) allozyme-based global molecular biogeographic analysis of American, European and Baltic salmon populations was the first comprehensive phylogeographic study. Ståhl's work has been geographically expanded and complemented by several subsequent studies using allozymes (Kazakov & Titov 1991, 1993, Elo 1993, Elo *et al*. 1994, Skaala *et al*. 1998, Makhrov *et al*. 1998, Koljonen *et al*. 1999).

Restriction fragment length polymorphism (RFLP) analysis of circular mitochondrial DNA was the next globally applied molecular method (Bermingham *et al*. 1991). Later on, polymerase chain reaction (PCR) and the subsequent restriction fragment length polymorphism (RFLP) analysis of a 2000 bp fragment covering part of 16S rRNA and *ND1* genes of mitochondrial DNA (mtDNA) has been widely adopted (Nielsen *et al*. 1996, 1998, Verspoor *et al*. 1999, King *et al*. 2000, Nilsson *et al*. 2001, Consuegra *et al*. 2002). This method was improved and standardized by sequencing the same "standard" segment of the common European and American RFLP haplotypes (Nilsson *et al*. 2001) and it now offers the largest connectible data set, accumulating a sample size in excess of 6200 specimens published from 115 European salmon populations, including the work of this paper.

This study expands earlier salmon population analyses into a region thus far poorly studied, to the north-east extreme of the species range. The westernmost population included is from the River Tana in northern Norway and Finland, and the easternmost samples are from tributaries of the Pechora, near the Ural mountains at the extreme eastern border of the natural distribution of Atlantic salmon. Due to their remote location, the northeastern populations are still relatively unaffected by human activities. It is, however, a region of great commercial importance, since a major proportion of the European natural salmon smolt production might come from the rivers included in this paper. The origins of open sea feeding salmon remains the main concern of praxis-oriented fish genetics (e.g., Bermingham *et al*. 1991, Cutler *et al*. 1991). Mixed-stock marine harvests are an international management problem, and the genetic characterization of productive populations is one of the first steps towards scientifically regulated fisheries.

Materials and methods

Fish sampling and laboratory methods

The sampling details for each river are presented in Table 1. The possible separate samples from tributaries, and repeated collections are pooled in calculations under the primary river. The microscale distribution and population structure is being studied by microsatellite markers and will be reported elsewhere.

For this survey, practically all the fish have been caught by electrofishing in rapids, and they have been of mixed age 0+ to 4+ fry or parr, seldom smolts or precocious males. Samples from the Tana #28 (the numbering corresponds to that in Table 1 and Fig. 1) and Kachkovka #19 rivers also contained wild-caught adult fish. The River Kola #7 sample consisted of hatchery raised juveniles, as did half of the fish in the Umba #12 sample. Our own population samples were complemented by two samples from Verspoor *et al*. (1999). The total number of first-degree rivers was 28, while the total number of separate population samples was 48. The total number of fish analysed was 1759.

After the catching and measuring of the fish, small cuts of fin tissue were stored in ethanol and transferred to the laboratory in Finland. The

Table 1. The *Salmo salar* populations surveyed, their sample size and mtDNA RFLP haplotype. The numbering of the populations corresponds to the numbering in the map in Fig. 1. The size class of the river is in estimated salmon stock: 1 = less than 100 ascending adults annually, 2 = 100–500, 3 = 500–1000, 4 = 1000–5000, 5 = 5000–10000, and 6 = more than ten thousand ascending spawners per year. *H* = haplotype diversity.

