Ann. Zool. Fennici 42: ¶–¶ ISSN 0003-455X
Helsinki ¶¶¶ 2005 © Finnish Zoological and Botanical Publishing Board 2005

Phylogeography of anadromous and non-anadromous Atlantic salmon (*Salmo salar*) from northern Europe

Anni Tonteri[1,2,3], Sergey Titov[4], Alexei Veselov[5], Alexander Zubchenko[6], Mikko T. Koskinen[1,7], David Lesbarrères[1], Svjatoslav Kaluzchin[8], Igor Bakhmet[5], Jaakko Lumme[2] & Craig R. Primmer[1,3]*

[1)] *Department of Biological and Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland*
[2)] *Department of Biology, P.O. Box 3000, FI-90014 University of Oulu, Finland*
[3)] *Current address: Department of Biology, FI-20014 University of Turku, Finland (*corresponding author's e-mail: craig.primmer@utu.fi)*
[4)] *State Research Institute on Lake and River Fisheries, Makarova 26, 199053 St. Petersburg, Russia*
[5)] *Institute of Biology, Karelian Research Centre, Pushkinskaya 11, 185610 Petrozavodsk, Russia*
[6)] *Knipovich Polar Research Institute of Marine Fisheries and Oceanography (PINRO), Knipovich 6, 183763 Murmansk, Russia*
[7)] *Current address: Finnzymes Diagnostics, Keilaranta 16, FI-02150 Espoo, Finland*
[8)] *Varzuga Scientific Research Centre of Ecology of Polar Ecosystems, Murmansk Region, Russia*

Received 18 Aug. 2004, revised version received 20 Oct. 2004, accepted 20 Oct. 2004

Tonteri, A., Titov, S., Veselov, A., Zubchenko, A., Koskinen, M. T., Lesbarrères, D., Kaluzchin, S., Bakhmet, I., Lumme, J. & Primmer, C. R. 2005: Phylogeography of anadromous and non-anadromous Atlantic salmon (*Salmo salar*) from northern Europe. — *Ann. Zool. Fennici* 42: ¶–¶.

The phylogeography of north European anadromous and non-anadromous Atlantic salmon (*Salmo salar*) populations was investigated using 21 nuclear (microsatellites and allozymes) loci and mitochondrial DNA haplotypes. A neighbour-joining population tree revealed several statistically supported groupings that generally corresponded well with the sampling regions. A comparison of F_{ST} and R_{ST} estimates with a novel allele size permutation method suggested that at least two of the groups had diverged from each other already prior to the ice receding after the last ice age, thus suggesting that north European Atlantic salmon are derived from at least two separate refugia. We propose that the anadromous and non-anadromous salmon populations from the Baltic Sea basin most likely originate from a southeastern ice-lake refugium. The present day White and Barents Sea basins have probably been colonized from multiple refugia.

Introduction

The last glaciation, dating from ~ 115 000 to 10 000 years ago (Andersen & Borns 1994) had a great influence on the biodiversity of northern Europe. As vast ice masses covered the north European region during this period, current species of the region originate from ancestors formerly resident in non-glaciated areas in the south of the continent. Comparative phylogeography surveys have identified major European refugia for terrestrial species in the

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Apennine, Iberian, and Balkan Peninsulas as well as in southwest Russia (Hewitt 1996, 1999, Taberlet *et al*. 1998). Considering aquatic species, refugia contributing to current-day freshwater fauna have been identified in central and eastern Europe (Nesbø *et al*. 1999, Koskinen *et al*. 2000, 2002, Kontula & Väinölä 2001), while marine and anadromous (sea migrating) fishes most likely have different histories due to their differing environmental requirements. Species with alternative life-history strategies, like Atlantic salmon (*Salmo salar*) and brown trout, (*S. trutta*) are especially intriguing and difficult to understand.

The Atlantic salmon (*Salmo salar* L.) is a spec ies that has re-colonised north European waters following the last ice age. Nowadays the European distribution area of anadromous Atlantic salmon ranges from northern Portugal to the Pechora River in northwest Russia, including Iceland, the British Isles, and the Baltic Sea (Parrish *et al*. 1998). In addition, there are non-anadromous (land-locked) salmon living in lakes both in Europe and in North America. These land-locked populations were formed as they became isolated from the sea or lake by rapids and waterfalls created by fast post-glacial land-upheaval (Berg 1985). The Atlantic salmon is globally substructured into genetically differentiated populations (Ståhl 1987). Using various classes of genetic markers, it has been demonstrated that there is a clear division between the North American and European salmon populations (Ståhl 1987, Bermingham *et al*. 1991, McConnell *et al*. 1995a, 1995b, Verspoor *et al*. 1999, King *et al*. 2001). European salmon are further divided into two major groups: the Atlantic and the Baltic salmon (e.g. Ståhl 1987, Bermingham *et al*. 1991, Bourke *et al*. 1997, Verspoor *et al*. 1999, Nilsson *et al*. 2001).

Despite numerous studies, the post-glacial origin of north European Atlantic salmon is still debated. Considering Baltic Sea salmon, both Atlantic (Verspoor *et al*. 1999) and eastern freshwater (Nilsson *et al*. 2001) refugia have been proposed, as well as combinations of the two (Koljonen *et al*. 1999). Similarly for salmon of northwestern Russia and the White Sea basin, refugial populations in the eastern Barents Sea area have been proposed (Kazakov & Titov 1991, Asplund *et al*. 2004), but evidence for later immigration from the Atlantic also exists (Asplund *et al*. 2004).

There are several potential reasons for this lack of concordance. It seems clear that there are several sources of postglacial re-colonisation, and the relative contribution of each is difficult to quantify. Further, some studies suffered from insufficient sampling of some north European regions, with dense coverage of northwest Russian populations being particularly rare (but *see* Kazakov & Titov 1991, Asplund *et al*. 2004). Earlier studies of north European Atlantic salmon phylogeography have generally relied on data from a single class of molecular marker: allozymes (Kazakov & Titov 1991, Koljonen *et al*. 1999) or mtDNA (Verspoor *et al*. 1999, Nilsson *et al*. 2001, Asplund *et al*. 2004). Therefore, inconsistent conclusions may be due to the fact that these markers have different modes of inheritance and differing mutation and divergence rates which, if information from different marker types is considered in isolation, may cause conflicting conclusions when determining whether the observed genetic structuring was created before or during the period of ice-induced isolation, or following post-glacial recolonization.

The analysis of multiple classes of molecular markers, in combination with dense and extensive sampling coverage, may provide a resolution to some of the difficulties outlined above. Firstly, combined analysis of nuclear markers with differing mutation rates such as microsatellites and allozymes may help to provide a clearer picture of population relationships as the different marker types may resolve relationships over different evolutionary time scales. Secondly, analysis of microsatellite data may be particularly useful for estimating the number of post-glacial refugia that contributed to the re-colonisation of northern Europe. This is due to the timing of the re-colonisation event, which occurred ~ 10 000 years, or 2500 Atlantic salmon generations, ago. This generation number is precisely the divergence time above which microsatellite mutations are expected to contribute to the genetic divergence of populations (Estoup & Angers 1998). Therefore, mutation should only have contributed to the microsatellite divergence between

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

north European salmon populations colonised from different post-glacial refugia. Whether or not mutation has contributed significantly to the genetic divergence between two groups of populations can be statistically tested by a recently developed allele size permutation test (Hardy *et al*. 2003). Hence, the method can be applied to test whether north European salmon populations from a particular region originate from one or more refugia. In this study, a combination of microsatellite, allozyme and mtDNA variation was used to infer the refugial Atlantic salmon lineages that contributed to the post-glacial colonisation of current day north European populations.

Material and methods

Sampled populations

A total of 901 Atlantic salmon individuals from 15 anadromous and eight non-anadromous north European populations were analysed (Table 1 and Fig. 1). The sampled individuals were either adult or juvenile salmon, which were caught by electrofishing, fly-fishing, drift nets or permanent traps between the years 1988 and 2000. Fin samples were preserved in 95% ethanol for DNA extraction while tissues for allozyme analysis were snap frozen in liquid nitrogen. For the majority of samples, DNA was isolated using the salt extraction method of Aljanabi and Martinez (1997).

Microsatellite data

Initially, seventeen microsatellite markers were screened: *Ssa14, Ssa289* (McConnell *et al*. 1995b), *SSOSL85, SSOSL311* (Slettan *et al*. 1995), *Ssa85, Ssa171, Ssa197, Ssa202* (O'Reilly *et al*. 1996), *SSOSL438* (Slettan *et al*. 1996), *SLEEI84, SLEEN82* (**unpublished**, GenBank accession numbers U86703 and U86706, respectively), *SS11* (Martinez *et al*. 1999), *Ssa412, Ssa422* (Cairney *et al*. 2000), *SS20.19, SSD30,* and *SSF43* (Sánchez *et al*. 1996). A new reverse primer was designed for *Ssa412* (5´-GTT TCT TGG TTA GTA CCG GAC ATG-3´) in order to obtain a longer PCR product. Of these seventeen loci, three were excluded from further analyses due to the amplification of a duplicated locus in some populations (*SS11*), due to the occurrence of null-alleles (*SS20.19*), or due to tight physical linkage with another locus in the dataset (*Ssa289*: linkage distance to *Ssa422* < 10 cM (Gharbi 2001). To facilitate genotyping through enhancing 3´ adenylation, a GTTT "PIGtail" (Brownstein *et al*. 1996) was added to the 5´ end of each non-labelled primer.

The 10 μl amplification reactions consisted of ca. 100 ng of extracted DNA, 0.1 to 0.5 μM of each primer, one of which was fluorescently labelled (Table 2), 1 × NH_4 reaction buffer (16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8), 0.01% Tween-20, Bioline), 1.5 mM $MgCl_2$, 250 μM dNTP and 0.1 U of BioTaq DNA polymerase (Bioline). In some cases, multiplex PCRs were also utilised. These reactions were otherwise the same except that some primer concentrations were modified (Table 2). All amplifications were carried out on a PTC100, PTC200 (MJ Research) or Mastercycler gradient (Eppendorf) thermal cyclers. The general PCR protocol used started with a 3-minute denaturation at 94 °C followed by 35 cycles of denaturation at 94 °C for 30 seconds (s), annealing at **x °C** (Table 2) for 30 s, and extension at 72 °C for 30 s. The protocol ended with a 5-minute final extension at 72 °C. For *SSOSL311* each repeated step was 60 seconds long.

In order to enable simultaneous electrophoresis of several markers in a single gel lane, the markers were divided into two groups (Table 2). The loci within a group were pooled together in different amounts in order to produce fluorescent signals of similar intensity (Table 2). The denaturated samples were electrophoresed on an ABI Prism™ 377 genetic analysis instrument along with the GeneScan 400HD ROX size standard (Applied Biosystems). GeneScan 3.1.2 and Genotyper 2.5 (Applied Biosystems) were used to analyse the DNA fragments and to score the genotypes. All genotypes were inspected visually and then data were exported to a spreadsheet program for further analysis. In total, 901 individuals were successfully genotyped with at least 12 of the 14 microsatellite markers and were hence included in further statistical analyses.

