Protein and microsatellite single locus variability in *Salmo salar* L. (Atlantic salmon)

J. A. SÁNCHEZ*, C. CLABBY†, D. RAMOS, G. BLANCO, F. FLAVIN†, E. VÁZQUEZ & R. POWELL†

Departamento de Biología Funcional, Area de Genética, Universidad de Oviedo, 33071 Oviedo, Spain and †Recombinant DNA Group, Department of Microbiology, University College Galway, Ireland

We describe genetic variation at three microsatellite single loci and six allozyme loci of seven natural Atlantic salmon populations from Ireland and Spain. A comparison of genetic variability detected at both types of loci is performed. Also, the relative value of microsatellite single locus variability with regard to the identification of individual salmon populations is assessed. Microsatellite loci are shown to display higher levels of variation than allozyme loci. Mean number of alleles (6 ± 1.53) and heterozygosity (0.46 ± 0.04) at microsatellite loci are greater than those found for allozymes (1.85 ± 0.05 and 0.21 ± 0.03, respectively), and some microsatellite alleles appear to be specific for a location or geographical area. Allozyme and microsatellite variation show the same pattern of differentiation between populations with Irish and Spanish populations grouped into different clusters. However, greater values of genetic distance were found among microsatellite ($D = 0.0747 \pm 0.011$) rather than among allozyme loci ($D = 0.0449 \pm 0.008$). These results provide evidence that microsatellite-based analysis of genetic variation will be useful in the identification of individual populations of Atlantic salmon.

Keywords: allozyme variation, microsatellite variation, population genetics, *Salmo salar*.

Introduction

A knowledge of genetic diversity and population structure is a prerequisite for the successful management of conservation programmes for exploited species such as Atlantic salmon (*Salmo salar*, L.). Over the past 20 years, most population genetic analyses of Atlantic salmon have been performed by studying variation at protein coding loci or allozymes (Cross & Ward, 1980; Ståhl, 1987; Verspoor, 1988; Davidson *et al.*, 1989; Koljonen, 1989; McElligott & Cross, 1991; Sánchez *et al.*, 1991, 1993; Blanco *et al.*, 1992; Jordan *et al.*, 1992; Elo *et al.*, 1994; O'Connell *et al.*, 1995). These studies have shown that the Atlantic salmon can be subgrouped into three races, North American, Western European and Baltic Sea (Ståhl, 1987; Verspoor, 1988). However, these studies also reveal that levels of genetic variability and differentiation within and among populations are low. In fact, of 50 or more protein coding loci examined, only 12 have been found to be polymorphic, each segregating for either two or three alleles,

and six of them ($_sAAT-4*$; *IDPH-3**; *MDH-3,4**; *m-MEP-2**; *IDDH-1** and *IDDH-2**) account for more than 98 per cent of the total absolute gene diversity of this species (Cross & Ward, 1980; Ståhl, 1987; Davidson *et al.*, 1989). Thus, allozyme data have limited values as diagnostic genetic markers and alternative methods which permit us to distinguish between salmon stocks are required (Davidson *et al.*, 1989; Blanco *et al.*, 1992).

The occurrence of repetitive DNA elements of various types have been evidenced by DNA reassociation experiments for more than two decades (Britten & Kohne, 1968). One type of repetitive element, referred to as microsatellites, corresponds to tandemly repeated DNA motifs of one to five base pairs in length evenly distributed in the genome of all investigated eukaryotic species (Hamada *et al.*, 1982; Weber, 1990; Wong *et al.*, 1990; Wintero *et al.*, 1992; Estoup *et al.*, 1993; Slettan *et al.*, 1993). Genetic variability at microsatellite single loci can be rapidly analysed using the polymerase chain reaction (PCR) directed by a pair of amplification oligodeoxynucleotide primers derived from unique sequences flanking the repetitive motif. The PCR

amplification products are then examined and sized after electrophoresis on agarose or polyacrylamide gels. In all investigated species, microsatellites have been found to be highly variable among individuals (Hamada *et al.*, 1982; Hearne *et al.*, 1992; Wintero *et al.*, 1992; Estoup *et al.*, 1993; Hughes & Queller, 1993). The high degree of polymorphism and random distribution in the genome make microsatellite loci promising candidates for use as genetic markers in a large scope of genetic studies such as linkage analysis, genome mapping, population structure and stock composition (Estoup *et al.*, 1993; Hughes & Queller, 1993).

This paper describes the analysis of microsatellite single locus variability from seven natural populations of Atlantic salmon sampled from Spanish and Irish rivers. Comparison is made with allozyme variation detected in these populations in order to assess the relative values of both techniques in population genetic studies and the potential application of microsatellites for identification of individual populations of Atlantic salmon.

Materials and methods

Samples

Adult fish (1 and 2 SW) from four Spanish rivers (Bidasoa, Sella, Narcea and Esva) and salmon parr (0$^+$ and 1$^+$) from three Irish rivers (River Bann on the Slaney system, River Bunree on the Moy and from the River Shannon) were collected and analysed for both protein and microsatellite polymorphisms. Figure 1 details the geographical locations of the sampled rivers and the numbers of individuals assayed per river are shown in Tables 2 and 3.