#	Population Location of sample, Lat. N, Long. E	*N*	Stock size class	Mitochondrial haplotypes *Ava*II, *Hae*III, *Hinf*I, *Rsa*I						*H*
				AAAA	AABA	BBBA	BBBB	NFLD	rare	
	Eastern Barents Sea									
1	Pechora: Un'ia, 61°32´, 58°15´	11	6		11					0
	Pechora: Pizhma, 64°53´, 51°17´	21	6		19		2			0.181
	White Sea: Arkhangelsk									
2	Megra, 66°03´, 41°43´	58			18	4	36			0.523
3	Zolotitsa, 65°30, 41°02´	7			7					0
4	S. Dvina: Padoma, 61°56´, 44°19´	56	6		10		46			0.299
	S. Dvina: Pyshentsa 63°26´, 46°55´	11	6		11					0
	S. Dvina: Niukhcha 63°40´, 46°16´	4	6		4					0
	White Sea: Karelian coast									
5	Suma, 64°14´, 35°25´	52	3		3		49			0.111
6	Kusema, 65°22´, 34°15´	14	2		3		11			0.363
7	Pon'goma, 65°18´, 34°02´	51	2		30		21			0.494
8	Pulon'ga Karelia, 66°18´, 33°17´	50	2		10		40			0.327
9	Nil'ma, 66°27´, 33°05´	43	1				43			0
10	Kolvitsa, 67°00´, 33°10´	45	2		21		24			0.509
11	Pila, 66°46´, 34°10´	41	2		13		28			0.444
POPULATION GROUP III White Sea		464			160	4	300			0.464
	Kola Peninsula									
12	Umba (hatchery) 66°47´, 34°25´	38			21	6	11			0.602
	Umba (wild) 66°49´, 34°17´	48	6		28	5	13		AABB 2	0.586
	Umba: Ukko 66°43, 34°18´	19			12		6		ADBA 1	0.526
13	Olenitsa, 66°28´, 35°21´	49	2		31	18				0.474
14	Varzuga 66°54´, 35°55´	40	6		28	6	6			0.477
	Varzuga (spawners 2001)	16			13	2	1			0.342
	Varzuga: Ser'ga 66°28´, 36°30´	31	2		23	5	3			0.428
	Varzuga: Aren'ga 66°32´, 32°11´	35	1		34	1				0.057
	Varzuga: Kitsa 66°21´, 36°53´	46	2		39	4	3			0.275
15	Pulon'ga Kola, 66°15´, 39°59´	47	3		35	10	1	1		0.408
16	Likhodeevka, 66°21´, 40°11´	45	2		31	9	2	2	AEAA 1	0.492
17	Babya, 66°23´, 40°18´	47	3		36	7	1	3		0.395
18	Ponoi, 67°02´, 39°14´	38	6		16	17	3	2		0.630
19	Kachkovka, ad. 67°26´, 40°57´	29	2	1	17	7	1	3		0.606
	Kachkovka, juv. 67°26´, 40°57´	49		4	11	10	15	8	AEAA 1	0.797
POPULATION GROUP II Kola		577		5	375	107	66	19	5	0.530
	Barents Sea West									
20	Drozdovka 68°20´, 38°25´	40	2		7	23	10			0.591
21	Varzina 68°22´, 38°10´	40	3			20	13	7		0.629
22	Murmansk fjord: Kola Hatchery	45	5		4	8	25	6	BBAA 2	0.646
23	Murmansk: Tuloma, 68°41´, 31°55´	36	4		1	7	26	2		0.449
	Murmansk: Pecha[1]	44	1		6		31	7		0.470
24	Ura, 69°18´, 32°48´	42	3	3	1	12	23	3		0.623
25	Zapadnaia Litsa, 69°23´, 32°11´	43	3			16	23	4		0.580
26	Titovka, 69°29´, 31°50´	40	3	1	2	8	25	3	ADBA 1	0.552
27	Neiden/Näätämö, 69°36´, 28°44´	119	2		21	7	80	11		0.509
	Näätämö: Silisjoki, 69°35´, 28°22´	30			14	6	10			0.652

continues

Table 1. Continued.

#	Population Location of sample, Lat. N, Long. E	*N*	Stock size class	Mitochondrial haplotypes *Ava*II, *Hae*III, *Hinf*I, *Rsa*I						*H*
				AAAA	AABA	BBBA	BBBB	NFLD	rare	
28	Tana, Anarjohka, 69°13´, 25°42´	14	1		2	4	7	1		0.692
	Iesjokka, 69°14´, 25°38´	17	1		4	10	3			0.603
	Karasjokka, 69°25´, 25°44´	15	1		5	4	6			0.705
	Lakselven	20	1				20			0
	Maskejohka	22	1			3	19			0.247
	Tana/Teno, 70°04´, 27°39´	15	6		3	4	8			0.648
	Tana 2000	30	6		3	12	14		BBAB 1	0.632
	Tana 2001	58	6		7	19	30	1	BBAB 1	0.607
	Tana[1]	48	6	11		32	5			0.503
POPULATION GROUP I Barents W		718		15	80	195	378	45	5	0.633

[1] Verspoor *et al.* (1999)