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Table 1. Details and genetic diversity indices of the Atlantic salmon populations included in the study. Populations from river Neva and lake Saimaa are of hatchery origin. Abbr. — the abbreviation of the population name. *N* — number of individuals. *A* — Average number of alleles in a population. A_r — Allelic richness. H_o — observed heterozygosity. H_e — expected heterozygosity.

Sample populations					Microsatellite diversity					Allozyme diversity		
Population	Abbr.	Basin	Coordinates	Migration behaviour	*N*	*A* (range)	A_r	H_o	H_e	*N*	*A*	H_e
Dee*	Dee	Atlantic Ocean	56°54´N, 3°27´W	Anadromous	48	10.6 (3–23)	6.9	0.72	0.73	47	1.9	0.19
Teno	Ten	Atlantic Ocean	70°30´N, 28°25´E	Anadromous	46	9.7 (2–20)	6.3	0.66	0.71	85	2.0	0.23
Tuloma	Tul	Atlantic Ocean	68°41´N, 31°55´E	Anadromous	42	9.1 (2–16)	6.1	0.67	0.72	–	–	–
Pizhma	Piz	Barents Sea	64°53´N, 51°17´E	Anadromous	21	5.5 (2–10)	4.7	0.62	0.62	50	1.6	0.23
Unja	Unj	Barents Sea	61°32´N, 58°15´E	Anadromous	11	4.1 (2–7)	4.0	0.62	0.54	50	1.7	0.18
Megra	Meg	White Sea	66°03´N, 41°43´E	Anadromous	48	7.8 (2–15)	5.5	0.68	0.66	60	2.0	0.30
Nilma	Nil	White Sea	66°27´N, 33°05´E	Anadromous	35	4.5 (2–7)	3.6	0.61	0.55	43	1.9	0.27
Pongoma	Pon	White Sea	65°18´N, 34°02´E	Anadromous	44	6.6 (2–12)	4.9	0.67	0.66	51	2.0	0.20
Pulonga	Pul	White Sea	66°18´N, 33°17´E	Anadromous	40	5.4 (1–9)	4.1	0.57	0.57	51	1.9	0.30
Suma	Sum	White Sea	64°14´N, 35°25´E	Anadromous	38	5.0 (2–11)	4.1	0.61	0.58	52	1.9	0.23
Kitsa	Kit	White Sea	66°25´N, 36°54´E	Anadromous	45	8.4 (2–17)	5.4	0.70	0.67	42	2.0	0.27
Varzuga	Var	White Sea	66°36´N, 36°35´E	Anadromous	47	8.4 (2–16)	5.3	0.61	0.65	29	2.0	0.33
Luzhma	Luz	White Sea	63°13´N, 33°18´E	Non-anadromous	40	3.9 (1–8)	3.2	0.47	0.45	36	1.4	0.14
Kamennoe	Kam	White Sea	64°28´N, 30°26´E	Non-anadromous	41	2.7 (1–5)	2.4	0.38	0.35	61	1.3	0.07
Pisto	Pis	White Sea	65°16´N, 30°35´E	Non-anadromous	53	3.6 (1–7)	2.9	0.43	0.41	58	1.4	0.08
Neva	Nev	Baltic Sea	59°58´N, 30°13´E	Anadromous	43	6.4 (2–12)	4.6	0.63	0.60	50	1.7	0.21
Tornio†	Tor	Baltic Sea	65°49´N, 24°08´E	Anadromous	42	7.4 (3–14)	4.9	0.59	0.59	644	2.3	0.15
Vindelälven‡	Vin	Baltic Sea	63°50´N, 20°05´E	Anadromous	47	4.8 (2–12)	3.6	0.47	0.48	100	1.4	0.10
Lizhma	Liz	Baltic Sea	62°25´N, 34°27´E	Non-anadromous	26	3.9 (2–8)	3.2	0.54	0.51	34	1.6	0.17
Saimaa	Sai	Baltic Sea	63°19´N, 30°01´E	Non-anadromous	44	2.5 (1–5)	2.2	0.29	0.28	90	1.4	0.08
Shuja	Shu	Baltic Sea	61°51´N, 34°09´E	Non-anadromous	20	4.4 (2–10)	3.8	0.57	0.55	34	1.6	0.18
Sysky	Sys	Baltic Sea	61°39´N, 31°16´E	Non-anadromous	42	4.3 (1–12)	3.4	0.49	0.47	42	1.6	0.22
Taipale	Tai	Baltic Sea	60°38´N, 30°30´E	Non-anadromous	38	6.3 (2–14)	4.5	0.59	0.57	39	1.9	0.19

* Allozyme data from Jordan *et al.* 1992, except for *ESTD-2** and *IDDH-1** E. Verspoor, personal comm.
† Allozyme data from Koljonen *et al.* 1999, except for *ESTD-2** Bourke *et al.* 1997
‡ Allozyme data from Koljonen *et al.* 1999, except for *ESTD-2** (*see* text for details)

Date and signature

Allozyme data

Seven allozyme loci were utilized in the study: *AAT-4**, *ESTD-2**, *IDDH-1**, *IDDH-2**, *IDHP-3**, *MDH-3**, and *MEP-2**. Sample preparation and electrophoresis were performed as described in Kazakov and Titov (1998). Tissue samples suitable for allozyme analysis were not available for the same individuals as analysed with microsatellites in the Tuloma, Dee, Tornio, and Vindelälven populations, hence, allozyme allele frequencies from earlier studies were used for these populations (Jordan *et al*. 1992, Bourke *et al*. 1997, Koljonen *et al*. 1999, E. Verspoor, pers. comm.). Published details of allele frequencies for the locus *ESTD-2** were not available for Vindelälven. Therefore, the locus was assumed to be monomorphic in this population since all the other studied rivers in the Baltic Sea are monomorphic for this locus (Bourke *et al*. 1997). As no allozyme data were available for Tuloma, this population was excluded from analyses, which included allozyme data.

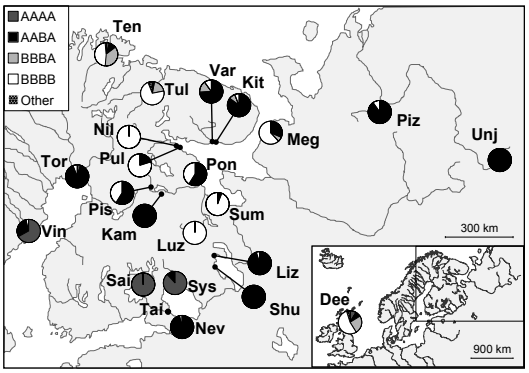

Fig. 1. A map indicating sampling locations and pie diagrams showing the distribution of mtDNA haplotypes among the studied Atlantic salmon populations. MtDNA haplotype data were not available for river Taipale population.

Mitochondrial DNA data

Frequencies of mtDNA haplotypes were obtained from previously published studies (Palva *et al*. 1989, Nielsen *et al*. 1996, Nilsson *et al*. 2001,

Table 2. Details of the PCR reactions and genetic diversity indices of the microsatellite markers employed in this study. Total *A* — the total number of alleles observed across populations. Mean *A* — average number of alleles per population. A_r — allelic richness. H_o — the average observed heterozygosity within a population. H_e — the average expected heterozygosity within a population.

Microsatellite	Group	Label	PCR details				Diversity indices					
			Primer conc. (μM)		Annealing temp. (°C)	Pooled vol. (μl)	Size range (bp)	Total *A*	Mean *A*	A_r	H_o	H_e
			Single	Multiplex								
Ssa85	1	HEX	0.5		55	2.0	115–181	27	8.4	8.9	0.75	0.72
*Ssa171**	1	FAM	0.1	0.2	55	1.5	206–260	25	7.6	7.7	0.72	0.69
Ssa197	1	HEX	0.4		60	1.5	159–271	27	10.2	10.0	0.80	0.79
Ssa202	1	NED	0.3		55	0.5	227–283	14	7.3	8.3	0.74	0.73
*SSOSL85**	1	NED	0.5	0.5	55	1.0	178–226	19	6.7	7.4	0.67	0.65
SSOSL311	1	FAM	0.4		53	1.0	120–186	31	9.7	10.1	0.75	0.74
*SSOSL438**	1	NED	0.2	0.11	55	0.5	117–151	16	4.4	4.9	0.53	0.54
SLEEI84†	2	FAM	0.3	0.4	55	1.0	171–229	28	7.6	7.1	0.60	0.58
SLEEN82†	2	NED	0.4	0.4	55	1.0	209–233	13	4.3	5.0	0.48	0.47
Ssa14†	2	HEX	0.4	0.4	55	1.0	146–152	4	2.2	2.1	0.34	0.33
Ssa412	2	FAM	0.2		60	1.0	282–306	8	3.3	3.4	0.44	0.44
Ssa422	2	HEX	0.6		60	1.0	180–210	12	5.1	5.6	0.65	0.61
SSD30	2	HEX	0.6		50	1.0	217–247	10	2.7	3.2	0.27	0.27
SSF43	2	HEX	0.4		55	1.0	106–126	8	3.1	3.6	0.30	0.32
Average of all loci								16.6	5.7	16.6	5.72	6.2

* Multiplex I.
† Multiplex II.

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Asplund *et al*. 2004) for all but three of the populations included in the study. Haplotypes for the land-locked populations from Kamennoe and Luzhma were generated using the methods described in Asplund *et al*. (2004). MtDNA data were not available for the population from the river Taipale.

Statistical analyses

Genetic diversity and Hardy-Weinberg and genotypic linkage equilibrium

For microsatellites, the observed number of alleles (A), the observed proportion of heterozygotes (H_o), and the expected level of gene diversity (H_e) were estimated using Mircosatellite Toolkit 3.1 (Park 2002). To account for unequal sample sizes, allelic richness (A_r) was calculated with FSTAT 2.9.3.2 (Goudet 2001). Since no individual allozyme data were available for samples from rivers Dee, Tornio, and Vindelälven, the observed number of alleles as well as unbiased gene diversity were calculated from previously published allele frequencies. GENEPOP version 3.4 (Raymond & Rousset 1995) was used to test for significant deviations from Hardy-Weinberg equilibrium (HWE) or genotypic linkage equilibrium. To correct for multiple significance tests, a sequential Bonferroni correction (Rice 1989) was employed.

Genetic differentiation

Genetic divergence among populations was studied using Wright's F_{ST} and its analogue ρ_{ST} (Rousset 1996), which is based on the stepwise mutation model. GENEPOP 3.4 was employed to compute F_{ST} and ρ_{ST} estimates over all populations and for all population pairs based on microsatellite data by using the θ estimator of Weir and Cockerham (1984) and the Φ estimator of Michalakis and Excoffier (1996), respectively. F_{ST} values were not estimated from the allozyme data since individual data were not available for all populations. For the computation of ρ_{ST} estimates, allele sizes in base pairs were transformed into numbers of repeat units. The loci *Ssa171* and *SLEEI84* had alleles that were not multiples of the reported repeat unit length. In *Ssa171* these half alleles were observed in up to 40% of individuals in seven populations and were marked as missing data. In *SLEEI84* the half alleles were more common, occurring in up to 78% of individuals in 12 populations, and thus the locus was excluded from the computation ρ_{ST} estimates.