Protein analysis

Tissue samples of liver and muscle were removed from specimens and analysed by horizontal starch gel (12 per cent) electrophoresis. The following six loci were studied: $_s$AAT-4*, IDPH-3*, MDH-3,4*, m-MEP-2*, IDDH-1* and IDDH-2*. Two buffer systems were used in the electrophoretic analysis; a continuous citrate-aminopropylmorpholine system (Clayton & Tretiak, 1972) for $_s$AAT-4*, IDPH-3*, MDH-3,4* and m-MEP-2*, and a discontinuous Tris-citrate-borate system (Ridgway *et al.*, 1970) for IDDH-1* and IDDH-2*. Specific staining for enzymes was carried out using the methods of Cross & Ward (1980) and Aebersold *et al.* (1987). The nomenclature for designation of loci and alleles
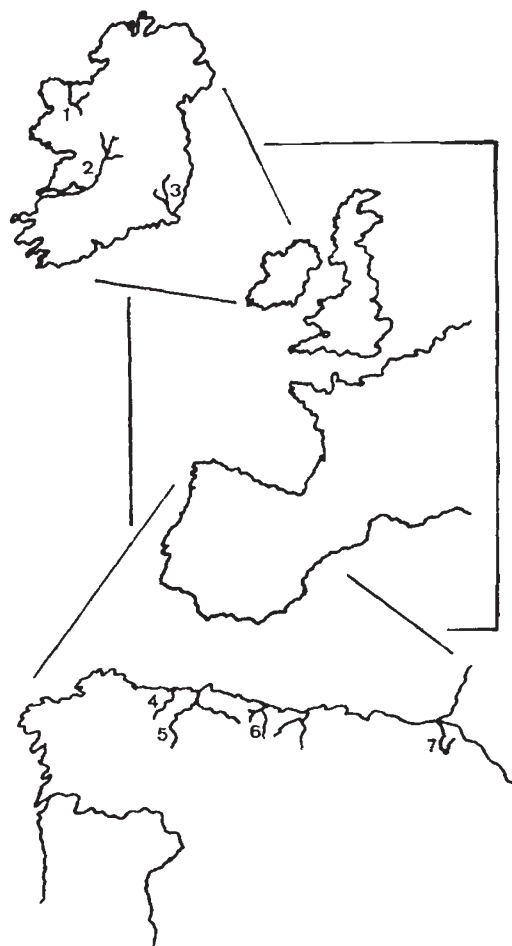
Fig. 1 Geographical location of samples. The numbers represent: 1, River Moy; 2, River Shannon; 3, River Slaney; 4, River Esva; 5, River Narcea; 6, River Sella and 7, River Bidasoa.

follows the recommendations of Shaklee *et al.* (1990).

Isolation and identification of microsatellites

Total salmon genomic DNA for the isolation of microsatellites and the examination of natural populations was prepared as described by Mooney *et al.* (1995). Briefly, 0.5–1.0 mL of whole blood was diluted to 50 mL in lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM EDTA, 2 mM MgCl$_2$, 400 mM NaCl, 0.2 per cent SDS and 0.2 mg mL^{-1} Protease K), incubated overnight at 37°C, followed by phenol/chloroform extraction and ethanol precipitation. Total genomic DNA was digested with *Alu*I and *Hae*III to generate small blunt-ended fragments. These fragments were fractionated on 1 per cent low melt agarose gel (FMC BioProducts, Rockland, USA). Fragments ranging between 100 and 500 bp in

length were excised from the gel, purified (Promega Magic Preps., Promega, Madison, USA) and ligated to the *Sma*I site of plasmid pBGS8 (Spratt *et al.*, 1986). The ligation mixture was used to transform competent *E. coli* JM109 by electroporation. Recombinant clones were detected as white colonies on LB plates containing 40 μg mL^{-1} XGal, 0.2 mM IPTG and 50 μg mL^{-1} kanamycin and were subsequently replica plated to nitrocellulose filters and screened with radiolabelled (GT)$_{10}$ or (CT)$_{10}$ oligodeoxynucleotides. Positive clones were confirmed by probing the plasmid inserts by Southern blot and the nucleotide sequence of the inserts was determined by the chain termination dideoxy method (Sanger *et al.*, 1977). Two PCR amplification primers were designed in the regions flanking the microsatellite repeats of clones F43, D30 and 20.19. The resulting amplified fragments were 121, 240 and 96 bp in length, respectively. The DNA sequences of clones F43, D30 and 20.19 have been submitted to GenBank and assigned accession numbers U37494, U37493 and U37490, respectively.

Microsatellite analysis

Microsatellite polymorphism was analysed by specific PCR performed in an Omnigene thermocycler (Omnigene, Hybaid, Teddington, UK). The 50 μL reaction mixture contained 50–100 ng of genomic DNA template, 100 ng of each specific primer, 200 μM of dCTP, dTTP, dGTP and dATP, 1.5 mM MgCl$_2$, 10X Promega reaction buffer and one unit of Promega Taq polymerase. After a denaturating step of 5 min at 94°C, samples were processed through 30 cycles consisting of 30 s at 94°C, 30 s at the specific annealing temperature and 30 s at 72°C. The final chain elongation step was lengthened to 10 min. Primer sequences and specific annealing temperatures are shown in Table 1. Aliquots of the amplification products were resolved on denaturing acrylamide gels containing 6 M urea and 30 per cent formamide. The products were

visualized by silver staining (Promega Silver Sequence™ DNA Staining).

Statistical analysis

Statistical analyses were performed using the BIOSYS-1 computer package (Swofford & Selander, 1989). Heterogeneity of allele frequencies among samples was tested using the Pearson χ^2 test. The amount of genetic variation in each locus and population was indicated by the observed and expected heterozygosities. As in most of the cases (42 per cent at enzyme and 90 per cent at microsatellite loci) the genotype number (observed or expected) was <5, comparison with the Hardy–Weinberg expectation was not performed. *F*-statistics of Wright (1978) were used in the genetic diversity analyses. Overall genetic differences between samples were expressed as genetic distances (Nei, 1972), and a UPGMA dendrogram (Sneath & Sokal, 1973) was constructed to visualize these differences.