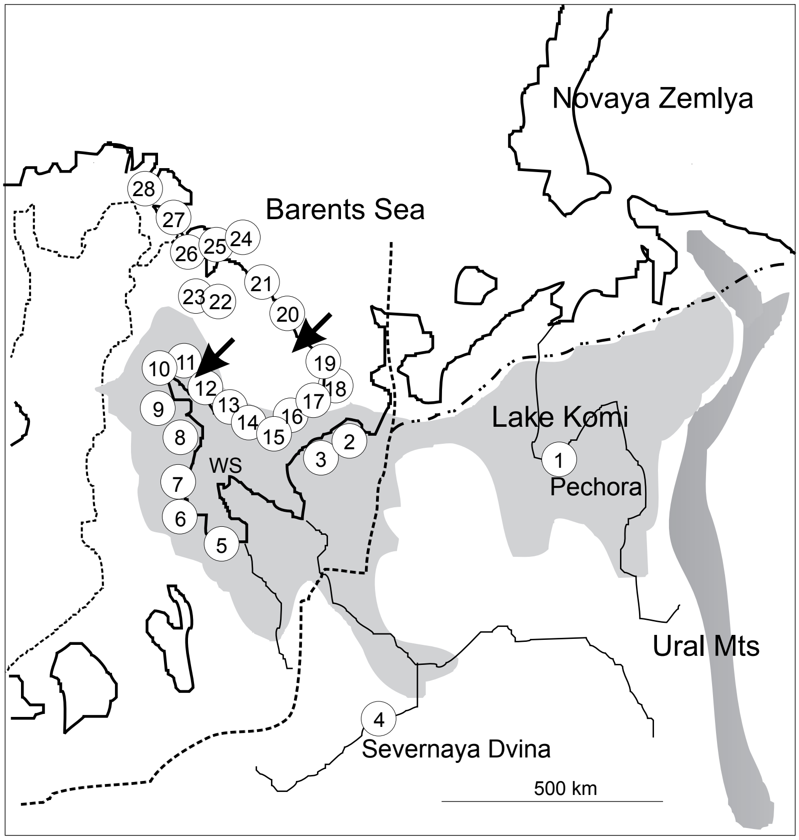

Fig. 1. The sampled first-degree rivers (details in Table 1). Arrows indicate the borders of the population groups explained in the text. The approximate extent of the lake at 90 000 BP is marked by shadowing. Lake Komi was dammed by the Early/Middle Weichselian Ice Shield of the Barents and Kara Seas (dash/double-dot line). The White Sea ice lake (WS) existed later, reaching its maximum *relative* height at about 10 000 BP. The dashed line shows the margin of the Scandinavian Ice Shield during the latest Weichselian maximum 20 000 BP, when the eastern Barents Sea was ice free (Svendsen *et al.* 1999, Mangerud *et al.* 2001, Maslennikova & Mangerud 2001).

DNA extraction, PCR and RFLP protocols of 2000 bp of 16S rRNA/*ND1* were as in Nilsson *et al.* (2001). The restriction enzymes used were *Ava*II, *Hae*III, *Hinf*I and *Rsa*I, which appear to resolve most of the European variation in this segment of the mtDNA.

Data analysis

The haplotype nomenclature was adopted from Nielsen *et al.* (1996) which was the first paper using this segment of mtDNA and the same set of restriction enzymes on a wide geographic

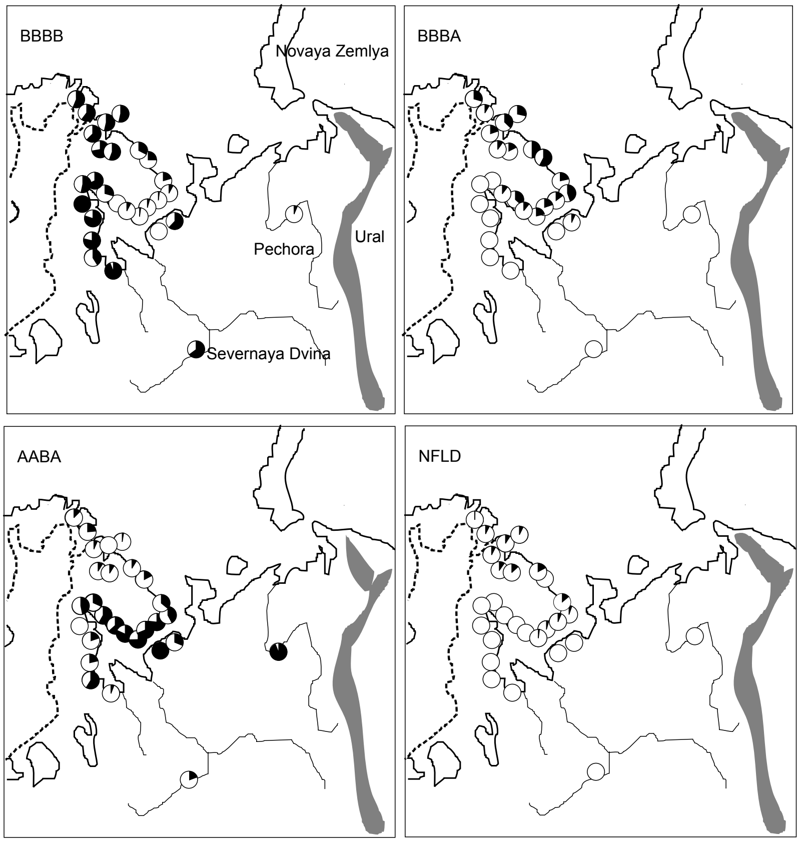

Fig. 2. Distribution of the major mtDNA haplotypes BBBB, BBBA, AABA and NFLD in the study area. Black segments indicate the proportion of the indicated haplotype in the pooled population sample of each primary river.

scale, and the naming system was readily applicable to Baltic data (Nilsson *et al.* 2001). The haplotype nomenclatures used by King *et al.* (2000), Verspoor *et al.* (1999) and Consuegra *et al.* (2002) are different, but all of them can be translated to the same system via the published DNA sequences. Haplotype diversity was estimated by $H = [n/(n - 1)](1 - \sum p^2)$. Conventional F_{ST} statistics based on haplotype frequencies were used to estimate levels of population differentiation. Haplotype diversity and population structure were analyzed using Arlequin 2.000 (Excoffier *et al.* 1992).