Isolation by distance

A Mantel's test was employed to examine if there was an association between the geographical and genetic divergence of the populations. Non-anadromous populations were excluded from the analysis since there are no migration possibilities between populations. Interpopulation geographical distances were calculated as shortest water distances between river-mouths and they were plotted against the estimates of $F_{ST}/(1 - F_{ST})$ (Rousset 1997) and D_{CE} distances calculated from the microsatellite data. Allozyme data were excluded from this analysis due to the unavailability of individual data for some populations. The Mantel's test was performed using the procedure of Smouse *et al*. (1986) implemented within the program GenAlEx 5.1 (Peakall & Smouse 2001). Statistical significance of the values was obtained via 999 random permutations.

Relationships between populations

To study the inter-population relationships and to infer potential post-glacial colonization routes of the Atlantic salmon, the Cavalli-Sforza and Edwards' (1967) chord distance (D_{CE}) was calculated. This distance measure was chosen because it has been shown to be one of the most efficient in obtaining correct tree topology using microsatellite data (Takezaki & Nei 1996). A Neighbor-Joining tree was constructed utilizing the programs Seqboot, Gendist, Neighbor, Consense, and Fitch from Phylip version 3.573c or 3.6b software packages (e.g. Felsenstein 1995). Trees were created for allozyme and microsatellite data separately, and also for data for both marker types combined. Phylogram reliability was estimated by 2000 bootstrap replicates over loci.

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Individual assignment tests

The extent of genetic differentiation among populations can also be measured with assignment tests (Cornuet *et al*. 1999, Hansen *et al*. 2001). GeneClass2 program (Piry *et al*. 2004) was used to assign individuals to their most likely population of origin based on their microsatellite multi-locus genotype. Again, allozyme data was excluded from this analysis since individual data were not available for all populations. The direct approach using the Bayesian method (Rannala & Mountain 1997) and the "leave one out" procedure were employed.

Analysis of molecular variance

A hierarchical analysis of molecular variance (AMOVA) was performed using microsatellite data by applying the Arlequin version 2.0 software (Schneider *et al*. 2000). In order to test alternative colonisation hypotheses, populations were grouped in three different combinations:

1. All non-anadromous populations were grouped together with the anadromous populations from the Baltic Sea to test for possible common ancestry. A second group was formed of the anadromous populations from the White and Barents Seas, and the Atlantic Ocean to account for their potential recolonization from a second refugium.
2. As above except that the three non-anadromous populations from rivers draining to the White Sea basin, Kamennoe, Pisto, and Luzhma, were grouped together with the anadromous populations from the White Sea basin.
3. As for grouping 1, except that the non-anadromous populations of Kamennoe, Pisto, and Luzhma were included as a third, separate group.

Allele size permutation test

To test whether stepwise-like mutations have contributed to genetic differentiation between groups of populations, the allele size randomization test (Hardy *et al*. 2003) implemented in the program SPAGeDi 1.1 (Hardy & Vekemans 2002) was utilized. This method can be interpreted as testing whether $F_{ST} = R_{ST,}$ where R_{ST} is an SMM-based measure of genetic differentiation based on microsatellite allele size variance, which is analogous to F_{ST}, and unbiased with respect to differences in sample size between populations and differences in variance between loci (Slatkin 1995). When the contribution of mutations to genetic differentiation is negligible as compared with genetic drift and migration, estimates of differentiation using F_{ST} and R_{ST} should be similar. On the other hand, if stepwise-like mutations have contributed significantly to divergence, $R_{ST} > F_{ST}$.

Taking into account the generation time of salmon and the approximate microsatellite mutation rate, stepwise-like mutations should not have contributed significantly to the divergence of populations colonised from the same glacial refugium, but should have contributed to the divergence of populations colonised from different refugia (Estoup & Angers 1998). Hence, allele size permutation provides a method for testing whether particular regions have been colonised from more than one refugium. For this purpose the populations were divided into three groups: (1) the Baltic Sea basin, (2) Kamennoe, Pisto, and Luzhma, and (3) White and Barents Seas, and the Atlantic Ocean. Due to the geographical distinctiveness of these regions, and also genetics data (*see* below), migration between these regions can be considered to be negligible. Again, the locus *SLEEI84* was excluded because of its high frequency of half alleles. It should be noted however that as a consequence of their stepwise-like mutation pattern, size homoplasy can occur in microsatellite loci, i.e. alleles of the same size are not necessarily derived from the same ancestral allele (for review *see* Estoup *et al*. 2002), which could lead to underestimation of population divergence. If the amount of size homoplasy was considerable, the allele size permutation test would be conservative, as the level of genetic differentiation as measured by R_{ST} would underestimate the level of population differentiation, and hence also underestimate the contribution of mutations to population divergence.

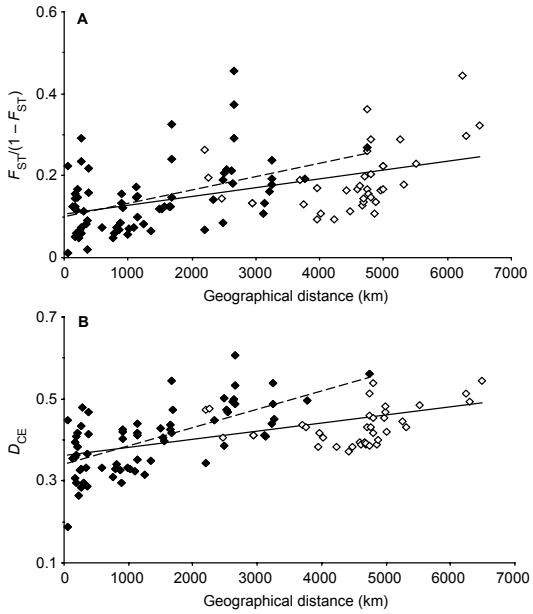

Fig. 2. The relationship between the geographical and genetic distances of anadromous Atlantic salmon populations. Genetic distances are given as a) $F_{ST}/(1- F_{ST})$, and b) D_{CE}. Black diamonds represent interpopulations distances between Baltic populations, grey diamonds represent interpopulation distances between Atlantic/Barents/White populations and white diamonds represent interpopulation distances between the two groups. The solid black line represents the regression slope of all interpopulation comparisons (F_{ST}: Mantel's r_{XY} = 0.477, P = 0.001; D_{CE}: Mantel's r_{XY} = 0.513, P = 0.001) and the grey hatched line that of Atlantic/Barents/White population comparisons. (F_{ST}: r_{XY} = 0.452, P = 0.048; D_{CE}: Mantel's r_{XY} = 0.651, P = 0.004).

Results

Microsatellite and allozyme diversity of Atlantic salmon from northern Europe

The average number of alleles per microsatellite locus within a population varied between 2.5 (Saimaa) and 10.6 (Dee; Table 1 and Appendix 1) with average observed heterozygosity ranging from 0.29 (Saimaa) to 0.72 (Dee). As expected, the variation observed at allozyme loci was lower than at microsatellite loci (Table 1 and Appendix 2), with the number of alleles per allozyme locus varying from 1.3 (Kamennoe) to 2.3 (Tornio) and expected heterozygosity from 0.07 (Kamennoe) to 0.33 (Varzuga).

Hardy-Weinberg and linkage equilibrium of microsatellite and allozyme loci

Considering data for each population, the null hypothesis of Hardy-Weinberg equilibrium (HWE) could not be rejected for any microsatellite or allozyme locus after a Bonferroni-type correction for multiple statistical tests. At the population level, following the correction for multiple tests, one deviation from Hardy-Weinberg proportions (Nilma) remained significant (Appendix 1). This was due to significant heterozygote excess at three of the 14 microsatellite loci (*SSOSL311, SSOSL438, Ssa14*). This may be due to hybridisation between Atlantic salmon and trout in this river system (J. Lumme unpubl. data).

After correcting for multiple statistical tests, five pairs of microsatellite loci were in significant linkage disequilibrium (LDE). However, as there are no indications that these loci are genetically linked (Gharbi 2001, Gilbey *et al*. 2004), it is unlikely that these cases affected our analyses. Regarding the allozyme data for which individual level genotypes were available, only one pair of loci (*sAAT-4** & *IDDH-2**) was not in linkage equilibrium.

Genetic differentiation and relationships between the North European salmon populations

Single locus estimates of F_{ST} and ρ_{ST} ranged from 0.127 to 0.309 and from 0.081 to 0.314, respectively. The lowest estimate of pairwise F_{ST} calculated over all loci was 0.011 (Kitsa–Varzuga) and the highest 0.500 (Saimaa–Kamennoe) with the mean value being 0.217 (Appendix 3). The pairwise difference measured with ρ_{ST} ranged between 0.007 (Kitsa–Varzuga) and 0.571 (Saimaa–Unja) with the mean value of 0.218 (Appendix 3). A significant association between geographical and genetic distance (both F_{ST} and D_{CE}: Appendix 4) was observed for all anadromous populations and also when the three Baltic populations were excluded from the analysis (Fig. 2).

Based on allozyme data alone, phylogeographic resolution was relatively limited with the

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

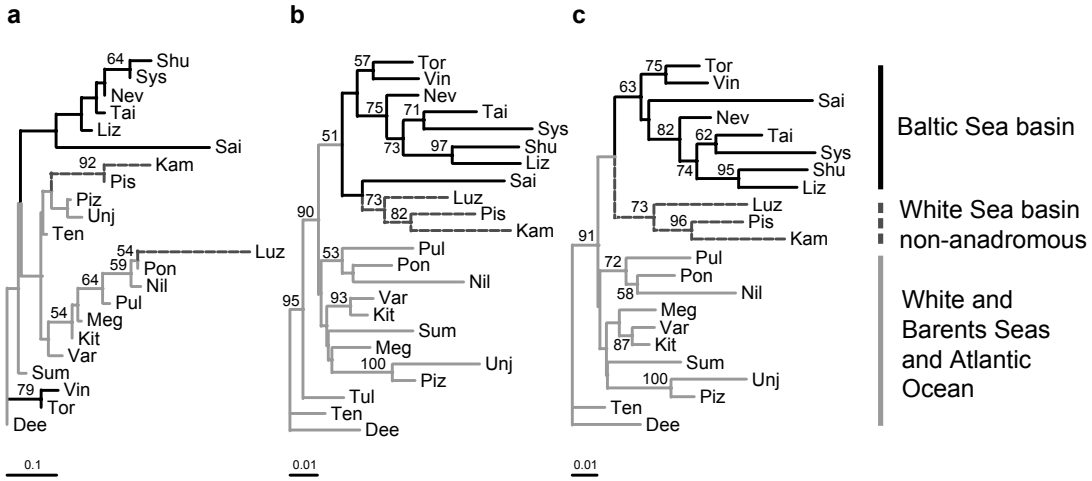

Fig. 3. Neighbour-joining phylograms based on (**a**) allozyme data, (**b**) microsatellite data, and (**c**) the combination of both marker types based on D_{CE} distances. The numbers indicate percent bootstrap support for each node over 2000 replications. Only values over 50% are shown.