Results

Protein analysis

The degrees of polymorphism observed for the six protein coding loci examined were similar to those described for other European natural populations of Atlantic salmon (Cross & Ward, 1980; Ståhl, 1987; Crozier & Moffet, 1989; Koljonen, 1989; McElligott & Cross, 1991; Sánchez *et al.*, 1991; Blanco *et al.*, 1992; Jordan *et al.*, 1992; Elo *et al.*, 1994; O'Connell *et al.*, 1995). Only two alleles were found at each locus, a common allele (100) and an alternative whose frequencies are shown in Table 2. Highly significant differences among all samples were found for five of the six loci ($\chi^2_6 = 15.722$, $P = 0.015$ for *MDH-3,4**; $\chi^2_6 = 86.884$, $P<0.001$ for *m-MEP-2**; $\chi^2_6 = 35.985$, $P<0.001$ for *IDPH-3**; $\chi^2_6 = 25.849$, $P<0.001$ for $_sAAT-4*$ and $\chi^2_6 = 15.852$, $P = 0.014$ for *IDDH-2**). Spanish populations showed little signi-

Table 1 Repeat motifs, PCR primer sequences and annealing temperatures

Microsatellite loci	Motif	Primer sequences	Annealing temperature (°C)
μ-F43*	(AC/TG)$_n$	5'-AGC GGC ATA ACG TGC TGT GT-3' 5'-GAG TCA CTC AAA GTG AGG CC-3'	60
μ-20.19*	(AC/TG)$_n$	5'-TCA ACC TGG TCT GCT TCG AC-3' 5'-CTA GTT TCC CCA GCA CAG CC-3'	62
μ-D30*	(AG/TC)$_n$	5'-AGC AGT AAA GAG AGA GAC TG-3' 5'-TGT TGA CTT CCT TCC CCA AG-3'	53

ficant differentiation ($\chi^2_{18} = 29.679$; $P < 0.049$) and no differences between gene frequencies at any locus were found. However, significant heterogeneity ($\chi^2_{10} = 39.175$; $P < 0.001$) was found among Irish samples, primarily because of significant differences in the frequencies at the loci $_sAAT$-4^* ($\chi^2_2 = 17.920$; $P < 0.001$) and $IDPH$-3^* ($\chi^2_2 = 7.393$; $P = 0.024$).

The most remarkable differences between the Irish and Spanish populations were observed at the MDH-$3,4^*$ and m-MEP-2^* loci. MDH-$3,4^*$ was monomorphic in Irish populations whereas Spanish populations were polymorphic (Table 2). This locus shows lower levels of variability in Irish and other European natural populations (Cross & Ward, 1980; Ståhl, 1987; Crozier & Moffet, 1989; McElligott & Cross, 1991) and the frequency of the alternative allele (87) in Spanish populations is much higher

than in other European populations. This may be interpreted as a special feature of natural populations of Spanish rivers (Sánchez *et al.*, 1991, 1993). For the m-MEP-2^* locus, one of the most widespread polymorphisms detected in Atlantic salmon, a cline of frequencies was described by Verspoor & Jordan (1989). They found that southern European populations show higher frequencies of the m-MEP-2^*-100 allele than do northern populations. In this study, Irish and Spanish populations also differ in this manner, but the Moy and Shannon samples from Ireland show lower levels of the 100 allele frequency than those observed in other studies made on these rivers (Cross & King, 1983; McElligott & Cross, 1991; Galvin *et al.*, 1994), and also in other populations of this area (Ireland and Britain) where both alleles (100 and 125) have frequencies of

Table 2 Observed allele frequencies at the six protein coding loci in the seven natural populations of Atlantic salmon

Locus	Allele	Populations						
		Moy (20)	Shannon (20)	Slaney (22)	Narcea (18)	Bidasoa (31)	Sella (43)	Esva (77)
MDH-$3,4^*$	100	1.000	1.000	1.000	0.972	0.984	0.895	0.948
	87	0.000	0.000	0.000	0.028	0.016	0.105	0.052
	H_o	0.000	0.000	0.000	0.056	0.032	0.116	0.078
	H_e	0.000	0.000	0.000	0.054	0.032	0.187	0.098
m-MEP-2^*	100	0.342	0.325	0.523	0.889	0.855	0.881	0.760
	125	0.658	0.675	0.477	0.111	0.145	0.119	0.240
	H_o	0.368	0.550	0.500	0.222	0.226	0.143	0.377
	H_e	0.450	0.439	0.499	0.198	0.248	0.210	0.365
$IDPH$-3^*	100	1.000	0.955	0.833	0.972	0.984	0.988	1.000
	116	0.000	0.045	0.167	0.028	0.016	0.012	0.000
	H_o	0.000	0.910	0.333	0.056	0.032	0.023	0.000
	H_e	0.000	0.087	0.278	0.054	0.032	0.023	0.000
$_sAAT$-4^*	100	0.975	0.611	0.614	0.750	0.581	0.605	0.737
	50	0.025	0.389	0.386	0.250	0.419	0.395	0.263
	H_o	0.050	0.444	0.591	0.357	0.323	0.465	0.421
	H_e	0.049	0.475	0.474	0.375	0.487	0.478	0.388
$IDDH$-1^*	100	0.842	0.588	0.684	0.650	0.621	0.605	0.667
	-72	0.158	0.412	0.316	0.350	0.379	0.395	0.333
	H_o	0.211	0.353	0.421	0.100	0.345	0.316	0.182
	H_e	0.266	0.484	0.432	0.455	0.471	0.478	0.444
$IDDH$-2^*	100	0.947	1.000	0.895	0.800	0.897	0.769	0.826
	28	0.053	0.000	0.105	0.200	0.103	0.231	0.174
	H_o	0.000	0.000	0.000	0.200	0.138	0.256	0.227
	H_e	0.100	0.000	0.188	0.320	0.185	0.355	0.288
	H_o(avg)	0.105 (± 0.062)	0.240 (± 0.098)	0.308 (± 0.103)	0.165 (± 0.048)	0.183 (± 0.056)	0.220 (± 0.065)	0.214 (± 0.067)
	H_e(avg)	0.144 (± 0.073)	0.248 (± 0.099)	0.312 (± 0.079)	0.243 (± 0.069)	0.243 (± 0.082)	0.289 (± 0.074)	0.264 (± 0.072)

