Journal of Heredity 2007:98(7):692–704
doi:10.1093/jhered/esm093
Advance Access publication November 5, 2007

A Comparison of Biallelic Markers and Microsatellites for the Estimation of Population and Conservation Genetic Parameters in Atlantic Salmon (*Salmo salar*)

Heikki J. Ryynänen, Anni Tonteri, Anti Vasemägi, and Craig R. Primmer

From the Department of Biological and Environmental Sciences, University of Helsinki, PO Box 65, FIN-00014 University of Helsinki, Finland (Ryynänen); and the Department of Biology, University of Turku, FIN-20014, Finland (Tonteri, Vasemägi, and Primmer).

Address correspondence to H. J. Ryynänen at the address above, or e-mail: heikki.j.ryynanen@helsinki.fi.

Abstract

Biallelic markers such as single nucleotide polymorphisms (SNPs) and insertion/deletion polymorphisms have become increasingly popular markers for various population genetics applications. However, the effort required to develop biallelic markers in nonmodel organisms is still substantial. In this study, we compared the estimation of various population genetic parameters (genetic divergence and structuring, isolation-by-distance, genetic diversity) using a limited number of biallelic markers (in total 7 loci) to those estimated with 14 microsatellite loci in 21 Atlantic salmon (*Salmo salar*) populations from northern Europe. Pairwise F_{ST} values were significantly correlated between biallelic loci and microsatellite datasets, as was overall heterozygosity when both anadromous and nonanadromous populations were analyzed together. However, when the anadromous and nonanadromous samples were analyzed separately, only genetic divergence correlations remained significant. Biallelic markers alone were not sufficient for reliable neighbor-joining clustering of populations but gave highly similar isolation-by-distance signals when compared with microsatellites. Finally, although several population prioritization measures for conservation exhibited significant correlation between different marker types, the specific populations highlighted as being most valuable for conservation purposes varied depending on the marker type and conservation criteria applied. This study demonstrates that a relatively small set of biallelic markers can be sufficient for obtaining concordant results in most of the analyses compared with microsatellites, although estimates of genetic distance are generally more concordant than estimates of genetic diversity. This suggests that a relatively small number of biallelic markers can provide useful information for various population genetic applications. However, we emphasize that the use of much higher number of loci is preferable, especially when the genetic differences between populations are subtle or individual multilocus genotype-based analyses are to be performed.

Single nucleotide polymorphisms (SNPs) together with insertion/deletion polymorphisms (indels) represent an important component of the genetic variation found in the genomes of vertebrates, and in recent years particularly SNPs have been claimed to become the 'new' genetic marker of choice in a variety of population genetic applications for several reasons (see Brumfield et al. 2003; Morin et al. 2004 for reviews). First, they occur at high frequencies in a wide range of species (Britten et al. 2003; Brumfield et al. 2003). Second, the mutation mechanism of SNPs is simpler and better understood compared with microsatellites (Vignal et al. 2002; Brumfield et al. 2003). Third, their genotyping error rates are relatively low (Ranade et al. 2001) and high-throughput analysis procedures enable large-scale population level analyses using very large number of markers (Kwok 2001; Syvänen 2001; Vignal et al. 2002; Chen and Sullivan 2003). Fourth, often the developed SNPs are linked to the coding regions of the genome, and it is assumed that such gene-linked loci can be more efficiently used to identify functionally important polymorphisms than

application of random genetic markers (van Tienderen et al. 2002; Luikart et al. 2003).

The potential advantages of biallelic markers are counterbalanced by the fact that a substantial amount of screening effort is required for their development and a considerably larger number of loci need to be assessed due to the rather low amount of information per marker (each locus normally possesses only two allelic variants) (Glaubitz et al. 2003). For example, it has been estimated that for accurate parentage determination in natural populations, up to 100 SNPs are required (Anderson and Garza 2006), whereas with highly polymorphic microsatellite loci, as few as 3 loci have proved to be sufficient for accurate parentage assignment (e.g., Saino et al. 1997). This limitation is especially important when using biallelic markers in nonmodel organisms where no large-scale sequencing projects have been initiated, and thus the polymorphism discovery step needs to be made through laboratory screening rather than using in silico searches (Morin et al. 2004). Another critical issue, when using biallelic markers such as SNPs in population genetic analyses of nonmodel organisms, is the choice of a suitable genotyping platform as many high-throughput methods require expensive and specialized equipment not commonly found in genetic laboratories working with nonmodel organisms. This fact was highlighted in a recent review by Schlötterer (2004), which provides an historical perspective on the use of different molecular marker types and suggests that when evaluating the feasibility of SNPs in nonmodel organisms it is essential to also assess the investment in marker development and genotyping costs. Furthermore, it is widely acknowledged that if SNPs have been identified from a small panel of individuals, rare SNPs will be under-represented, which is known as ascertainment bias (e.g., Nielsen 2000). In spite of these potential limitations, biallelic markers are expected to represent useful tools for various applications in molecular ecology and conservation genetics (Morin et al. 2004). Growing EST databases and novel SNP detection methods (Comai et al. 2004; Orsini et al. 2007) are expected to further increase the efficiency and speed of SNP discovery. In addition, from a conservation perspective, SNPs may have some technical advantages over microsatellites as they can be amplified from highly degraded DNA (Budowle 2004), with some of the recent SNP genotyping methods allowing polymorphism detection with fragment sizes of only tens of base pairs (see Chen and Sullivan 2003 for a review).