NJ phylogram indicating only a small number of nodes supported by bootstrap values > 50% (Fig. 3a). The geographic population structure based on microsatellite data alone was clearer and generally corresponded with the geographical sampling regions (Fig. 3b). The analysis of the combined microsatellite–allozyme data set identified the same three groups as the microsatellite data alone (Fig. 3c). However, the bootstrap support of some key nodes rose and the within-cluster population relationships were clearer. With a bootstrap support of 63%, Lake Saimaa and rivers Tornio and Vindelälven clustered together with Neva and the Lake Onega and Ladoga populations, all from the Baltic Sea basin. Again, the three Karelian non-anadromous populations from the White Sea basin formed a well-supported group of their own. The bootstrap support for the affinity between the Baltic and White Sea basin non-anadromous population groups (36%) was lower than for microsatellite data alone (51%) indicating that at least some microsatellite and allozyme loci suggest differing affinities. The remaining populations included all those from the White and Barents Seas, and the Atlantic Ocean. Support for this cluster as a separate group was not high, and highly supported nodes within the cluster tended to be for populations situated geographically close to each other (Fig. 3c).

Overall, 91.2% of the individuals were assigned to the correct population of origin and 99.6% to the group of origin: (1) the Baltic Sea basin, (2) Kamennoe, Pisto, and Luzhma, and (3) White and Barents Seas, and the Atlantic Ocean (Table 3). At the population level, 100% assignment success was achieved for seven out of eight non-anadromous populations compared to only one of 15 anadromous populations. Assignment efficiency was above 80% for all populations except Kitsa (68.9%) and Varzuga (57.4%), two rivers that share the same estuary and had the lowest pair-wise F_{ST} observed in the study. Accordingly, the majority of incorrect assignments in these two rivers were reciprocal.

Partitioning of microsatellite variance

The hierarchical analysis of molecular variance (AMOVA) revealed that the proportion of variance residing between groups was highest (8.4%) when the populations were divided into three groups: (1) the Baltic Sea basin, (2) Kamennoe, Pisto, and Luzhma, and (3) White and Barents Seas, and the Atlantic Ocean (Table 4). With other population groupings, the proportion of between-group variance was 34% to 45% less (Table 4). Most of the variation, ca. 77%, resided at the within population level.

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Table 3. Number of individuals assigned to different populations based on their multi-locus microsatellite genotype. The populations are organized according to their region of origin. Assignments within groups are shaded in grey **[?]**.

	Atlantic Ocean			Barents Sea		White Sea										Baltic Sea									All
						Anadr.							Non-anadr.			Anadr.			Non-anadr.						
	Dee	Ten	Tul	Piz	Unj	Meg	Nil	Pon	Pul	Sum	Kit	Var	Luz	Kam	Pis	Nev	Tor	Vin	Liz	Sai	Shu	Sys	Tai		
Dee	47	1																							48
Ten		37	3			1		1																	42
Tul		6	32				1	2									1								42
Piz				21	1							1													23
Unj					10																				10
Meg		2	2			43					2														49
Nil							33		1																34
Pon			1					36		1		3													41
Pul	1		1						39		1														42
Sum								1		37	1														39
Kit			1			1	1	3			31	16													53
Var			1			2		1			10	27				1									42
Luz													40		1										42
Kam														41											40
Pis															52										52
Nev																42		1							43
Tor			1			1											39	1							42
Vin																	1	45							46
Liz																	1		26						27
Sai																				44					44
Shu																					20				20
Sys																						42			42
Tai																							38		38
Total	48	46	42	21	11	48	35	44	40	38	45	47	40	41	53	43	42	47	26	44	20	42	38		901
Correct (%)	97.9	80.4	76.2	100	90.9	89.6	94.3	81.8	97.5	97.4	68. 9	57.4	100	100	98.1	97.7	92.9	95.7	100	100	100	100	100		91.2

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Comparison of F_{ST} and R_{ST} estimates — has mutation contributed to genetic differentiation?

For the allele size permutation test, the populations were divided into three groups based on the result of the AMOVA: (1) the Baltic Sea basin, (2) Kamennoe, Pisto, and Luzhma, and (3) White and Barents Seas, and the Atlantic Ocean. The global multilocus estimates of F_{ST} and R_{ST} between groups were 0.103 and 0.118, respectively (Fig. 4). The observed multilocus R_{ST} lay above the upper limit of the 95% confidence interval of the null distribution of the permuted pR_{ST} and was statistically significant ($P = 0.016$). Considering single loci, the global R_{ST} values for *Ssa202* and *SSOSL85* were significantly larger than the pR_{ST} ($P = 0.042$ and $P = 0.038$, respectively). This indicates that stepwise-like mutations have contributed to genetic divergence and therefore, postglacial colonisation of northern Europe by more than one glacial refugium is statistically supported. The result remained statistically significant ($P = 0.017$) when the three non-anadromous populations from the White Sea basin (Kam, Pis, Luz) were excluded from the analysis. However, for other group pairings, there was no indication that stepwise-like mutations have contributed to genetic divergence (Fig. 4). The same was true at the within group level.

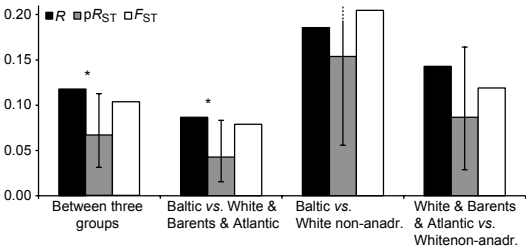

Fig. 4. Global R_{ST}, pR_{ST}, and F_{ST} estimates of the three groups. Allele size permutation test results for testing the contribution of mutation to differentiation between the three groups of North European Atlantic salmon. *See* text for details of the groupings. The 95% confidence intervals are given for pR_{ST}. cases where global R_{ST}, was significantly ($0.05 > P > 0.01$) larger than the permuted null distribution pR_{ST}, are indicated with an asterix.

Discussion

Genetic structure of the north European Atlantic salmon

The level of genetic divergence among anadromous populations, as measured using F_{ST} (0.123), was similar to that observed in other studies investigating the genetic relationships of European Atlantic salmon populations using nuclear markers (0.092 to 0.136, Bourke *et al.* 1997, King *et al.* 2001, Koljonen *et al.* 2002) and also similar to divergence estimates in other anadromous fish species (seven species average $G_{ST} = 0.108$, Ward *et al.* 1994). Global F_{ST} estimates for different regions were relatively similar (White Sea: 0.103, Barents Sea and Atlantic Ocean: 0.095, and Baltic Sea: 0.089). The level of genetic divergence between non-anadromous populations was considerably higher (0.336) and similar to the divergence levels observed between other freshwater salmonid populations (0.360 to 0.482, e.g. García-Marin *et al.* 1999, Primmer *et al.* 1999) and freshwater species in general (49 species average $G_{ST} = 0.222$, Ward *et al.* 1994). This difference, and the signal

Table 4. Hierarchical analysis of molecular variance (AMOVA) results for three alternative groupings of the studied populations. Percentages of total variation explained (%) and fixation indices are given for three hierarchical levels.

Partitioning of the populations			Among groups			Within groups			Within populations		
Group 1*	Group 2†	Group 3	%	P‡	F_{CT}	%	P‡	F_{SC}	%	P‡	F_{ST}
Baltic + Luz, Kam, Pis	Whi, Bar, Atl	–	4.7	***	0.05	17.8	***	0.19	77.6	***	0.22
Baltic	Whi, Bar, Atl + Luz, Kam, Pis	–	5.6	***	0.06	17.5	***	0.19	77.0	***	0.23
Baltic	Whi, Bar, Atl	Luz, Kam, Pis	8.4	***	0.08	14.8	***	0.16	76.9	***	0.23

* Baltic, Baltic Sea basin; Luz, Luzhma; Kam, Kamennoe; Pis, Pisto: † Whi, White Sea; Bar, Barents Sea; Atl, Atlantic Ocean; ‡ *** $P < 0.001$.

of isolation by distance observed for anadromous populations (Fig. 2), highlights the effect of straying i.e. migration on the population genetic structure of Atlantic salmon. The moderate to high level of population differentiation is also evidenced in the individual assignment test results (Table 3). Overall 91.2% of the individuals were assigned to the correct population of origin and 99.3% to 100% to the correct group of origin. As expected, due to their higher level of population differentiation, all but one non-anadromous population had all their individuals correctly assigned. In comparison, the assignment efficiency (57.4% to 100%) among the populations from the White/Barents/Atlantic populations was lower.

Even though individuals could be correctly assigned to their group of origin, analysis of molecular variance indicated that the proportion of genetic variation explained at the between-group level was relatively low (8.4%; Table 4). Nevertheless, this proportion is similar to that observed in an earlier study of Atlantic salmon (6.1% between European countries and 3.2% between North American provinces; King *et al*. 2001). In contrast, a higher proportion of molecular variance, 21.9%, was found to reside between the two continents (King *et al*. 2001).

Relationships between populations

Interpopulation relationships based on allozyme data alone were inconclusive, with few nodes supported by bootstrap values above 50% (Fig. 3a). On the other hand, microsatellite data divided the populations into groups corresponding relatively well to their geographical origin but the relationship of some populations to the groups still remained unclear (Fig. 3b). Combining allozyme and microsatellite data, the relationships became clearer (Fig. 3c): all populations from the Baltic Sea basin grouped together with a bootstrap support value of 63%. Within this cluster, populations from the Gulf of Finland and the Gulf of Bothnia formed separate, well supported, groups. Interestingly, the anadromous River Neva population shared a higher affinity with geographically closer Lake Ladoga and Lake Onega populations than with anadromous populations from the northern Baltic Sea basin, Bothnian Bay (Fig. 3c). The three non-anadromous Karelian populations from the White Sea basin formed a second distinct group (Fig. 3c). The clustering of populations from the White and Barents Seas and the Atlantic Ocean was more highly supported by microsatellite data alone than by the combined data-set (bootstrap support of 51% *vs*. 36%, Fig. 3) and thus the relationship between these populations remains somewhat unclear. Overall however, the level of phylogram bootstrap support is higher than that observed in earlier studies (e.g. Koljonen *et al*. 1999, King *et al*. 2001).