Number of fish examined per river is in parenthesis.
H_o, H_e: observed and expected heterozygosity at each locus.
H_o (avg) and H_e (avg): average heterozygosity per population.
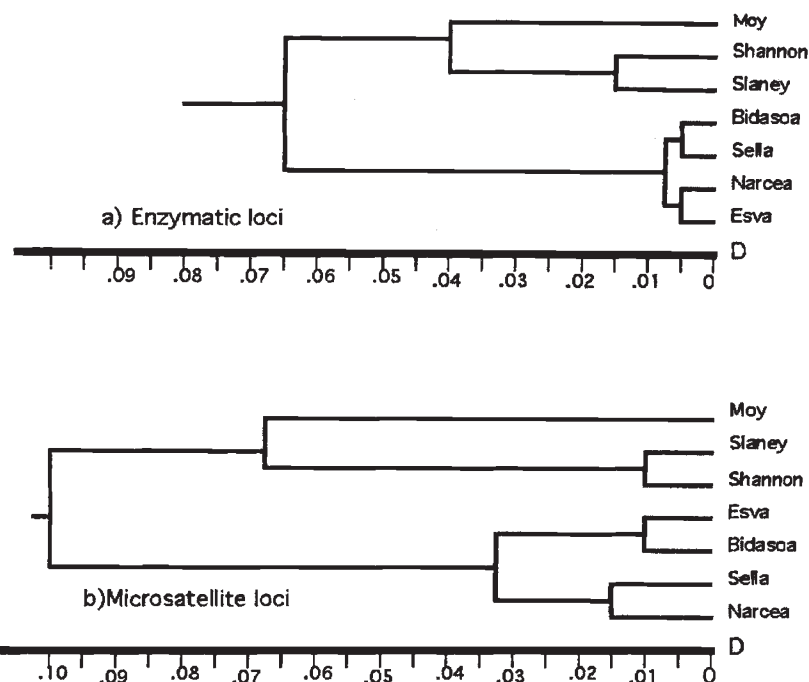
Fig. 2 UPGMA dendrogram for Atlantic salmon based on genetic distance estimated from allozymes (a) and microsatellite (b) loci.

around 0.5 (Cross & Ward, 1980; Cross & King, 1983; Crozier & Moffet, 1989; Hovey *et al.*, 1989; Verspoor & Jordan, 1989; Jordan *et al.*, 1992). These differences may result from the number of individuals analysed in this work.

The amount of genetic variation, measured by average heterozygosity, varies between 10.5 and 30.8 per cent and shows lower levels of variation among Spanish populations (19.5 per cent) compared with Irish populations (21.7 per cent) (Table 2). Genetic distances (Nei, 1972) between populations ranged from 0.004 (Esva vs. Narcea) to 0.114 (Moy vs. Sella). Spanish populations have a lower mean genetic distance between them ($D = 0.007$) than that observed between Irish populations ($D = 0.033$). The UPGMA dendrogram (Fig. 2a) constructed using these genetic distances shows that the seven salmon populations grouped into two regional clusters, and the mean genetic distance between the Spanish and Irish populations was estimated as $D = 0.067$. Similar to previously described data (Blanco *et al.*, 1992), it appears that genetic distance between populations increases with increasing geographical distance ($r = 0.769$).

Microsatellite analysis

The three microsatellite loci examined display a high level of polymorphism in the seven investigated populations (Table 3). As expected, alleles differed

by multiples of two base-pairs in length which suggests that all alleles at these loci differ through variation in the number of repeat units in the microsatellite motif. At the μ-20.19* locus four alleles were found ranging in length from 96 to 102 bp (Table 3). The μ-20.19*-96 allele is the most common allele in all populations, ranging in frequency from 0.5 to 0.714 (Table 3). Although no significant heterogeneity was found from comparisons of all samples ($\chi^2_{18} = 21.092$, $P = 0.275$), and among rivers within the same geographical area ($\chi^2_9 = 4.865$, $P = 0.845$; $\chi^2_6 = 6.211$, $P = 0.400$), significant differences between both areas exist ($\chi^2_3 = 11.70$, $P = 0.011$). Thus, in Irish populations the μ-20.19*-98 allele has the second highest frequency (mean frequency 0.254) whereas the allele μ-20.19*-102 has the second highest frequency (0.145) in Spanish populations (Table 3). The observed heterozygosity varies between 50 and 79 per cent and in all cases these values are higher than the expected ones (Table 3).