Results

Geographical structure of populations

The most common haplotypes AABA (N = 615, overall frequency 35.0%) and BBBB (N = 744, 42.3%) were found in all geographical segments of the coastline, in variable frequencies. Other haplotypes had less continuous distributions. The haplotype marked as NFLD (Newfoundland, N = 64, 3.6%) was restricted to the region between the Tana #28, and the Pulon'ga/Kola #15, which is situated in the narrowest pass of the Gorlo strait (White Sea strait). This haplotype represents the divergent phylogenetic clade of Atlantic salmon, which is most common in North America. The haplotype BBBA (N = 306, 17.4%) was common in populations on the Atlantic side of the continuum, found in all rivers along the coast from the Tana #28 to the Umba #12. The only outliers were detected from the Megra #2, which is situated on the southeastern coast in the Gorlo strait, less than 60 km from the opposite shore.

The cartographic visualization of haplotype frequencies (Fig. 2) and the pairwise similarity matrix suggested a division of populations into three regions, and the partitioning was tested by F_{ST} analysis (Table 2). The Barents West–Kola group border was set between the Drozdovka

#20 and the Kachkovka #19. This coincides with the navigation border of the White Sea, which is the line connecting Svyatoi Nos and Kanin Nos (Berger *et al.* 2001). The border between Kola group and the "deep" White Sea group was placed between the Pila #11 and the Umba #12, suggested by the haplotype frequencies. The "White Sea group" therefore included the populations on Karelian and Arkhangelsk coast of the White Sea. In statistical analyses, also the Pechora #1 was included in this group. AMOVA results presented in Table 2 indicated that the partitioning of the populations into these three groups, following the order of rivers along the coast, explained 22.1% of the variation ($P = 0.0000$). Only 9.9% of the variation then remained as variation among populations within groups (also highly significant), and 68.0% remained as variation within populations. Thus, the geographical structuring was very strong. We also tested the structure by equalizing the sample sizes. By using this balanced procedure, the variance components were 21.2%, 16.5% and 62.5%, respectively. The increase of intragroup component is mainly due to overemphasis of small monomorphic samples. The optimal division of the populations into three groups was also tested by moving the group borders (data not shown).

Pairwise comparisons of population groups showed that Barents West and White Sea groups were genetically more similar to each other ($F_{ST} = 0.091$) than they were with Kola group which is situated between them ($F_{ST} = 0.273$ in both comparisons (Table 2)).

The region "Barents Sea West" included 718 fish from the nine westernmost rivers, from the Tana #28 to the Drozdovka #20, and it was characterized by the highest proportion of the ubiquitous BBBB haplotype (53%), a high proportion of the BBBA haplotype (27%) and the presence of the North American NFLD haplotype in low but consistent frequencies. The haplotype diversity in this region was the highest, 0.633 and the mean pairwise F_{ST} between the populations within the Barents West group was 0.032 (Table 2).

The next region was formed by eight populations along the eastern and southern coasts of the Kola Peninsula, from the River Kachkovka #19 to the River Umba #12. Characteristic for the region was the high frequency of AABA (65%), which is the haplotype showing the most "eastern" distribution. The haplotype diversity of this group was 0.530 and the mean F_{ST} between populations was 0.081, indicating that this region was less uniform than the neighboring regions. Indeed, the samples Ponoi #18 and Kachkovka #19 differed significantly from the other populations in this group, showing transient frequencies of the haplotypes.

Near the westernmost point of the White Sea, between the Umba #12 and Pila #11 rivers there was a qualitative change in haplotype ratios. The haplotype BBBA disappeared. This third region of nine populations was given the name White Sea, including the rivers on the Karelian coast and the samples from the Arkhangelsk region. The region begins from the River Pila #11 and continues along the coast to the southern part of

Table 2. Hierarchical AMOVA over population groups and populations. $H = [n/(n-1)](1 - \Sigma p^2)$ was used as diversity index; and mean divergence between populations was estimated by F_{ST}.

Comparison	Sample N	Number of haplotypes	H	N samples	F_{ST}	% of variation
Among 3 groups				3	0.221	22.1
Barents/White Sea				2	0.091	
Barents/Kola				2	0.273	
Kola/White Sea				2	0.272	
Among populations within groups				27	0.320	9.9
Barents W	718	8	0.632	8	0.075	
Kola	577	8	0.529	8	0.096	
White Sea	464	3	0.463	11	0.274	
Within populations						68.3
Total	1759	10	0.667	27		

the White Sea (Suma #5), and further to the Severnaya Dvina #4 and the Megra #2. Haplotype diversity in this region was lowest of all, 0.464, and only two haplotypes, AABA and BBBB were present. The mean pairwise F_{ST} between the Karelian and Arkhangelsk populations was 0.015, i.e., one tenth of the overall mean F_{ST} between populations (0.154). In the AMOVA, the small Pechora samples were included in this group; their contribution is, however, small.