Post-glacial origin of north European Atlantic salmon

Evidence of a single glacial refugium for the Baltic Sea

The colonisation of the Baltic Sea and lakes Ladoga and Onega from a single refugium is supported by the grouping of all Baltic Basin populations together with a moderately high bootstrap support (Fig. 3c). The most likely source of post-glacial colonisation is the Baltic Ice Lake, the predecessor of the present day Baltic Sea and lake Ladoga, which was situated at the southern and southeastern edge of the Scandinavian Ice Sheet ~ 12 600 to 10 300 years ago (e.g. Björck 1995). Lake Onega was not a part of the Baltic Ice Lake but it has been suggested to have been connected to it (Saarnisto *et al*. 1995), thus allowing colonisation of Lake Onega also from this refugium. This post-glacial colonisation scenario is in line with that proposed by Nilsson *et al*. (2001), based on mtDNA, but is not concordant with colonisation scenarios where a significant contribution of salmon from a western (North Sea) refugium has been proposed for recolonising part (Koljonen *et al*. 1999) or all (Verspoor *et al*. 1999) of the Baltic region. Given the highly supported (91% bootstrap support, Fig. 3c) separation of all Baltic Sea basin populations from the Scottish River Dee population, any significant contribution of North Sea stocks to recolonisation of the Baltic seems highly unlikely.

Date and signature
I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Several additional glacial lineages in the White, Barents, and Atlantic basins

Both the population phylogram (Fig. 3c) and the allele size permutation test (Fig. 4) indicate that the populations from the White, Barents, and Atlantic basins most likely originate from different glacial refugia than the Baltic populations. A likely location of one refugium is the eastern Barents Sea. During the Late Weichselian, from ca. 25 000 to 10 000 years ago, the western Barents Sea and White Sea were still covered by ice but survival in the eastern Barents Sea should have been feasible (Mangerud *et al*. 2001). As the glacier receded it would have been possible for salmon to spread to the rivers of the Kola Peninsula and to the White Sea (Kazakov & Titov 1991, Asplund *et al*. 2004). The mitochondrial haplotype (AABA) which is observed in high frequencies in both the Baltic and White Sea basins (Fig. 1, Appendix 5; Nilsson *et al*. 2001, Asplund *et al*. 2004) is most likely due to its presence in inhabitants of the Komi Ice Lake about 90 000 years ago (Mangerud *et al*. 2001), which has been connected to both the Barents and Baltic Ice Lakes at differing times (Maslenikova & Mangerud 2001).

A second potential source of the White and Barents Sea salmon could be the Atlantic Ocean as has been suggested earlier (Verspoor *et al*. 1999, Asplund *et al*. 2004). While neither the nuclear locus phylogram, nor the allele size permutation test, lend strong support to this theory (Fig. 3c), the consistent rare occurrence of a western Atlantic allozyme allele **80* at the locus *ESTD-2** (Bourke *et al*. 1997) in populations from the White and Barents Sea basins, including two non-anadromous populations (Appendix 2), supports early immigration from the western Atlantic Ocean. Similarly, the mtDNA haplotype BBBB commonly occurs in both the Atlantic Ocean and the White Sea populations (Fig. 1, Appendix 5; Nilsson *et al*. 2001, Asplund *et al*. 2004). Under this scenario, salmon would have dispersed via a stepping stone model from the North Sea or even as far as from the Iberian Peninsula along the coast of Norway to the Barents Sea and eventually to the White Sea (Verspoor *et al*. 1999, Asplund *et al*. 2004), to supplement the apparent contribution of an eastern refugium.

Non-anadromous populations possibly colonised from at least two different refugia

Both the nuclear locus phylogram (Fig. 3) and the AMOVA analysis (Table 4) indicate that the eight non-anadromous populations can be separated into two distinct groups, one including five populations from the Baltic Sea basin (Lakes Saimaa, Ladoga, and Onega) and a second including the remaining three Karelian populations from the White Sea basin. While the Baltic Sea basin non-anadromous populations formed a group with the anadromous populations (*see* above), the three non-anadromous populations from the White Sea basin formed a distinct cluster (Fig. 3). The distribution of mitochondrial haplotypes (Asplund *et al*. 2004, and the results presented here), as well as the nuclear AMOVA results indicate that Kamennoe, Pisto, and Luzhma are likely derived from the same refugium as the rest of the populations in the White Sea basin (Table 4), but perhaps isolated before significant Atlantic immigration. Thus the origin of the White Sea non-anadromous populations could also be the eastern Barents Sea refugium. The occurrence of an allozyme allele (*ESTD-2*80*) commonly observed in North American populations (Bourke *et al*. 1997) indicates contribution of an Atlantic refugium.

Indications of male-driven gene flow

While studies with maternally inherited mitochondrial DNA markers have identified several clear borders defining the geographical distribution of mtDNA haplotypes (Asplund *et al*. 2004), nuclear markers failed to detect the same clear structure (Fig. 3). More specifically, the relationships between the White and Barents Sea populations based on nuclear markers are unclear (Fig. 3), whilst with mtDNA markers they can be divided into three distinct groups according to abrupt changes in mtDNA haplotype frequencies (*see* fig. 2 in Asplund *et al*. 2004). This could indicate that the homing of female Atlantic salmon to their natal river is stronger than that of males, and hence the population genetic structuring of bi-parentally inherited markers is weaker than maternally inherited markers due to male-

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)
Date and signature

biased gene flow. Strong homing of females makes sense in biological terms. To maximize the size and number of offspring, hatching success and larval survival, the female must lay her eggs at an optimal site (Resetarits 1996). As an egg is essentially a female cell, laying the eggs at a site where the female was incubated and hatched itself — a place that is evidently suitable — maximises her reproductive success and consequently parental fitness.

Acknowledgements

This study was funded by the Finnish Academy and the Finnish Ministry of Agriculture and Forestry. We wish to thank Petri Heinimaa, Jan Nilsson, Jorma Piironen, Igor Studenov and John Taggart for kindly providing population samples, and two anonymous reviewers for their comments on a previous version of the manuscript.

References

Aljanabi, S. M. & Martinez, I. 1997: Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. — *Nucleic Acids Res*. 25: 4692–4693.

Andersen, B. G. & Borns, H. W. 1994: *The ice age world: an introduction to Quaternary history and research with emphasis on North America and Northern Europe during the last 2.5 million years*. — Scandinavian University Press, Oslo.

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.

Berg, O. K. 1985: The formation of non-anadromous populations of Atlantic salmon, *Salmon salar* L., in Europe. — *J. Fish Biol*. 27: 805–815.

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.

Björck, S. 1995: A review of the history of the Baltic Sea, 130-8.0 ka BP. — *Quatern. Int*. 27: 19–40.

Bourke, E. A., Coughlan, J., Jansson, H., Galvin, P. & Cross, T. F. 1997: Allozyme variation in populations of Atlantic salmon located throughout Europe: diversity that could be compromised by introductions of reared fish. — *ICES J. Mar. Sci*. 54: 974–985.

Brownstein, M. J., Carpten, J. D. & Smith, J. R. 1996: Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. — *BioTechniques* 20: 1004–1010.Cairney, M., Taggart, B. J. & Hoyheim, B. 2000: Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. — *Mol. Ecol*. 9: 2175–2178.

Cavalli-Sforza, L. L. & Edwards, A. W. F. 1967: Phylogenetic analysis: models and estimation procedures. — *Am. J. Hum. Genet*. 19: 233–257.

Cornuet, J., Piry, S., Luikart, G., Estoup, A. & Solignac, M. 1999: New methods employing multilocus genotypes to select or exclude populations as origins of individuals. — *Genetics* 153: 1989–2000.

Estoup, A. & Angers, B. 1998: Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. — In: Carvalho, G. R. (ed.), *Advances in molecular ecology,* 1st ed.: 55–85. IOS Press, Amsterdam.

Estoup, A., Jarne, P. & Cornuet J.-M. 2002: Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. — *Mol. Ecol*. 11: 1591–1604.

Felsenstein, J. 1995: *Phylip (Phylogeny Inference Package) version 3.573c*. — Department of Genetics, University of Washington, Seattle.

García-Marin, J.-L., Utter, F. M. & Pla, C. 1999: Postglacial colonization of brown trout in Europe based on distribution of allozyme variants. — *Heredity* 82: 46–56.

Gharbi, K. 2001: Construction d'une carte génétique partielle du génome tétraploïde de la truite commune (*Salmo trutta*). Cartographie comparée des régions paralogues et alignements avec les cartes du saumon atlantique (*Salmo salar*) et de la truite arc-en-ciel (*Oncorhynchus mykiss*). — Ph.D. thesis. Institut National Agronomique Paris-Grignon. [**In French with English abstract**].

Gilbey, J., Verspoor, E., McLay, A. & Houlihan, D. 2004: A microsatellite linkage map for Atlantic salmon (*Salmo salar*). — *Anim. Genet*. 35: 98–105.

Goudet, J. 2001: *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)*. — Available on the web at http://www.unil.ch/izea/softwares/fsat.html.

Hansen, M. M., Kenchington, E. & Nielsen, E. E. 2001: Assigning individual fish to populations using microsatellite DNA markers. — *Fish. Fish*. 2: 93–112.

Hardy, O. J. & Vekemans, X. 2002: SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. — *Mol. Ecol. Notes* 2: 618–620.

Hardy, O. J., Charbonnel, N., Freville, H. & Heuertz, M. 2003: Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. — *Genetics* 163: 1467–1482.

Hewitt, G. M. 1996: Some genetic consequences of ice ages, and their role in divergence and speciation. — *Biol. J. Linn. Soc*. 58: 247–276.

Hewitt, G. M. 1999: Post-glacial re-colonization of European biota. — *Biol. J. Linn. Soc*. 68: 87–112.

Jordan, W. C., Youngson, A. F., Hay, D. W. & Ferguson, A. 1992: Genetic protein variation in natural populations of Atlantic salmon (*Salmo salar*) in Scotland: Temporal

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

and spatial variation. — *Can. J. Fish. Aquat. Sci.* 49: 1863–1872.

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. [Казаков, Р. В. & Титов, С. Ф.] 1998: [Population genetic structure of Atlantic salmon *Salmo salar*]. — In: Kazakov, R. V. [Казаков, Р. В.] (ed.), [*Atlantic salmon*], 1st ed.: 43–72. Nauka, Sankt-Petersburg, Russia. [In Russian].

King, T. L., Kalinowski, S. T., Schill, W. B., Spidle, A. P. & Lubinski, B. A. 2001: Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. — *Mol. Ecol.* 10: 807–821.

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.

Koljonen, M.-L., Tähtinen, J., Säisä, M. & Koskiniemi, J. 2002: Maintenance of genetic diversity of Atlantic salmon (*Salmo salar*) by captive breeding programmes and the geographic distribution of microsatellite variation. — *Aquaculture* 212: 69–92.

Kontula, T. & Väinölä, R. 2001: Postglacial colonization of Northern Europe by distinct phylogeographic lineages of the bullhead, *Cottus gobio*. — *Mol. Ecol.* 10: 1983–2002.

Koskinen, M. T., Nilsson, J., Veselov, A. J., Potutkin, A. G., Ranta, E. & Primmer, C. R. 2002: Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus*, Salmonidae. — *Heredity* 88: 391–401.