Nine alleles were found at the μ-F.43* locus, ranging in allelic forms from 103 to 127 bp in length (Table 3). Another four possible alleles not found in these populations (105, 107, 111 and 113 bp in length) could be expected at this locus. Significant frequency differences were found in the comparison between all populations ($\chi^2_{48} = 57.988$, $P < 0.001$), but no significant heterogeneity was found in the comparison of populations of the same geographical

Table 3 Observed allele frequencies and heterozygosity at three microsatellite loci in the seven natural populations of Atlantic salmon

Locus	Allele	Populations						
		Moy (19)	Shannon (19)	Slaney (22)	Narcea (9)	Bidasoa (16)	Sella (14)	Esva (12)
μ-20.19*	96	0.666	0.500	0.568	0.555	0.714	0.679	0.546
	98	0.222	0.237	0.295	0.111	0.036	0.107	0.136
	100	0.056	0.132	0.114	0.167	0.071	0.143	0.136
	102	0.056	0.132	0.023	0.167	0.179	0.071	0.182
	H_o	0.500	0.789	0.636	0.778	0.571	0.500	0.636
	H_e	0.500	0.659	0.576	0.623	0.442	0.503	0.632
μ-F-43*	103	0.237	0.395	0.409	0.055	0.187	0.036	0.208
	109	0.000	0.026	0.000	0.111	0.000	0.036	0.000
	115	0.237	0.211	0.205	0.556	0.469	0.750	0.458
	117	0.026	0.000	0.000	0.000	0.000	0.000	0.000
	119	0.000	0.000	0.000	0.000	0.000	0.000	0.042
	121	0.500	0.342	0.295	0.167	0.313	0.142	0.292
	123	0.000	0.026	0.068	0.000	0.000	0.000	0.000
	125	0.000	0.000	0.023	0.111	0.031	0.000	0.000
	127	0.000	0.000	0.000	0.000	0.000	0.036	0.000
	H_o	0.737	0.632	0.818	0.333	0.688	0.429	0.833
	H_e	0.637	0.681	0.698	0.636	0.646	0.413	0.660
μ-D-30*	226	0.139	0.000	0.000	0.000	0.000	0.000	0.000
	234	0.639	0.917	0.881	1.000	1.000	1.000	0.958
	236	0.056	0.083	0.048	0.000	0.000	0.000	0.042
	240	0.083	0.000	0.071	0.000	0.000	0.000	0.000
	242	0.083	0.000	0.000	0.000	0.000	0.000	0.000
	H_o	0.500	0.167	0.143	0.000	0.000	0.000	0.083
	H_e	0.556	0.153	0.217	0.000	0.000	0.000	0.080
	H_o(avg)	0.579 ($\pm$0.079)	0.529 ($\pm$0.187)	0.532 ($\pm$0.202)	0.370 ($\pm$0.225)	0.420 ($\pm$0.212)	0.310 ($\pm$0.156)	0.518 ($\pm$0.224)
	H_e(avg)	0.564 ($\pm$0.040)	0.498 ($\pm$0.173)	0.497 ($\pm$0.145)	0.420 ($\pm$0.210)	0.366 ($\pm$0.191)	0.305 ($\pm$0.155)	0.457 ($\pm$0.189)

Number of fish examined is in parenthesis.
H_o, H_e: observed and expected heterozygosity at each locus.
H_o (avg), H_e (avg): average heterozygosity per population.

area ($\chi^2_{12} = 13.554$, $P = 0.3297$ for Irish and $\chi^2_{18} = 27.239$, $P = 0.074$ for Spanish populations). The largest differences between both areas are because in Irish populations the most common alleles are the 103 and 121 alleles, whereas the 115 allele is the most common in Spanish populations (Table 3). Also, some alleles were detected only in Irish (117 and 123) or Spanish (119 and 127) populations, but these 'specific' alleles show lower frequency levels among the respective populations (Table 3). At this microsatellite locus, all populations showed high levels of heterozygosity, particularly in the Esva population where 83 per cent of individuals were heterozygous and, with the exception of the Shannon and Narcea populations, observed heterozygosity is higher than expected (Table 3).

At the μ-D30* locus, five alleles were found, ranging from 226 to 242 bp in length (Table 3). As for locus μ-F-43*, other alleles not found in this work could be expected at the μ-D30* locus. This locus showed the lowest level of variability of the microsatellite loci examined. Thus three Spanish populations were monomorphic and more than two alleles were found only in the Irish Moy and Slaney populations (Table 3). The μ-D30*-234 allele is the most common allele in all populations. The test for homogeneity of allele frequencies among all populations showed statistically significant differences ($\chi^2_{24} = 59.880$, $P < 0.001$). Significant differences among Irish populations also exist ($\chi^2_8 = 22.635$, $P = 0.004$) but not among Spanish populations ($\chi^2_3 = 3.282$, $P = 0.350$). Heterozygosity values vary

between 0 and 50 per cent with lower levels in Spanish (2 per cent) than in Irish populations (27 per cent) (Table 3).

Average heterozygosity at the microsatellite loci is much higher than that detected at allozyme loci and varies between 31 and 58 per cent (Table 3).

Using alleles frequencies determined at the three microsatellite loci, genetic distances (Nei, 1972) between samples were calculated and an UPGMA dendrogram was constructed (Fig. 2b). The dendrogram shows a similar arrangement to that constructed with allozyme frequencies with the natural populations grouped in two regional clusters. Mean genetic distance between the Spanish populations ($D = 0.0268$) is lower than that found between the Irish populations ($D = 0.0493$) and the mean genetic distance between the Spanish and Irish populations is 0.0747.

Discussion

Since the early 1960s, the dominant method for the analysis of genetic variation in natural populations has been the screening of protein variation using electrophoretic techniques. One of the first identified protein loci showing polymorphisms in Atlantic salmon was transferrin, a serum protein (Payne *et al.*, 1971). Subsequently, Atlantic salmon populations have been extensively analysed for genetic variability using approximately 50 protein loci (see reviews of Ståhl, 1987; Davidson *et al.*, 1989). These studies provide evidence that Atlantic salmon populations are substructured into multiple, genetically distinct stocks more or less reproductively isolated. However, the low levels of variation detected at these allozyme loci do not allow clear differentiation between Atlantic salmon populations and other types of genetic marker are required (Davidson *et*

al., 1989; Taggart & Ferguson, 1990; Blanco *et al.*, 1992).