The mean pairwise F_{ST} between nearest neighbour populations was 0.070, while the overall mean pairwise F_{ST} was 0.154. Within the population groups, the pairwise F_{ST} distances were between 0.015 and 0.081. The mean geographic distance of the populations was 616 km, when roughly measured by additive direct lines parallel to the coast from river mouth to river mouth. The regression coefficient of pairwise F_{ST} versus the geographic distance between populations was low and not statistically significant in the total data ($b = 0.00003$), nor within any region.

Discussion

Nature of the mitochondrial variation in northwestern Russian salmon populations

The number of variable restriction sites is too low for us to be able to use the genetic distances between the haplotypes in analyses, and the evolution of the haplotypes probably predates the Weichselian geological events in the area. Thus, the effect of latest glaciation cycle is suggested to be only lineage sorting by random drift, and postglacial development is slowly mixing the haplotypes again. There is a strong geographical structure, but because it is not formed by unimodal factors, but caused by heterogeneous origins of the populations, the differentiation is not a unimodal or linear function of distance. No equilibrium between founder effects, migration and drift has been reached, and therefore, standard Nm values (number of migrants, Hudson *et al.* 1992) between populations or regions have not been estimated.

The partitioning of populations into three regions, corresponding strictly and in linear order with the segments of coastal line, explained 22% of the genetic variation in AMOVA. This division was *ad hoc*, but it avoids producing artificial combinations of geographically distant populations, which by random drift or by other reasons happen to be similar. For more explicit groupings, a number of additional loci are needed. The high proportion of the explained molecular variance supports the division made.

A key haplotype border zone is found in the narrowest pass of the White Sea, the Gorlo strait. The River Megra #2, on the Arkhangelsk coast of the Gorlo strait, is located only 60 km away from neighbors on the opposite coast, but the haplotype composition of the population is more concordant with neighbors on the same southern White Sea coast hundreds of kilometers away. Only four fish with Atlantic BBBA type were found in the Megra #2 (7%), while the frequency of BBBA was 45% in the Ponoi #18 and 15% in the Babya #17 on the opposite shore. Females thus seldom stray over the strait, and the shoreline and/or the sea currents may serve as homing cues (if the White Sea group fish migrate out of the Gorlo strait at all!). The current of warm and less salty (24‰) water out from the White Sea occurs along the eastern coast, and the inward current brings colder and saltier (29‰) water along the Kola coast (Berger *et al.* 2001). Such differences vary seasonally, but certainly there appear to be enough indicators for female salmon to find their natal river. The sea-time distribution of White Sea salmon has not been studied.

The geographical (regional) genetic structuring revealed in this study was more pronounced than in earlier studies on Atlantic salmon (the Baltic and West–East Atlantic differences excluded). Nielsen *et al.* (1996) reported a variance component of only 0.2% among the three large areas: Swedish west coast, Scotland and Ireland. As a contrast, the variance among populations within groups in the North Sea–Atlantic was much higher, 20.3%. There are several hypothetical explanations for this difference in geographical structuring. The first may be that the populations studied by Nielsen *et al.* (1996) were all smaller and much more disturbed, which may sweep away the original postglacial structuring and increase the population differentiation by random drift. The total haplotype diversity in the compound North Sea sample was $H = 0.648$, i.e., lower, but not significantly so, than the Northwest Russian total $H = 0.667$. The high overall diversity was understandable: the North Sea coasts and the British Isles are the central meeting region, connecting all possible postglacial colonizations into a diverse, if not homogeneous mixture. Another, simple explanation might be that the populations in the North Sea and the British Isles are all derived from relatively large and therefore variable marine refugium.

The mean genetic differentiation (F_{ST}) between pairwise nearest neighbor populations in the NW Russian data set was 0.070 (Table 2). The mean F_{ST} divergence between populations within the regions was 0.015 among the White Sea group, 0.032 in the Barents W group and 0.081 among the Kola group. This emphasizes the success of grouping, because the mean F_{ST} among all pairwise population comparisons was 0.154, ten times higher than in the least diverse White Sea group. The highest haplotype diversity was found in the Barents W group (pooled $H = 0.633$). The western Barents Sea coast is at the crossroads of the eastern and western expansions. The western influence was diminished in the Kola group ($H = 0.530$) and completely missing in the White Sea population group, which is suggested to represent the original colonization from the Komi Ice Lake (Karelian $H = 0.463$, only two haplotypes present, excluding the four occurrences of BBBA in the Megra #2).

The haplotype BBBA is a qualitative marker of the marine Atlantic faunal component. In the Danish straits connecting the Atlantic and the Baltic Sea, the BBBA haplotype is found near the narrowest point, in the River Lagan, but not further in the Baltic (Nilsson *et al.* 2001). In the White Sea, a similarly abrupt border is found between the Umba #12 and the Pila #11, their river mouths being less than 15 km apart. Further into less salty water, BBBA is almost completely absent (Fig. 2). Another qualitatively oceanic and western marker is the phylogenetically distant haplotype NFLD, only found on the western Barents coast and in North America. In addition, the haplotype AAAA ($N = 20$, 1.1%) was rare in northwestern Russia. This haplotype was predominant in the River Skjern, in Denmark (Nielsen *et al.* 1996).