Recently, several studies have focused on comparing the advantages and disadvantages of microsatellites and SNP markers for various genetic analysis strategies in humans (see Bailey-Wilson et al. 2005). For example, based on simulations, Glaubitz et al. (2003) reported that at least 5 times more SNPs than microsatellites are needed to reliably determine genetic relationships between human populations, and for linkage studies the amount needed has been reported to be approximately 3-fold (Kruglyak 1997). The use of various molecular markers in animal genetics has recently been reviewed by Vignal et al. (2002) but thus far only few empirical studies have compared the utility of microsatellites and SNPs for population genetic studies of nonmodel organisms. For example, recent studies with flycatchers and gray wolves applied SNPs to address ecological, evolutionary, or conservation issues (Saetre et al. 2003; Seddon et al. 2005), whereas Rengmark et al. (2006) evaluated the efficiency of SNPs in terms of genetic assignment and parentage testing in Atlantic salmon.

The identification of a large number of novel and unambiguous SNPs or indels in salmonid fishes is particularly challenging due to the duplicated nature of their genome (Allendorf and Thorgaard 1984; Smith et al. 2005; Ryynänen and Primmer 2006; Hayes et al. 2007). Recently, we identified SNPs and indels in a number of Atlantic salmon genes (Ryynänen and Primmer 2004a, 2006). In this study, 9 of these polymorphisms were chosen to analyze 667 individuals originating from 21 northern European populations that had earlier been studied using 14 microsatellite markers (Tonteri et al. 2005). This enabled a direct comparison of the gene-linked biallelic marker data with the previous results obtained with presumably neutral microsatellite markers in respect of an estimation of various population genetic parameters as well as prioritization of particular populations for conservation purposes. Furthermore, possible deviations from neutral expectations were assessed in both datasets to identify putative outlier loci potentially affected by natural selection.

Materials and Methods

Samples and Loci Assessed

A total of 667 individuals from 21 populations in northern Europe (Figure 1, Table 1) were analyzed in this study. These samples were a subset of those included also in the phylogeography study of Tonteri et al. (2005), which enabled a direct comparison of two datasets consisting of different types of markers. The sampling and DNA extraction methodologies as well as the analysis methods for the 14 microsatellite loci are described in Tonteri et al. (2005).

A total of 9 gene-associated biallelic polymorphisms were assessed. These included 6 of the SNPs or indels reported in Ryynänen and Primmer (2006) (Tables 2 and 3), 1 nonsynonymous sequence polymorphism in the *MHC II* gene (Table 2), and 2 indels identified in the *GH1* gene of Atlantic salmon (Ryynänen and Primmer 2004a) (Table 3). The polymorphic loci were initially identified by screening 15 populations for SNP/indel markers and 9 populations for *GH1* indels, covering a wide range of the species' distribution using 1 individual per population (Ryynänen and Primmer 2004a, 2006). Furthermore, in attempts to minimize ascertainment bias in the subsequent analyses, selection of the loci was based on their suitability for secondary screening methods rather than on their polymorphism levels within or between populations.