Microsatellite loci, regions of DNA containing tandem repeats of a short sequence motif, have been shown to be a rich source of of highly polymorphic genetic markers in humans and in other mammalian species and they appear to be promising markers for the detection of genetic variation in species where other methods have only revealed low variation levels (Estoup *et al.*, 1993; Hughes & Queller, 1993; Presa *et al.*, 1994). In this study, all the microsatellite loci screened were polymorphic, showing between four and nine alleles per locus, with an average of 6 ± 1.53 alleles per locus. For two of the microsatellite loci, μ-F-43* and μ-D30, some of the alleles appear specific for a location or geographical area (Tables 3 and 4). The microsatellite data expand the data found in the same populations using protein-coding loci, which exhibit only one or two alleles per locus (mean 1.85 ± 0.05) which are common for all populations (Tables 2 and 4). Also, at the microsatellite loci, the mean heterozygosity per population (0.46 ± 0.04) was significantly higher ($t_5 = 5.997$, $P = 0.01$) than that detected at the allozyme loci (0.21 ± 0.03) (Tables 2–4). Note that the levels of polymorphism described for the allozyme loci are an overestimation of the general level of variability at allozyme loci, as these particular loci were chosen for this analysis because they show variation in most previous investigations of salmon populations and are thought to account for more than 98 per cent of the total absolute gene diversity of the species (Cross & Ward, 1980; Ståhl, 1987; Davidson *et al.*, 1989). More realistic data about the level of variability for allozyme loci can be obtained in a review of other studies where more loci (between 19 and 54), chosen at random, have been analysed. Thus, in these works, the number of alleles per locus, the

Table 4 Comparison between genetic variability of Atlantic salmon at microsatellite and allozyme loci

	Microsatellite loci		Enzyme loci	
	Mean	Range	Mean	Range
Alleles/locus	6 ± 1.53	4–9	2	2
Alleles/locus/population	3.57 ± 0.17	3–4.33	1.85 ± 0.05	1.8–2
Loci polymorphic: no criterion	87.71 ± 6.73	66.67–100	85.71 ± 5.66	66.67–100
0.95 criterion	80.95 ± 6.73	66.67–100	69.05 ± 5.66	50–83.33
Heterozygosity per population: observed	0.46 ± 0.04	0.31–0.58	0.21 ± 0.03	0.11–0.31
expected	0.44 ± 0.03	0.31–0.56	0.24 ± 0.02	0.14–0.31
Genetic distance (Nei, 1972)	0.075 ± 0.011	0.011–0.176	0.045 ± 0.008	0.004–0.114
Diagnostic value (Ayala & Powell, 1972)	0.687 ± 0.035	0.65–0.73	0.593 ± 0.039	0.53–0.78

percentage of polymorphic loci and the average percentage heterozygosity vary between 1.08 and 1.28, 7 and 28.94 per cent and between 1.1 and 7.2 per cent, respectively (Cross & Ward, 1980; Ståhl, 1987; Koljonen, 1989; McElligott & Cross, 1991; Sánchez *et al.*, 1991; Jordan *et al.*, 1992; Elo *et al.*, 1994; O'Connell *et al.*, 1995). As this study is the first report of a microsatellite-based analysis of genetic variation in Atlantic salmon populations there are insufficient data to state categorically that levels of polymorphism at microsatellite loci are appreciably greater than those found at allozyme loci; however, it does appear so (Table 4). Also, in the sister species brown trout (*Salmo trutta*), known to show higher levels of polymorphism than salmon populations at allozyme loci, the average heterozygosity (0.31) and the number of alleles per locus (4.6) for microsatellite loci are higher than observed at allozyme loci (0.06 and 1.55, respectively) (Presa *et al.*, 1994).

There are differences in both the amount and the distribution of genetic variation between allozymes and microsatellites (Table 5). The total absolute gene diversity for allozyme loci is approximately 42 per cent lower than that of microsatellite loci, and the relative genetic divergence between samples is approximatey 30 per cent lower for microsatellite loci (7.92 per cent vs. 10.43 per cent) (Table 5). However, the estimations by both allozymes and microsatellites of the relative genetic divergence between geographical areas (62.70 per cent and 67.30 per cent for allozymes and microsatellites, respectively) and among rivers within areas (37.29 per cent and 32.70 per cent) are similar (Table 5). Both allozyme and microsatellite variation showed the same pattern of differentiation between areas with the Irish and Spanish populations grouped in different clusters (Fig. 2). However, greater values for genetic distances between populations were found using the allelic frequencies of microsatellite loci (mean genetic distances 0.0449 ± 0.008 using

allozyme data and 0.0747 ± 0.011 for microsatellite data, Table 4), indicating that microsatellite loci can be more useful genetic markers for the identification of an individual's population. In fact, mean diagnostic value (estimated as the probability of assigning correctly an individual to one of two populations (see Ayala & Powell, 1972) was estimated as 0.593 ± 0.039 for the five allozyme loci (ranging between 0.53 and 0.78) and 0.687 ± 0.035 (ranging between 0.65 and 0.73) for the three microsatellite loci (Table 4). Certainly, these values are lower than the 0.99 probability in the Criterion I (Ayala & Powell, 1972) used to define a locus as diagnostic. However, the possibility of finding loci with higher values of diagnosis appears more feasible using microsatellites rather than allozymes. Although two protocols have been recently adapted to analyse new allozyme loci (Verspoor & Jordan, 1994; Wilson *et al.*, 1995), it is apparent that a greater number of new allozyme loci will not be available for analysis in the short-term. In contrast, the total number of (dG-dT) microsatellites in the salmon haploid genome has been estimated to be 6.3×10^4 (Slettan *et al.*, 1993) and other microsatellites with different repeat motifs or isolated from other related species may also be useful for salmon studies (Estoup *et al.*, 1993, Presa *et al.*, 1994).