A number of rare haplotypes, some of them not previously described in any parts of Europe, were observed in this study. The haplotypes AABB, AEAA, BBAA, and BBAB were each observed in two individuals. They were all observed in the western Barents and Kola coastal segment characterized by the presence of BBBA, and perhaps also reflect the Atlantic contribution to the high genetic diversity of the Barents area. Of those rare types, ADBA has been reported earlier along the Swedish west coast (Nilsson *et al.* 2001).

Elo *et al.* (1994) compared two of the Barents rivers, the Tana #28 and the Neiden #27, using allozyme polymorphism. Only 0.3% of the genetic diversity was attributed to the difference between the two rivers, 2.6% to tributary systems within drainages, 3.5% to sampling localities within the tributaries, and 93.6% to sampling locations. The authors explained the results as indicative for local adaptation. In the framework of the present study, it seems that the local adaptation is indeed possible, at least in a more coarse grained environment. Female haplotype distribution in our large study area indicates more structuring than expected on the basis of the study of Elo *et al.* (1994).

Transatlantic migration

The distribution of the haplotype named here as NFLD (from "Newfoundland") was interesting. The haplotype represents the highly diverged North American clade in the salmon phylogeny (Nilsson *et al.* 2001). Total number of "North American" haplotypes registered among ~6200 European salmon is now 64, and they were all found in the study area of the present paper, along the Barents Sea coast, between the Tana #28 and the Pulon'ga/Kola #15 rivers. Information from further west and south along the Norwegian coast is scarce or missing, but the "North American" haplotypes have not been found along the west coast of Sweden, nor in the British Isles and Iceland (Nielsen *et al.* 1996, Verspoor *et al.* 1999). Verspoor *et al.* (1999) recorded the North American mitochondrial haplotype in the River Pecha (a Tuloma tributary). Curiously, they observed two RFLP types "NANAA" (haplotype VI, 11.4%) and "NANAB" (V, 4.5% =

2 individuals), separated by a *Rsa*I restriction site. The restriction enzyme *Rsa*I was used in the present study as well, without detecting this polymorphism among the NFLD haplotypes.

It is tempting to conclude that the colonization from North America to the Barents coast is postglacial, but not recent. The maternally inherited mitochondrial marker has had time to disperse more than a thousand kilometers along the coast. Further evidence of the considerable age of the colonization comes from the nuclear allozyme marker of the predominantly North American distribution, ESTD*92. Kazakov and Titov (1991, 1993) recorded this slow allele in many rivers on the Kola Peninsula and in the White Sea. Skaala *et al.* (1998) confirmed this allele, and the same group of authors stated that it is almost fixed (diagnostic) in North American populations (Makhrov *et al.* 1998). The frequencies of ESTD*92 were higher than the frequencies of the NFLD haplotype of mitochondria, up to 51% in the Ponoi #18. ESTD*92 allele goes also deeper into the White Sea — the Umba #12, 24%, the Keret 17%. The River Keret is between the Pulon'ga/Karelia #8 and the Pon'goma #7 on the Karelian coast. Thus, the geographical distributions of nuclear and mitochondrial markers of immigrants have already diverged, indicating also a difference in female and male migratory behavior.

It is possible that the original NFLD female fish joined the Barents fish when feeding together on the west Greenland banks (Bermingham *et al.* 1991). Whether the gene flow between North America and the Barents coast is continuous or historical, remains to be studied. The observations presented here suggest that it was a lucky, single long distance colonization event, and offers a challenge and possibility for more accurate timing.

Reciprocally, haplotypes phylogenetically belonging to the "European" clade occur in Canadian rivers. King *et al.* (2000) reported fifteen individuals with a BBBA haplotype (36%) in the River Gander, in Newfoundland. Bermingham *et al.* (1991) detected "European" haplotypes in the landlocked population in Gumbo, and in Nilsson *et al.* (2001), the 16SrDNA/*ND1* sequences of three Ouananiche Brook (Newfoundland, Canada) dwarfed salmon were confirmed by sequencing to be identical with corresponding "European" haplotypes (BBBA, BBBB). Evidently, some of the "European" clade salmon have been isolated thousands of years in Northern America, perhaps since deglaciation.

Postglacial colonization from an eastern refugium: a hypothesis confirmed

The hypothesis of the north-eastern refugium for the easternmost Atlantic salmons was suggested on the basis of the allozyme polymorphisms by Kazakov and Titov (1991, 1993). A candidate for this refugium is the early Weichselian proglacial Komi Ice Lake, which was dammed by the Barents Ice Sheet about 90 000 BP (Mangerud *et al.* 2001) and which most probably had an outlet eastwards of the present lake Onega (Maslennikova & Mangerud, 2001) to the dislocated and freshwater Baltic [Sea] Ice Lake. The latter is perhaps the more southern freshwater refugium proposed as the origin for the salmon populations in the southern Baltic Sea (Koljonen *et al.* 1999, Nilsson *et al.* 2001).