The SNPs were genotyped using restriction fragment length polymorphism analysis of the PCR products characterized by the presence/absence of a diagnostic

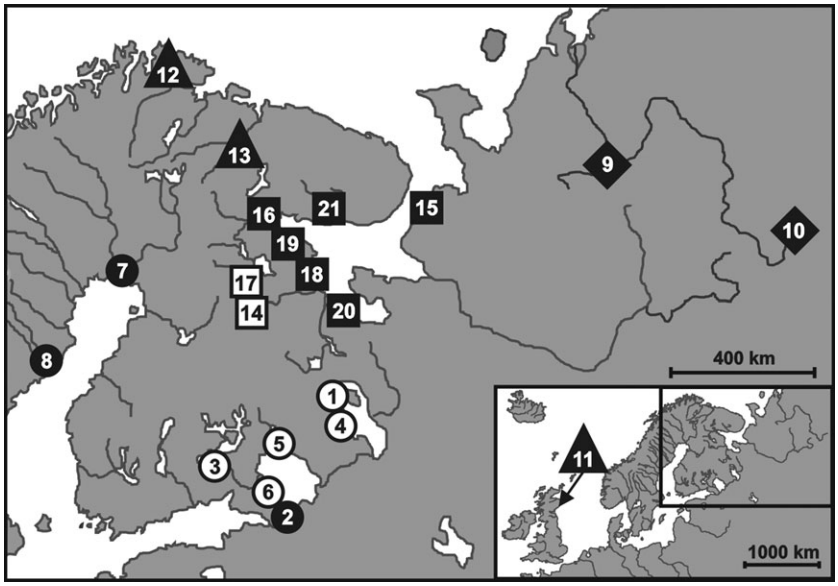

Figure 1. Sample locations of Atlantic salmon populations analyzed in this study. Circles, squares, diamonds, and triangles represent the Baltic Sea, White Sea, Barents Sea, and the Atlantic Ocean, respectively. Furthermore, black and white symbols indicate anadromous and nonanadromous populations, respectively. See Table 1 for names and detailed locations of the populations.

restriction endonuclease recognition site (Table 2). The PCR reaction mixtures (20 μl total) contained 50–100 ng of template DNA, 1× PCR buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 1.5 mM MgCl2, 0.25 mM of each dNTP (Finnzymes, Espoo, Finland), 0.5 U of *Taq* polymerase (BioTaq, BIOLINE, London, UK) and 0.5 μM of each primer (Table 2). The PCR amplifications were performed in PTC-100 or PTC-200 Thermocyclers (MJ Research, Watertown, MA), with a cycle consisting of an initial denaturation step at 94°C for 2 min followed by 35 cycles of 95°C for 30 s, annealing at one of various temperatures (see Table 2) for 30 s, 72°C for 45 s, followed by a final extension at 72°C for 5 min. The amplified PCR fragments were then digested for 1.5 h in a 20 μl total volume using 2–5 μl of PCR product, 2 U of specific restriction enzyme (Promega, Madison, WI; or MBI Fermentas GmbH, St. Leon-Rot, Germany) at the recommended incubation temperature. The digested PCR products were visualized using 1.2% agarose gel electrophoresis followed by ethidium bromide staining, and genotypes were determined by eye. The indels were analyzed as outlined in Ryynänen and Primmer (2004b), except for using the primers introduced in Table 3. Only individuals that were genotyped with at least 6 of the 9 SNPs/indels and 12 out of 14 microsatellites (Tonteri et al. 2005) were included in further analyses (see Table 1).

Population Genetic Analyses

As expected, the 2 *GH1* polymorphisms (SsGH1_indel01 and SsGH1_indel03, Table 3) exhibited highly significant genotypic disequilibrium ($P < 0.0001$) calculated using GENEPOP v.3.4 software (Raymond and Rousset 1995). Therefore, the program PHASE 2.0.2 (Stephens et al. 2001; Stephens and Donnelly 2003) was used to infer haplotypes for the 2 *GH1* indels, and the inferred haplotypes were subsequently treated as a single locus. Due to the a priori expectation that the MHC class II beta chain in *Salmon salar* may be affected by selective forces (see Landry and Bernatchez 2001; Langefors et al. 2001; Lohm et al. 2002; de Eyto et al. 2007), and because the analyzed polymorphic site in *MHC* gene was a nonsynonymous substitution, this locus was excluded from all analyses assuming neutrality but was included in the analysis aimed at detecting the potential effects of selection (see below). Therefore, the subsequent analyses included either 6 or 7 biallelic loci plus the *GH1* haplotype, depending on the analysis applied. Hereafter, this set of markers consisting both SNPs and indels is referred to as "biallelic loci".