The comparison of allozyme and microsatellite heterozygosity shows a positive, but statistically insignificant, correlation between the two sets of data ($r^2 = 0.011$, $P = 0.8814$). The Spanish populations which showed lower levels of heterozygosity at allozyme loci also displayed the lowest levels at microsatellite loci (Tables 2–4). However, comparisons between both levels of variation should be made with caution as different mechanisms are involved in the generation and maintenance of both classes of polymorphism. Microsatellite polymorphism is generated by length mutations mainly from DNA slippage and unequal crossing-over and these are distinct mechanisms from those that generate

Table 5 Hierarchical gene diversity analysis for Atlantic salmon (Wright, 1978)

	Microsatellite loci			Enzyme loci		
	Total variance	Relative (%)	Relative (%)	Total variance	Relative (%)	Relative (%)
Total	0.47997	100		0.27792	100	
Within samples	0.44195	92.08		0.24893	89.57	
Between samples	0.03802	7.92	100	0.02899	10.43	100
Between areas	0.02558		67.30	0.01818		62.70
Among rivers within area	0.01244		32.70	0.01081		37.29

point mutational change. In contrast, fast and slow alleles of alcohol dehydrogenase of *Drosophila melanogaster* differ by a single amino acid, Thr vs. Lys at codon 192, and only one of the polymorphisms found in the sequencing of this locus results in this change (Kreitman, 1983). Also of relevance is the fact that microsatellite polymorphism appears to be selectively neutral whereas for allozyme variation strict neutrality may not be true at all loci (*m-MEP-2** and *ₛAAT-4** in Atlantic salmon) and differences between populations may reflect special adaptation to specific environments.

In summary, these preliminary results show that a high level of polymorphism can be detected at microsatellite loci of Atlantic salmon. This microsatellite variation magnifies the differentiation of salmon populations previously observed from allozyme loci. The data highlight the potential of microsatellite single locus variation analysis for the study of Atlantic salmon population genetics.

Acknowledgements

References

AEBERSOLD, P. B., WINANS, G. A., TEEL, D. J., MILNER, G. B. AND UTTER, F. M. 1987. *Manual for Starch Gel Electrophoresis: a Method for the Detection of Genetic Variation.* Noaa Tech. Rep. NMFS. National Marine Fisheries Service. Seattle, WA.

AYALA, F. J. AND POWELL, J. R. 1972. Allozymes as diagnostic characters of sibling species of *Drosophila. Proc. Natl. Acad. Sci. U.S.A.*, **69**, 1094–1096.

BLANCO, G., SÁNCHEZ, J. A., VÁZQUEZ, E., RUBIO, J. AND UTTER, F. M. 1992. Genetic differentiation among natural European populations of Atlantic salmon, *Salmo salar* L., from drainages of the Atlantic Ocean. *Anim. Genet.*, **23**, 11–18.

BRITTEN, R. J. AND KOHNE, D. 1968. Repeated sequences in DNA. *Science*, **161**, 529–540.

CLAYTON, J. W. AND TRETIAK, D. N. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. *J. Fish. Res. Board Can.*, **29**, 1169–1172.

CROSS, T. F. AND KING, J. 1983. Genetic effects of hatchery rearing in Atlantic salmon. *Aquaculture*, **33**, 33–40.

CROSS, T. F. AND WARD, R. D. 1980. Protein variation and duplicate loci in the Atlantic salmon, *Salmo salar* L. *Genet. Res.*, **36**, 147–165.

CROZIER, W. W. AND MOFFET, I. J. J. 1989. Amount and distribution of biochemical and genetic variation among wild populations and hatchery stocks of Atlantic salmon, *(Salmo salar* L.), from north-east Ireland. *J. Fish Biol.*, **35**, 665–677.

DAVIDSON, W. S., BIRT, T. P. AND GREEN, J. M. 1989. A review of genetic variation in Atlantic salmon, *Salmo salar* L., and its importance for stock identification, enhancement programmes and aquaculture. *J. Fish Biol.*, **34**, 547–560.

ELO, K., VOURINEN, J. A. AND NIEMELÄ, E. 1994. Genetic resources of Atlantic salmon *(Salmo salar* L.) in Teno and Näätämö Rivers, northernmost Europe. *Hereditas*, **120**, 19–28.

ESTOUP, A., PRESA, P., KRIEG, F., VAIMAN, D. AND GUYOMARD, R. 1993. $(CT)_n$ and $(GT)_n$ microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity*, **71**, 488–496.

GALVIN, P., CROSS, T.F. AND FERGUSON, A. 1994. Genetic differentiation and gene flow in Atlantic salmon, *Salmo salar* L.: a case study of the River Shannon system in Ireland. *Aquacult. Fish. Manage.*, **25** (suppl. 2), 131–145.

HAMADA, H., PETRINO, M. G. AND KAKUNAGA, T. 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 6465–6469.

HEARNE, C. M., GHOSH, S. AND TODD, J. A. 1992. Microsatellites for linkage analysis of genetic traits. *Trends Genet.*, **8**, 288–294.

HOVEY, S. J., KING, D. P. F, THOMPSON, D. AND SCOTT, A. 1989. Mitochondrial DNA and allozyme analysis of Atlantic salmon, *Salmo salar* L., in England and Wales. *J. Fish Biol.*, **35**, 253–260.

HUGHES, C. R. AND QUELLER, D. C. 1993. Detection of highly polymorphic microsatellite loci in species with little allozyme polymorphism. *Mol. Ecol.*, **2**, 131–137.