Koskinen, M. T., Ranta, E., Piironen, J., Veselov, A., Titov, S., Haugen, T. O., Nilsson, J., Carlstein, M. & Primmer, C. R. 2000: Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. — *Mol. Ecol.* 9: 1609–1624.

Mangerud, J., Astakhov, V. I., Murray, A. & Svendsen, J. I. 2001: The chronology of a large ice-dammed lake and the Barents-Kara Ice Sheet advances, Northern Russia. — *Global Planet. Change* 31: 321–336.

Martinez, J. L., Moran, P. & Garcia-Vazquez, E. 1999: Dinucleotide repeat polymorphism at the *SS4*, *SS6* and *SS11* loci in Atlantic salmon (*Salmo salar*). — *Animal genetics* 30: 464–465.

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

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

McConnell, S., Hamilton, L., Morris, D., Cook, D., Paquet, D., Bentzen, P. & Wright, J. **1995a**: Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. — *Aquaculture* 137: 19–30.

Michalakis, Y. & Excoffier, L. 1996: A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. — *Genetics* 142: 1061–1064.

Nesbø, C. L., Fossheim, T., Vøllestad, L. A. & Jakobsen, K. S. 1999: Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. — *Mol. Ecol.* 8: 1387–1404.

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

Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., Kohlmann, 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. — *Mol. Ecol.* 10: 89–102.

O'Reilly, P. T., Hamilton, L. C., McConnell, S. K. & Wright, J. M. 1996: Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. — *Can. J. Fish. Aquat. Sci.* 53: 2292–2298.

Palva, T. K., Lehväsaho, H. & Palva E. T. 1989: Identification of anadromous and non-anadromous salmon stocks in Finland by mitochondrial DNA analysis. — *Aquaculture* 81: 237–244.

Park, S. D. E. 2002: *Trypanotolerance in West African cattle and the population genetic effects of selection*. — Ph.D. thesis, University of Dublin.

Parrish, D. L., Behnke, R. J., Gephard, S. R., McCormick, S. D. & Reeves, G. H. 1998: Why aren't there more Atlantic salmon (*Salmo salar*)? — *Can. J. Fish. Aquat. Sci.* 55 (suppl. 1): 281–287.

Peakall, R. & Smouse, P. E. 2001. *GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research*. — Australian National University, Canberra, Australia.

Piry, S., Alapetite, A., Cornuet, J., Paetkau, D., Baudouin, L. & Estoup, A. 2004: GeneClass2: a software for genetic assignment and first generation migrants detection. — *J. Hered.* 95: 536–539.

Primmer, C. R., Aho, T., Piironen, J. & Ranta, E. 1999: Examination of microsatellite variability in Nordic land-locked arctic charr, *Salvelinus alpinus* populations: implications for conservation. — *Hereditas* 130: 277–289.

Rannala, B. & Mountain, J. L. 1997: Detecting immigration by using multilocus genotypes. — *Proc. Natl. Acad. Sci. USA* 94: 9197–9201.

Raymond, M. & Rousset, F. 1995: GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. — *J. Hered.* 86: 248–249.

Resetarits, W. J. Jr. 1996: Oviposition site choice and life history evolution. — *Am. Zool.* 36: 205–215.

Rice, W. R. 1989: Analyzing tables of statistical tests. — *Evolution* 43: 223–225.

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Rousset, F. 1996: Equilibrium values of measures of population subdivision for stepwise mutation processes. — *Genetics* 142: 1357–1362.

Rousset, F. 1997: Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. — *Genetics* 145: 1219–1228.

Saarnisto, M., Gronlund, T. & Ekman, I. 1995: Lateglacial of Lake Onega — contribution to the history of the eastern Baltic Basin. — *Quatern. Int.* 27: 111–120.

Sánchez, J. A., Clabby, C., Ramos, D., Blanco, G., Flavin, F., Vazquez, E. & Powell, R. 1996: Protein and microsatellite single locus variability in *Salmo salar* L. (Atlantic salmon). — *Heredity* 77: 423–432.

Schneider, S., Roessli, D. & Excoffier, L. 2000: *Arlequin ver. 2.000: A software for population genetics data analysis*. — Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Slatkin, M. 1995: A measure of population subdivision based on microsatellite allele frequencies. — *Genetics* 139: 457–462.

Slettan, A., Olsaker, I. & Lie, O. 1995: Atlantic salmon, *Salmo salar*, microsatellites at the *SSOSL25*, *SSOSL85*, *SSOSL311*, *SSOSL417* loci. — *Animal genetics* 26: 281–282.

Slettan, A., Olsaker, I. & Lie, O. 1996: Polymorphic Atlantic salmon, *Salmo salar* L., microsatellites at the *SSOSL438*, *SSOSL439* and *SSOSL444* loci. — *Animal genetics* 27: 57–58.

Smouse, P. E., Long, J. C. & Sokal, R. R. 1986: Multiple regression and correlation extensions of the Mantel test of matrix correspondence. — *Syst. Zool.* 35: 627–632.

Sokal, R. R. & Rolof, F. J. 1995: *Biometry: the principles and practice of statistics in biological research.* — W.H. Freeman and company, New York.

Ståhl, G. 1987: Genetic population structure of Atlantic salmon. — In: Ryman, N. & Utter, F. (eds.), *Population genetics and fishery management*, 1st ed.: 121–140. University of Washington Press, Seattle.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G. & Cosson, J. F. 1998: Comparative phylogeography and postglacial colonization routes in Europe. — *Mol. Ecol.* 7: 453–464.

Takezaki, N. & Nei, M. 1996: Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. — *Genetics* 144: 389–399.

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. — *Biol. J. Linn. Soc.* 68: 129–146.

Ward, R. D., Woodwark, M. & Skibinski, D. O. F. 1994: A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. — *J. Fish Biol.* 44: 213–232.

Weir, B. S. & Cockerham, C. C. 1984: Estimating F-statistics for the analysis of population structure. — *Evolution* 38: 1358–1370.

Date and signature

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Appendix 1. Microsatellite diversity indices of the 23 studied Atlantic salmon populations and the results of the Hardy-Weinberg equilibrium test. *N* — number of individuals. *A* — Average number of alleles in a population. H_o — observed heterozygosity. H_e — expected heterozygosity. HWE — the *P* value of the Hardy-Weinberg equilibrium test after. The *P* values over all populations and loci have been corrected for multiple significant tests.

	Dee	Kam	Kit	Liz	Luz	Meg	Nev	Nil	Pis	Piz	Pon	Pul	Sai	Shu	Sum	Sys	Tai	Ten	Tor	Tul	Unj	Var	Vin	All
Ssa85																								
N	48	41	45	25	40	46	42	34	53	21	43	40	41	20	38	42	38	41	42	41	11	47	47	
A	14	3	12	6	5	7	12	7	3	7	9	9	1	6	11	12	14	12	8	8	6	14	7	8.4
H_o	0.81	0.68	0.91	0.92	0.38	0.76	0.91	0.85	0.45	0.76	0.84	0.68	0.00	0.80	0.95	0.79	1.00	0.76	0.79	0.68	0.91	0.87	0.83	0.75
H_e	0.91	0.57	0.82	0.76	0.37	0.80	0.88	0.73	0.50	0.79	0.82	0.78	0.00	0.70	0.81	0.78	0.87	0.77	0.77	0.64	0.82	0.84	0.81	0.72
HWE	0.018	ns	ns	ns	ns	ns	ns	ns	0.041	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Ssa171																								
N	47	40	38	26	40	42	43	34	51	21	37	39	34	20	34	42	38	44	42	33	11	41	47	
A	12	5	8	3	8	8	12	5	5	6	6	6	4	4	7	6	10	13	11	16	6	9	5	7.6
H_o	0.87	0.70	0.79	0.39	0.80	0.71	0.56	0.91	0.80	0.91	0.65	0.77	0.77	0.55	0.68	0.88	0.84	0.73	0.76	0.85	0.91	0.56	0.19	0.72
H_e	0.85	0.71	0.70	0.34	0.82	0.75	0.61	0.73	0.70	0.72	0.60	0.69	0.71	0.51	0.78	0.75	0.84	0.85	0.70	0.89	0.84	0.68	0.18	0.69
HWE	ns	0.015	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.034	ns	ns	ns	ns	ns	0.050	ns	ns	ns	ns	ns	ns
Ssa197																								
N	47	40	45	25	40	42	41	30	45	20	41	40	40	20	36	36	36	43	42	41	10	46	47	
A	23	4	14	4	7	15	10	7	7	7	12	9	4	10	7	6	6	20	14	16	7	13	12	10.2
H_o	0.92	0.88	0.96	0.76	0.83	0.93	0.88	0.73	0.60	0.75	0.73	0.80	0.38	0.90	0.69	0.67	0.67	0.84	0.91	0.88	1.00	0.85	0.77	0.80
H_e	0.92	0.71	0.88	0.64	0.76	0.89	0.83	0.77	0.69	0.76	0.84	0.78	0.40	0.84	0.77	0.76	0.65	0.91	0.89	0.90	0.87	0.87	0.85	0.79
HWE	ns	ns	ns	ns	ns	ns	ns	ns	0.003	ns	ns	ns	ns	ns	ns	0.003	ns	ns	ns	ns	ns	ns	ns	ns
Ssa202																								
N	47	34	45	26	39	45	40	34	53	21	37	40	44	20	35	42	37	43	42	39	11	47	46	
A	11	5	12	7	6	11	7	5	4	7	10	9	4	5	5	4	6	11	8	10	4	11	6	7.3
H_o	0.96	0.47	0.93	0.65	0.69	0.89	0.88	0.59	0.64	0.76	0.87	0.70	0.75	0.50	0.83	0.50	0.76	0.79	0.62	0.82	0.91	0.87	0.54	0.74
H_e	0.89	0.48	0.90	0.73	0.75	0.85	0.76	0.62	0.55	0.80	0.85	0.75	0.72	0.63	0.75	0.53	0.74	0.88	0.67	0.87	0.62	0.89	0.57	0.73
HWE	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.03	ns	ns	ns	ns	ns	0.047	ns	ns	ns	ns
SSOSL85																								
N	46	39	45	26	39	47	43	35	53	21	38	39	39	20	36	42	38	45	42	42	11	47	47	
A	12	4	11	3	4	8	8	4	6	6	10	5	2	5	4	4	6	12	10	10	4	11	5	6.7
H_o	0.89	0.74	0.67	0.39	0.31	0.60	0.81	0.63	0.70	0.76	0.79	0.82	0.39	0.55	0.58	0.60	0.63	0.87	0.71	0.74	0.82	0.66	0.68	0.67
H_e	0.86	0.68	0.71	0.43	0.28	0.56	0.77	0.66	0.67	0.72	0.78	0.71	0.40	0.48	0.55	0.58	0.60	0.85	0.80	0.83	0.62	0.71	0.61	0.65
HWE	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.02	ns	ns	ns	ns	ns	ns	ns	0.010	ns	ns	ns	ns
SSOSL311																								
N	47	40	45	26	40	48	42	35	52	20	43	39	43	20	38	41	38	44	40	42	11	47	44	
A	18	3	17	5	7	13	12	7	4	9	9	8	3	4	6	9	14	15	13	15	7	16	8	9.7
H_o	0.87	0.68	0.84	0.73	0.68	0.85	0.86	0.89	0.60	0.70	0.81	0.64	0.12	0.40	0.87	0.78	0.97	0.82	0.90	0.83	0.82	0.77	0.84	0.75
H_e	0.91	0.64	0.83	0.64	0.72	0.87	0.84	0.76	0.61	0.85	0.75	0.70	0.11	0.49	0.81	0.77	0.91	0.86	0.84	0.85	0.73	0.79	0.77	0.74
HWE	ns	ns	ns	ns	ns	ns	ns	0.002	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Continued

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Appendix 1. Continued.