The salmon populations in lake Onega, lake Ladoga and in the southern Baltic Sea differ from the White Sea populations by being monomorphic for the haplotype AABA (Nilsson *et al.* 2001). The haplotype AABA was even named as "Baltic" in earlier studies (Bermingham *et al.* 1991). The White Sea segment of the northern coast of Russia was populated by BBBB and AABA haplotypes, which are also found in the few adjacent landlocked relict populations in Russian Karelia (Nilsson *et al.* 2001). These two haplotypes are also the only ones found in the rivers draining from the western part of the postulated Komi Ice Lake, namely the Severnaya Dvina #4 and Pechora #4 (Maslennikova & Mangerud 2001, Mangerud *et al.* 2001).

While there was a strong Atlantic component in populations along the east and south coasts of the Kola Peninsula, the proportion of the eastern haplotype AABA was much higher here than in other regions. This might indicate that the primary colonization of this area occurred from the east, from populations released to the eastern Barents Sea from the Komi Ice Lake perhaps 20 000 BP (Fig. 1). During the late Pleistocene glacial maximum, the eastern Barents Sea was free of ice, and the Pechora #1 and neighboring river populations had an access to the sea again (Svendsen *et al*. 1999). Thus, the colonization of the Kola Peninsula may have occurred along the ice-free sea. The Atlantic immigration occurred later. Most Norwegian rivers are quite short, and the ice cap stayed the longest (until 6000 BP) in the Scandinavian mountains. This may have slowed down the step-by-step immigration from southern Atlantic refugia, when the demographic pressure from the North Sea coasts and the British Isles was still moderate.

The samples from the Pechora #1 and the Severnaya Dvina #4 analyzed in this study were obtained from spawning grounds far from the river mouths. The tributaries of the Severnaya Dvina #4 are the Padoma (900 km from the sea), the Pyshentsa (670 km) and the Niukhcha (680 km). The tributaries of the Pechora #1 are the Un'ia (1750 km from the sea) and the Pizhma (> 600 km). The samples are small, but it is suggested that they represent the homing descendants of original refugial populations. The lower parts of these long rivers flow through lowlands formed by fine-grained ice lake sediments, offering no spawning sites at all.

The best evidence to support the hypothesis of a refugial ice lake origin is the purity (= low diversity) of the populations on the western and southern coasts of the White Sea and in the Severnaya Dvina #4 and the Pechora #1. The Atlantic maternal lineages have not yet arrived in this historically isolated area. It remains to be studied which ecological and behavioral factors have prevented the introgression. Perhaps the salinity in river mouths guides the homing female spawners so precisely that this geographic structure has persisted for thousands of years.

Acknowledgements

The help and support from the Northern Russian local Federal State Boards of Fisheries — Murmanrybvod and Karelrybvod — and their heads, Boris Prishchepa and Vyatcheslav Movchan is gratefully acknowledged. During the field expeditions, the help of Dimitry Stepanov, Jussi Kuusela, Eero Lumme, Mikko Koskinen, Dimitry Sendek and Marek Ziętara was of great value. Kaj Kujala supplied some of the Tana samples, and the Finnish Game and Fisheries Research Institute those from Neiden/Näätämö and Kola. The technical laboratory help of Svetlana Riannel, Minna Rinta-Pukkila, Anita Ala-Poikela, Hannele Syri and Sanna Pottala is acknowledged. The Finnish Academy supported the work (Grants 44124, 63797 & 172964), as did the Finnish Ministry for Agriculture and Forestry, and Wihuri Foundation (to T.A.). We dedicate this paper to the memory of George V. Solomonov, a good fellow and a member of our first Karelian expedition team in 1999.

References

Berger, V., Dahle, S., Galaktionov, K., Kosobokova, X., Naumov, A., Rat'kova, T., Savinov, V. & Savinova, T. 2001: *White Sea. Ecology and environment*. — Derzavets Publisher. St. Petersburg–Tromsø.

Bermingham, E., Forbes, S. H., Friedland, K. & Pla, C. 1991: Discrimination between Atlantic salmon (*Salmo salar*) of North American and European origin using restriction analyses of mitochondrial DNA. — *Can. J. Fish. Aquat. Sci*. 48: 884–893.

Consuegra, S., Garcia de Leaniz, C., Serdio, A., Gonzáles Morales, M., Straus, G., Knox, D. & Verspoor, E. 2002: Mitochondrial DNA variation in Pleistocene and modern Atlantic salmon from the Iberian glacial refugium. — *Molecular Ecology* 11: 2037–2048.

Cutler, M. G., Bartlett, S. E., Hartley, S. E. & Davidson, W. S. 1991: A polymorphism in the ribosomal RNA genes distinguishes Atlantic salmon (*Salmo salar*) from North America and Europe. — *Can. J. Fish. Aquat. Sci*. 48: 1655–1661.

Elo, K. 1993: Gene flow and conservation of genetic variation in anadromous Atlantic salmon (*Salmo salar*). — *Hereditas* 119: 149–159.