JORDAN, W. C., YOUNGSON, A. F., HAY, D. W. AND FERGUSON, A. 1992. Genetic protein variation in natural populations of Atlantic salmon *(Salmo salar* L.) in Scotland: temporal and spatial variation. *Can. J. Fish. Aquat. Sci.*, **49**, 1863–1872.

KOJONEN, M.-L. 1989. Electrophoretically detectable genetic variation in natural and hatchery stocks of Atlantic salmon in Finland. *Hereditas*, **110**, 23–35.

KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster. Nature*, **304**, 412–417.

McELLIGOTT, E. A. AND CROSS, T. F. 1991. Protein variation in wild Atlantic salmon, with particular reference to southern Ireland. *J. Fish Biol.*, **39**, 35–42.

MOONEY, J., POWELL, E., CLABBY, C. AND POWELL, R. 1995. Detection of *Aeromonas salmonicida* in wild Atlantic salmon using a specific DNA probe test. *Dis. aquat.*

Org., **21**, 131–135.

NEI, M. 1972. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3321–3323.

O'CONNELL, M., SKIBINSKY, D. O. F. AND BEARDMORE, J. A. 1995. Mitochondrial DNA and allozyme variation in Atlantic salmon (*Salmo salar*) populations in Wales. *Can. J. Fish. Aquat. Sci.*, **52**, 171–178.

PAYNE, R. H., CHILD, A. R. AND FORREST, A. 1971. Geographical variation in the Atlantic salmon. *Nature*, **231**, 250–252.

PRESA, P., KRIEG, F., ESTOUP, A. AND GUYOMARD, R. 1994. Diversité et gestion génétique de la truite commune: apport de l'étude du polymorphisme des locus protéiques et microsatellites. *Génét. Sél. Évol.*, **26**, 183–202.

RIDGWAY, G. J., SHERBURNE, S. W. AND LEWIS, R. D. 1970. Polymorphism in esterases of Atlantic herring. *Trans. Am. Fish. Soc.*, **99**, 147–151.

SÁNCHEZ, J. A., BLANCO, G. AND VÁZQUEZ, E. 1993. Genetic status of Atlantic salmon in Asturias (northern Spain). In: Cloud, J. G. and Thorgaard, G. H. (eds) *Genetic Conservation of Salmonid Fishes*, pp. 219–225. Plenum Press, New York.

SÁNCHEZ, J. A., BLANCO, G., VÁZQUEZ, E., GARCÍA, E. AND RUBIO, J. 1991. Allozyme variation in natural populations of Atlantic salmon in Asturias (northern Spain). *Aquaculture*, **93**, 291–298.

SANGER, F., NICKLEN, S. AND COULSON, A. R. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 5463–5467.

SHAKLEE, J. B., ALLENDORF, F. W., MORIZOT, D. C. AND WHITT, G. S. 1990. Gene nomenclature for protein-coding loci in fish. *Trans. Am. Fish. Soc.*, **119**, 2–15.

SLETTAN, A., OLSAKER, I. AND LIE, O. 1993. Isolation and characterization of variable (GT)n repetitive sequences from Atlantic salmon, *Salmo salar* L. *Anim. Genet.*, **24**, 195–197.

SNEATH, P. H. A. AND SOKAL, R. R. 1973. *Numerical Taxonomy*. W.H. Freeman, San Francisco, CA.

SPRATT, B. G., HEDGE, P. J., HEESESEN, S., EDELMAN, A. AND BROOME-SMITH, J. K. 1986. Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8 and pEMBL9. *Gene*, **41**, 337–342.

STÅHL, G. 1987. Genetic population structure of Atlantic salmon. In: Ryman, N. and Utter, F., (eds) *Population Genetics and Fishery Management*, pp. 121–140. University of Washington Press, Seattle.

SWOFFORD, D. L. AND SELANDER, R. B. 1989. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *J. Hered.*, **72**, 282–302.

TAGGART, J. B. AND FERGUSON, A. 1990. Hypervariable minisatellite DNA single locus probes for the Atlantic salmon, *Salmo salar* L. *J. Fish Biol.*, **37**, 991–993.

VERSPOOR, E. 1988. Identification of stocks in the Atlantic salmon. In: Stroud, R. H. (ed.) *Proceedings of the Symposium on Future Atlantic Salmon Management*, pp. 37–46. Marine Recreational Fisheries Series, Savannah, GA.

VERSPOOR, E. AND JORDAN, W. C. 1989. Genetic variation at the *Me-2* locus in the Atlantic salmon within and between rivers: evidence for selective maintenance. *J. Fish Biol.*, **35**, 205–213.

VERSPOOR, E. AND JORDAN, W. C. 1994. Detection of an NAD^+-dependent malic enzyme locus in the Atlantic salmon, *Salmo salar*, and other salmonid fish. *Biochem. Genet.*, **32**, 105–117.

WEBER, J. L. 1990. Informativeness of human (dC-dA)n (dG-dT)n polymorphism. *Genomics*, **7**, 524–530.

WILSON, I. F., BOURKE, E. A. AND CROSS, T. F. 1995. A triose-phosphate isomerase polymorphism in the Atlantic salmon, *Salmo salar* L. *Biochem. Genet.*, **33**, 25–33.

WINTERO, A. K., FREDHOLM, M. AND THOMSEN, P. D. 1992. Variable (dG-dT)n, (dC-dA)n sequences in the porcine genome. *Genomics*, **12**, 281–288.

WONG, A. K. C., YEE, H. A., VAN DE SANDE, J. H. AND RATTNER, J. B. 1990. Distribution of CT-rich tracts is conserved in vertebrate chromosome. *Chromosoma*, **99**, 344–351.

WRIGHT, S. 1978. *Evolution and the Genetics of Populations*, vol. 4, *Variability Within and Among Natural Populations*. University of Chicago, Chicago.