	Dee	Kam	Kit	Liz	Luz	Meg	Nev	Nil	Pis	Piz	Pon	Pul	Sai	Shu	Sum	Sys	Tai	Ten	Tor	Tul	Unj	Var	Vin	All
SSOSL438																								
N	46	40	45	23	40	47	43	33	53	21	41	39	42	20	38	42	33	45	42	41	11	42	46	
A	9	2	6	3	3	6	5	4	2	4	6	5	3	3	3	2	5	6	5	7	3	5	4	4.4
H_o	0.67	0.05	0.62	0.70	0.38	0.83	0.65	0.82	0.06	0.62	0.73	0.72	0.50	0.40	0.55	0.02	0.46	0.58	0.38	0.66	0.64	0.57	0.65	0.53
H_e	0.78	0.05	0.67	0.60	0.31	0.74	0.69	0.66	0.06	0.59	0.64	0.76	0.51	0.43	0.50	0.02	0.67	0.67	0.44	0.74	0.52	0.65	0.69	0.54
HWE	ns	ns	ns	ns	ns	ns	ns	0.01	ns	ns	ns	ns	ns	ns	ns	ns	0.022	ns	ns	ns	ns	ns	ns	ns
SLEEI84																								
N	46	38	45	26	33	43	41	27	53	16	41	39	44	20	37	42	31	42	35	42	11	47	46	
A	19	1	10	8	2	12	8	6	5	10	6	5	2	7	5	4	7	17	10	11	4	12	4	7.6
H_o	0.91	0.00	0.71	0.39	0.21	0.86	0.78	0.63	0.30	0.88	0.73	0.72	0.02	0.70	0.49	0.71	0.68	0.79	0.89	0.81	0.27	0.68	0.54	0.60
H_e	0.91	0.00	0.72	0.51	0.24	0.86	0.72	0.58	0.27	0.86	0.67	0.63	0.02	0.67	0.43	0.65	0.64	0.88	0.78	0.82	0.26	0.66	0.55	0.58
HWE	ns	ns	ns	0.008	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
SLEEN82																								
N	47	35	44	26	40	45	40	34	52	20	44	40	44	20	37	42	38	43	36	40	11	47	44	
A	7	3	6	4	2	6	2	4	2	5	6	5	1	2	5	1	4	9	3	9	5	6	2	4.3
H_o	0.77	0.37	0.71	0.69	0.53	0.69	0.13	0.38	0.06	0.60	0.84	0.55	0.00	0.65	0.62	0.00	0.21	0.77	0.33	0.70	0.55	0.72	0.16	0.48
H_e	0.74	0.33	0.71	0.53	0.47	0.70	0.12	0.38	0.06	0.67	0.80	0.52	0.00	0.50	0.55	0.00	0.24	0.79	0.34	0.79	0.70	0.79	0.15	0.47
HWE	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Ssa14																								
N	48	41	45	26	40	44	41	35	52	20	41	40	44	19	36	42	37	45	42	42	10	47	45	
A	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	4	2	2	2	2	2.2
H_o	0.35	0.05	0.62	0.12	0.40	0.39	0.44	0.74	0.42	0.15	0.42	0.25	0.14	0.26	0.50	0.02	0.14	0.51	0.45	0.38	0.30	0.53	0.20	0.34
H_e	0.38	0.05	0.51	0.11	0.40	0.34	0.37	0.50	0.40	0.14	0.41	0.40	0.13	0.24	0.50	0.02	0.13	0.48	0.46	0.50	0.27	0.50	0.28	0.33
HWE	ns	ns	ns	ns	ns	ns	ns	0.005	ns	ns	ns	0.039	ns	ns	ns	ns	ns	ns	0.001	ns	ns	ns	ns	ns
Ssa412																								
N	38	36	44	26	40	41	37	33	52	16	41	37	43	17	36	40	38	44	38	40	11	46	47	
A	3	1	5	3	3	6	2	3	4	3	4	4	2	3	4	1	2	3	4	6	2	5	3	3.3
H_o	0.40	0.00	0.71	0.31	0.45	0.73	0.49	0.27	0.73	0.56	0.51	0.62	0.14	0.41	0.58	0.00	0.55	0.48	0.47	0.55	0.09	0.52	0.49	0.44
H_e	0.42	0.00	0.67	0.49	0.38	0.66	0.48	0.29	0.55	0.46	0.67	0.57	0.13	0.35	0.52	0.00	0.46	0.59	0.51	0.65	0.09	0.65	0.54	0.44
HWE	ns	ns	ns	0.043	ns	ns	ns	ns	0.004	ns	ns	ns	ns	ns	0.015	ns	ns	ns	ns	ns	ns	ns	ns	ns
Ssa422																								
N	45	40	45	26	39	42	41	33	48	21	40	35	43	20	37	42	38	43	42	41	10	46	47	
A	10	2	8	2	3	6	4	3	5	5	5	4	5	4	4	3	4	9	8	9	2	7	5	5.1
H_o	0.87	0.35	0.69	0.42	0.74	0.71	0.81	0.64	0.63	0.52	0.53	0.23	0.91	0.75	0.60	0.93	0.74	0.74	0.76	0.73	0.40	0.61	0.66	0.65
H_e	0.85	0.32	0.67	0.38	0.65	0.68	0.72	0.58	0.71	0.61	0.55	0.21	0.78	0.73	0.49	0.67	0.64	0.80	0.73	0.76	0.34	0.59	0.65	0.61
HWE	0.029	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.000	ns	0.034	ns	ns	ns	ns	ns	ns

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

SSD30																								
N	48	38	45	26	40	47	42	35	53	19	44	40	43	20	37	42	38	44	38	38	11	46	47	
A	3	1	4	2	1	5	3	2	1	2	3	1	1	2	4	3	4	4	3	4	2	4	2	2.7
H_o	0.13	0.00	0.22	0.58	0.00	0.34	0.48	0.23	0.00	0.37	0.50	0.00	0.00	0.55	0.24	0.52	0.42	0.25	0.24	0.37	0.64	0.11	0.02	0.27
H_e	0.12	0.00	0.21	0.51	0.00	0.32	0.54	0.21	0.00	0.42	0.43	0.00	0.00	0.48	0.31	0.57	0.46	0.29	0.34	0.38	0.46	0.11	0.02	0.27
HWE	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.035	ns	ns	ns	ns	ns
SSF43																								
N	48	41	45	26	40	46	43	35	52	19	44	40	42	20	38	41	38	46	42	40	11	46	47	
A	5	2	3	2	2	4	3	4	1	4	4	3	1	4	3	3	3	3	3	5	4	3	2	3.1
H_o	0.69	0.32	0.42	0.54	0.15	0.28	0.12	0.26	0.00	0.32	0.48	0.50	0.00	0.60	0.32	0.44	0.18	0.26	0.07	0.35	0.46	0.22	0.02	0.30
H_e	0.69	0.33	0.45	0.50	0.14	0.26	0.15	0.30	0.00	0.29	0.50	0.53	0.00	0.61	0.33	0.52	0.17	0.31	0.07	0.40	0.40	0.35	0.02	0.32
HWE	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.021	ns	ns
Average of all loci																								
A	10.6	2.7	8.4	3.9	3.9	7.8	6.4	4.5	3.6	5.5	6.6	5.4	2.5	4.4	5.0	4.3	6.3	9.7	7.4	9.1	4.1	8.4	4.8	
H_o	0.72	0.38	0.70	0.54	0.47	0.68	0.63	0.61	0.43	0.62	0.67	0.57	0.29	0.57	0.61	0.49	0.59	0.66	0.59	0.67	0.62	0.61	0.47	
H_e	0.73	0.35	0.68	0.51	0.45	0.66	0.61	0.55	0.41	0.62	0.66	0.57	0.28	0.55	0.58	0.47	0.57	0.71	0.60	0.72	0.54	0.65	0.48	
HWE	ns	ns	ns	ns	ns	ns	ns	0.002	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Appendix 2. Allozyme allele frequencies and sample sizes of the 22 studied Atlantic salmon populations. *N* — number of individuals.

	Dee*	Kam	Kit	Liz	Luz	Meg	Nev	Nil	Pis	Piz	Pon	Pul	Sai	Shu	Sum	Sys	Tai	Ten	Tor†	Unj	Var	Vin‡
*sAAT-4**																						
N	47	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	641	50	29	100
100	0.837	1	0.890	0.824	0.653	0.700	0.420	0.988	1	0.530	0.950	0.941	0.778	0.559	0.625	0.393	0.321	0.866	0.732	0.450	0.750	0.850
50	0.123	0	0.098	0.059	0.347	0.275	0.090	0.012	0	0.470	0.040	0.059	0.222	0.221	0.365	0.155	0.333	0.134	0.003	0.550	0.179	0
25	0.040	0	0.012	0.118	0	0.025	0.490	0	0	0	0.010	0	0	0.221	0.010	0.452	0.346	0	0.265	0	0.071	0.150
*ESTD-2**																						
N	–	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	50	50	29	–
100	1	0.792	0.845	1	1	0.908	1	1	0.948	1	1	0.912	1	1	1	1	1	0.929	1	1	0.655	1
80	0	0.208	0.155	0	0	0.092	0	0	0.052	0	0	0.088	0	0	0	0	0	0.071	0	0	0.345	0
*IDHP-3**																						
N	47	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	645	50	29	100
116	0.121	0	0.195	0	1	0.217	0	0.762	0	0	0.740	0.471	0	0	0.144	0	0	0.012	0.003	0	0.052	0
100	0.879	1	0.805	1	0	0.783	1	0.238	1	1	0.260	0.529	1	1	0.856	1	1	0.988	0.997	1	0.948	1
*IDDH-1**																						
N	–	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	645	50	29	100
100	1	1	0.854	0.712	1	0.750	0.730	0.613	1	0.840	0.931	0.627	1	0.353	0.990	0.634	0.936	0.873	0.998	0.980	0.776	1
28	0	0	0.146	0.288	0	0.250	0.270	0.388	0	0.160	0.069	0.373	0	0.647	0.010	0.366	0.064	0.127	0.002	0.020	0.224	0
*IDDH-2**																						
N	47	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	643	50	29	100
100	0.724	0.098	0.720	0	0.306	0.833	0.200	0.538	0.241	0.610	0.670	0.775	0.006	0	0.680	0	0.013	0.645	0.658	0.320	0.517	0.760
–72	0.276	0.902	0.280	1	0.694	0.167	0.800	0.450	0.759	0.390	0.300	0.225	0.117	1	0.320	1	0.936	0.313	0.342	0.670	0.483	0.240
70	0	0	0	0	0	0	0	0.013	0	0	0.030	0	0.878	0	0	0	0.051	0.042	0	0	0	0
115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.010	0	0
*sMDH-3,4**§																						
N	47	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	645	50	29	100
100	0.993	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.992	1	1	1
115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.005	0	0	0
80	0.007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.003	0	0	0
*mMEP-2**																						
N	47	61	42	34	36	60	50	43	58	50	51	51	90	34	52	42	39	85	645	50	29	100
125	0.316	1	0.625	0.632	0.971	0.600	0.092	0.581	0.939	0.770	0.794	0.510	0	0.118	0.317	0.298	0.308	0.582	0.102	0.850	0.500	0.050
100	0.684	0	0.375	0.368	0.029	0.400	0.908	0.419	0.061	0.230	0.206	0.490	1	0.882	0.683	0.702	0.692	0.418	0.898	0.150	0.500	0.950

* Jordan *et al.* 1992 except for *ESTD-2** and *IDDH-1** E. Verspoor, personal comm.