The observed and expected heterozygosities for biallelic markers and microsatellite loci were estimated using a Microsatellite toolkit version 3.1 (Park 2001). To evaluate the similarity of genetic diversity estimates measured as expected heterozygosity and allelic richness between 2 marker types, Pearson correlation coefficient was calculated. Tests for significant deviations from Hardy–Weinberg equilibrium were conducted using the GENEPOP v3.4 software, and significance levels were corrected for multiple comparisons by employing a Bonferroni correction (Rice 1989). In addition, a genetic differentiation test based on allele frequency distribution was performed using GENEPOP v3.4. The permutation test implemented in FSTAT v2.9.3 (Goudet 2001) was applied to test for the significance of differences in genetic diversity between anadromous and nonanadromous populations. The FSTAT program was also

Table 1. Names, origins, and genetic diversity indices for the Atlantic salmon populations included in this study

						SNPs diversity			Microsatellite diversity[a]		
No.	Population	Abbreviation	Basin	Latitude/longitude	*N*	Rs	H_o	H_e	Rs	H_o	H_e
1	Lizhma	Liz	Baltic Sea[b]	62°25′N 34°27′E	26	1.30	0.04	0.08	2.71	0.54	0.51
2	Neva	Nev	Baltic Sea	59°58′N 30°13′E	42	1.51	0.18	0.18	3.62	0.63	0.61
3	Saimaa	Sai	Baltic Sea[b]	63°19′N 30°01′E	44	1.35	0.11	0.10	1.99	0.29	0.28
4	Shuja	Shu	Baltic Sea[b]	61°51′N 34°09′E	17	1.26	0.08	0.07	3.12	0.57	0.55
5	Sysky	Sys	Baltic Sea[b]	61°39′N 31°16′E	10	1.40	0.14	0.12	2.83	0.47	0.47
6	Taipale	Tai	Baltic Sea[b]	60°38′N 30°30′E	36	1.47	0.23	0.19	3.52	0.60	0.57
7	Tornio	Tor	Baltic Sea	65°49′N 24°08′E	36	1.53	0.15	0.15	3.72	0.60	0.60
8	Vindelälven	Vin	Baltic Sea	63°50′N 20°05′E	44	1.42	0.10	0.12	2.94	0.47	0.48
9	Pizhma	Piz	Barents Sea	64°53′N 51°17′E	18	1.96	0.31	0.33	3.53	0.59	0.60
10	Unja	Unj	Barents Sea	61°32′N 58°15′E	11	1.70	0.36	0.29	3.26	0.62	0.54
11	Dee	Dee	The Atlantic Ocean	56°54′N 3°27′W	48	1.71	0.27	0.26	5.02	0.72	0.73
12	Teno	Ten	The Atlantic Ocean	70°30′N 28°25′E	42	1.99	0.38	0.37	4.60	0.66	0.70
13	Tuloma	Tul	The Atlantic Ocean	68°41′N 31°55′E	39	1.96	0.28	0.28	4.50	0.67	0.72
14	Kamennoe	Kam	White Sea[b]	64°28′N 30°26′E	40	1.40	0.14	0.14	2.15	0.37	0.35
15	Megra	Meg	White Sea	66°03′N 41°43′E	48	2.00	0.38	0.34	4.17	0.68	0.66
16	Nilma	Nil	White Sea	66°27′N 33°05′E	11	1.83	0.32	0.27	2.99	0.57	0.55
17	Pisto	Pis	White Sea[b]	65°16′N 30°35′E	24	1.49	0.17	0.16	2.55	0.44	0.42
18	Pongoma	Pon	White Sea	65°18′N 34°02′E	37	1.96	0.27	0.30	3.91	0.66	0.67
19	Pulonga	Pul	White Sea	66°18′N 33°17′E	19	1.85	0.39	0.34	2.92	0.52	0.53
20	Suma	Sum	White Sea	64°14′N 35°25′E	33	1.56	0.20	0.21	3.36	0.60	0.58
21	Varzuga	Var	White Sea	66°36′N 36°35′E	42	1.95	0.32	0.33	4.03	0.61	0.65

Populations from Neva and Saimaa were of hatchery origin.

N, number of individuals; Rs, average allelic richness in a population; H_o, observed heterozygosity; H_e, expected heterozygosity.

[a] Reanalyzed data from Tonteri et al. (2005).