Elo, K., Vuorinen, J. A. & Niemelä, E. 1994: Genetic resources of Atlantic salmon (*Salmo salar* L.) in Teno and Näätämö rivers, northernmost Europe. — *Hereditas* 120: 19–28.

Hudson, R. R., Slatkin, M. & Maddison, W. P. 1992: Estimation of gene flow from DNA sequence data. — *Genetics* 132: 583–589.

Kazakov, R. V. & Titov, S. F. 1991: Geographical patterns in the population genetics of Atlantic salmon, *Salmo salar* L., on U.S.S.R. territory, as evidence for colonization routes. — *J. Fish Biol*. 39: 1–6.

Kazakov, R. V. & Titov, S. F. 1993: Population genetics of salmon, *Salmo salar* L., in northern Russia. — *Aquaculture and Fisheries Management* 24: 495–506.

King, T. L., Spidle, A. P., Eackles, M. S., Lubinski, B. A. & Schill, W. B. 2000: Mitochondrial DNA diversity in North American and European salmon with emphasis on the Downeast rivers of Maine. — *J. Fish Biol*. 57: 614–630.

Koljonen, M.-L., Jansson, H., Paaver, T., Vasin, O. & Koskiniemi, J. 1999: Phylogeographic lineages and differentiation pattern of Atlantic salmon (*Salmo salar*) in the Baltic Sea with management implications. — *Can. J. Fish. Aquat. Sci*. 56: 1766–1780.

Makhrov, A. A., Skaala, O., Altukhov, Y. P. & Saunders, R. L. 1998: The allozyme ESTD* locus as a marker of genetic differentiation between European and North American populations of the Atlantic salmon (*Salmo salar* L.) — *Doklady Biological Sciences* 360: 281–283. Interperiodica Publishing.

Mangerud, J., Astakhov, V., Jakobsson, M. & Svendsen, J. I. 2001: Huge ice-age lakes in Russia. — *Journal of Quaternary Science* 16: 773–777.

Maslenikova, O. & Mangerud, J. 2001: Where was the outlet of the ice-dammed Lake Komi, Northern Russia? — *Global and Planetary Change* 31: 337–345.

McConnell, S. K., O'Reilly, P., Hamilton, L., Wright, J. M. & Bentzen, P. 1995: Polymorphich microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North American and European populations. — *Can. J. Fish. Aquat. Sci*. 52: 1863–1872.

Meinilä, M., Kuusela, J., Ziętara, M. S. & Lumme, J. 2004: Initial steps of speciation by geographic isolation and host switch in salmonid pathogen *Gyrodactylus salaris* (Monogenea: Gyrodactylidae). — *International Journal for Parasitology*: 34: 515–526.

Nielsen, E. E., Hansen, M. M. & Loeshke, V. 1996: Genetic structure of European populations of *Salmo salar* L. (Atlantic salmon) inferred from mitchondrial DNA. — *Heredity* 77: 351–358.

Nielsen, E. E., Hansen, M. M. & Mensberg, K.-L. D. 1998: Improved primer sequences for the mitochondrial ND1, ND3/4 and ND5/6 segments in salmonid fishes: application to RFLP analysis of Atlantic salmon. — *J. Fish Biol*. 53: 216–220.

Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., Kohlman, K. Löytynoja, A., Nielsen, E. E., Paaver, T., Primmer, C. R., Titov, S., Vasemägi, A., Veselov, A., Öst, T. & Lumme, J. 2001: Matrilinear phylogeography of Atlantic salmon (*Salmo salar* L.) in Europe and postglacial colonization of the Baltic Sea area. — *Molecular Ecology* 10: 89–102.

Skaala, O, Makhrov, A. A., Karlsen, T., Jorstad, K. E., Altukhov, Y. P., Politov, D.-V., Kuzishin, K. V. & Novikov, G. G. 1998: Genetic comparison of salmon from the White Sea and north-western Atlantic Ocean. — *J. Fish Biol*. 53: 569–580.

Ståhl, G. 1987: Genetic population structure of Atlantic salmon. — In: Ryman, N. & Utter, F. (eds.), *Population genetics and its application to fisheries managament*: 121–140. University of Washington Press, Seattle, WA.

Svendsen, J. I., Astakhov, V. I., Bolshiyanov, D. Yu., Demidov, I., Dowdeswell, J. A., Gataullin, V., Hjort, C., Hubberten, H. W., Larsen, E., Mangerud, J., Melles, M., Möller, P., Saarnisto, M. & Siegert, M. 1999: Maximum extent of the Eurasian ice sheet in the Barents and Kara Sea region during the Weichselian. — *Boreas* 28: 234–242.

Verspoor, E., McCarthy, E. M., Knox, D., Bourke, E. A. & Cross, T. F. 1999: The phylogeography of European Atlantic salmon (*Salmo salar* L.) based on RFLP analysis of the ND1/16sRNA region of the mtDNA. — *Biological Journal of Linnean Society* 68: 129–146.