† Bourke *et al.* 1997; Koljonen *et al.* 1999

‡ Koljonen *et al.* 1999 except for Vindelälven (see text for details)

§ *mMDH-3,4** in Jordan *et al.* 1992

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Appendix 3. Pairwise genetic distances as measured with F_{ST} (below diagonal) and **?**$_{ST}$ (above diagonal).

	Dee	Kam	Kit	Liz	Luz	Meg	Nev	Nil	Pis	Piz	Pon	Pul	Sai	Shu	Sum	Sys	Tai	Ten	Tor	Tul	Unj	Var	Vin
Dee		0.112	0.052	0.081	0.159	0.049	0.044	0.112	0.140	0.046	0.069	0.071	0.168	0.087	0.115	0.292	0.318	0.041	0.074	0.034	0.115	0.047	0.052
Kam	0.268		0.217	0.210	0.216	0.168	0.184	0.182	0.126	0.165	0.189	0.212	0.410	0.395	0.327	0.424	0.440	0.099	0.149	0.144	0.231	0.218	0.219
Kit	0.097	0.262		0.212	0.302	0.039	0.118	0.104	0.295	0.098	0.066	0.067	0.394	0.223	0.059	0.347	0.354	0.092	0.165	0.104	0.147	0.007	0.197
Liz	0.238	0.413	0.223		0.212	0.187	0.099	0.288	0.147	0.183	0.208	0.164	0.367	0.173	0.252	0.294	0.359	0.110	0.168	0.195	0.313	0.239	0.185
Luz	0.234	0.304	0.192	0.332		0.254	0.149	0.393	0.171	0.277	0.304	0.245	0.331	0.381	0.399	0.351	0.356	0.129	0.125	0.237	0.385	0.317	0.191
Meg	0.118	0.275	0.046	0.194	0.196		0.115	0.161	0.269	0.050	0.081	0.131	0.367	0.188	0.110	0.351	0.362	0.053	0.149	0.049	0.096	0.038	0.154
Nev	0.163	0.325	0.114	0.138	0.224	0.101		0.202	0.208	0.115	0.174	0.084	0.178	0.119	0.205	0.187	0.216	0.081	0.047	0.133	0.142	0.134	0.089
Nil	0.193	0.334	0.127	0.317	0.274	0.178	0.225		0.313	0.232	0.100	0.132	0.492	0.455	0.235	0.423	0.438	0.137	0.203	0.163	0.269	0.119	0.304
Pis	0.245	0.226	0.227	0.389	0.209	0.249	0.263	0.315		0.251	0.242	0.251	0.366	0.385	0.374	0.438	0.469	0.129	0.227	0.227	0.409	0.304	0.269
Piz	0.162	0.354	0.107	0.231	0.230	0.062	0.151	0.245	0.301		0.178	0.159	0.429	0.173	0.126	0.394	0.377	0.106	0.173	0.098	0.062	0.075	0.136
Pon	0.138	0.328	0.050	0.232	0.207	0.077	0.136	0.134	0.283	0.111		0.118	0.430	0.348	0.181	0.413	0.441	0.039	0.191	0.082	0.277	0.086	0.238
Pul	0.151	0.368	0.103	0.269	0.252	0.137	0.129	0.184	0.272	0.194	0.111		0.369	0.225	0.136	0.290	0.303	0.109	0.135	0.183	0.228	0.108	0.194
Sai	0.329	0.500	0.303	0.469	0.384	0.319	0.306	0.448	0.350	0.388	0.353	0.353		0.566	0.500	0.468	0.480	0.221	0.196	0.278	0.571	0.370	0.220
Shu	0.220	0.443	0.227	0.132	0.329	0.206	0.168	0.313	0.425	0.240	0.225	0.279	0.497		0.200	0.378	0.382	0.183	0.265	0.227	0.361	0.214	0.184
Sum	0.161	0.310	0.064	0.301	0.215	0.083	0.170	0.226	0.270	0.128	0.111	0.190	0.362	0.295		0.400	0.392	0.190	0.297	0.209	0.220	0.073	0.295
Sys	0.287	0.452	0.246	0.214	0.334	0.250	0.158	0.343	0.415	0.271	0.265	0.284	0.450	0.236	0.315		0.043	0.279	0.264	0.400	0.370	0.388	0.403
Tai	0.219	0.348	0.181	0.181	0.221	0.175	0.090	0.270	0.267	0.180	0.205	0.207	0.320	0.200	0.224	0.170		0.312	0.285	0.417	0.330	0.389	0.416
Ten	0.064	0.261	0.054	0.208	0.180	0.069	0.114	0.147	0.222	0.110	0.068	0.130	0.286	0.217	0.128	0.255	0.180		0.087	0.035	0.156	0.101	0.125
Tor	0.127	0.297	0.097	0.214	0.239	0.119	0.089	0.182	0.225	0.187	0.140	0.143	0.318	0.237	0.183	0.245	0.186	0.086		0.133	0.161	0.178	0.077
Tul	0.079	0.249	0.046	0.205	0.157	0.069	0.097	0.136	0.203	0.107	0.064	0.108	0.280	0.216	0.117	0.230	0.159	0.019	0.087		0.163	0.075	0.110
Unj	0.213	0.474	0.170	0.305	0.353	0.125	0.228	0.313	0.413	0.077	0.174	0.272	0.494	0.335	0.225	0.357	0.279	0.154	0.244	0.159		0.122	0.222
Var	0.118	0.274	0.011	0.241	0.199	0.056	0.126	0.144	0.237	0.111	0.055	0.129	0.324	0.245	0.068	0.269	0.202	0.066	0.119	0.057	0.178		0.177
Vin	0.209	0.399	0.142	0.251	0.272	0.140	0.078	0.265	0.299	0.225	0.165	0.143	0.312	0.268	0.207	0.279	0.206	0.160	0.101	0.144	0.308	0.149	

I accept this page: ❑ without changes, ❑ with changes (❑ I want to see the corrected version)

Date and signature

Appendix 4. Pairwise geographic distances (km, below the diagonal) and genetic distances (D_{CE}, above the diagonal) of anadromous populations.

	Dee	Kit	Meg	Nev	Nil	Piz	Pon	Pul	Sum	Ten	Tor	Tul	Unj	Var	Vin
Dee		0.411	0.412	0.477	0.539	0.497	0.441	0.451	0.487	0.342	0.406	0.387	0.561	0.409	0.474
Kit	3113		0.265	0.395	0.409	0.405	0.307	0.362	0.327	0.332	0.389	0.309	0.474	0.189	0.394
Meg	2937	237		0.382	0.467	0.350	0.331	0.413	0.365	0.341	0.389	0.331	0.447	0.283	0.372
Nev	2260	4659	4483		0.540	0.430	0.431	0.417	0.452	0.430	0.351	0.406	0.494	0.391	0.294
Nil	3253	181	377	4799		0.546	0.394	0.449	0.479	0.439	0.468	0.426	0.607	0.418	0.513
Piz	3771	1537	1353	5317	1677		0.426	0.473	0.417	0.437	0.483	0.429	0.315	0.397	0.446
Pon	3217	173	340	4763	169	1641		0.358	0.355	0.324	0.452	0.327	0.493	0.297	0.431
Pul	3256	183	380	4802	48	1680	164		0.433	0.418	0.419	0.403	0.534	0.382	0.385
Sum	3253	250	363	4799	277	1677	130	272		0.411	0.481	0.419	0.489	0.330	0.461
Ten	2207	995	819	3753	1135	1656	1099	1138	1135		0.382	0.288	0.499	0.328	0.438
Tor	2459	4858	4682	1143	4998	5516	4962	5001	4998	3952		0.383	0.544	0.399	0.295
Tul	2479	766	590	4025	906	1500	870	909	906	361	4224		0.502	0.328	0.418
Unj	4749	2515	2331	6295	2655	1248	2619	2658	2655	2634	6494	2478		0.467	0.512
Var	3136	63	260	4682	204	1560	196	206	273	1018	4881	789	2538		0.389
Vin	2197	4596	4420	889	4736	5254	4700	4739	4736	3690	300	3962	6232	4619	

Appendix 5. Mitochondrial DNA haplotype frequencies and sample sizes (*N*) of the studied Atlantic salmon populations (Nielsen *et al.* 1996, Nilsson *et al.* 2001, Asplund *et al.* 2004, except for Luzhma and Kamennoe).

	Dee†	Kam	Kit§	Liz‡	Luz	Meg§	Nev‡	Nil§	Pis§	Piz§	Pon§	Pul§	Sai*‡	Shu‡	Sum§	Sys§	Tul§	Ten§	Tor‡	Unj§	Var§	Vin‡
N	111	61	46	20	40	58	45	43	59	21	51	50	81	17	52	42	36	239	85	11	31	108
AAAA	0.09	0	0	0	0	0	0	0	0	0	0	0	0.99	0	0	0.88	0	0.05	0.05	0	0	0.67
AABA	0.07	1.0	0.85	0.95	0	0.31	0.96	0	0.59	0.90	0.59	0.20	0.01	1.0	0.06	0.12	0.03	0.10	0.86	1.0	0.74	0.29
BBBA	0.25	0	0.09	0	0	0.07	0	0	0	0	0	0	0	0	0	0	0.19	0.37	0	0	0.16	0.04
BBBB	0.54	0	0.06	0.05	1.0	0.62	0.04	1.0	0.41	0.10	0.41	0.80	0	0	0.94	0	0.72	0.47	0.06	0	0.10	0
Other	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.06	0.02	0	0	0	0

* Palva *et al.* 1989
† Nielsen *et al.* 1996
‡ Nilsson *et al.* 2001
§ Asplund *et al.* 2004