[b] A nonanadromous migration behavior.

used to estimate allelic richness within populations and pairwise F_{ST} between populations separately for biallelic and microsatellite loci. The standardized pairwise G'_{ST} estimates (Hedrick 2005) over biallelic and microsatellite loci were calculated by dividing the original pairwise F_{ST} values by the maximum pairwise F_{ST} values estimated with the program RECODEDATA (Meirmans 2006). A Mantel test for matrix correspondence was used to test for correlation of the pairwise F_{ST} and G'_{ST} estimates calculated between 2 different marker types as implemented in the GENALEX 6 program (Peakall and Smouse 2005). The significance level was based on 999 random permutations. Isolation-by-distance (IBD) was assessed for all anadromous populations based on biallelic loci and microsatellite data separately using a Mantel's test in the GENALEX 6 program as described above. The genetic distances were calculated as

Table 2. Details of the polymorphic loci used for restriction fragment length polymorphisms of Atlantic salmon populations in this study

Locus (GenBank access no.)	SNP site and substitution[a]	Primers (5′–3′)	Annealing Tm (°C)	Restriction enzyme (recognition site)	Allele	Fragment sizes (bp)	Allele frequency[b]
FGF6_snp02	510	F: CAACCTATTTTTACACTGGCTCCT	45	PacI	G	593	91.05
(DQ834854)	G/T	R: ACCAAATAGCCTACCATTCCATTA		(ttaat/taa)	T	534 59	8.95
IgM_snp01	423	F: ACCTTAGGGCAAATTAAACAATCA	42	SmlI	T	538	93.49
(DQ834865)	T/C	R: CAAAAACCAGAGATTGCAAAGTTA		(c/tyrag)	C	435 103	6.51
IL-1 beta2_snp01	85	F: ATAATTGTCTGTTGAGGCTGGAGT	56	BsmI	T	358 68	54.80
(DQ834861)	T/C	R: TGGAGAGAGAGAGAGGGAGAGATA		(gaatg/c)	C	318 68 40	45.20
MHC II_snp02	62[c]	F: GATCTGTATTATGTTTTCCTTCCAG	60	RsaI	T	298	87.11
(DQ863499)	T/C	R: CACCTGTCTTGTCCAGTATG		(gt/ac)	C	215 83	12.89

Details of these polymorphisms can be found in the GenBank and dbSNP databases (accession numbers provided in parentheses).

[a] Location of polymorphic site in GenBank accession.

[b] Average allele frequency over all populations.

[c] A nonsynonymous substitution.

Table 3. Details of the polymorphic loci used for length polymorphism screening of Atlantic salmon populations in this study

Locus (GenBank or dbSNP access no.)	Polymorphic site (size of indel in bp)	Primers (5′–3′)[a]	Fragment sizes (bp)	Allele frequency[b]
rps24_Indel01	661[c]	F: ATGAAGACCATCTTAGGTCTGAGC	121	31.39
(DQ834866)	(6)	R: GTTTGGGACAGGATAGATAATTGG	115	68.61
EF1a_Indel01	597[c]	F: TACCTTATACAAATGGCCGTAATG	151	91.30
(DQ834871)	(11)	R: GTTTCCATTTAACTACTCAGCCACAC	140	8.70
SsGH1_Indel01[d]	147	F: CCATAGGACATTCAATTTGACAAT	217	55.38
(ss24735153)	(12)	R: GTTTGAATTAGGGTCAAAACACAACACA	205	44.62
TGF-beta_Indel01	347[e]	F: CGATCAATTCATGTCTCTTAAATG	251	10.14
(DQ834857)	(9)	R: GTTTCGGAAGTCAATGTAAAGTCTTCG	242	89.86
SsGH1_Indel03[d]	4212	F: CAAGCTGTACAATACAACGCAAC	265	19.90
(ss24735155)	(2)	R: GTTTTGGTTGGAGTTTCTGACCATTAG	263	80.10

Details of these polymorphisms can be found in the GenBank and dbSNP databases (accession numbers provided in parentheses).

[a] FAM-label was incorporated in the forward primers. A-GTTT 'tail' was designed for the 5′ end of each reverse primer (added bases underlined) to reduce stuttering.

[b] Average allele frequency over all populations.

[c] Location of polymorphic site as it appears in GenBank.

[d] Taken from Ryynänen and Primmer (2004b).

[e] A longer insertion (9 bp) was detected during the population screening step and, for practical purposes, this was chosen instead of the original 1-bp deletion reported (Ryynänen and Primmer 2006).

$F_{ST}/(1 - F_{ST})$ (Rousset 1997), and interpopulation geographical distances were measured as the shortest sea distance between river mouths.

Phylogenetic Analysis

Neighbor-Joining trees were constructed for the analyzed datasets using the SEQBOOT, GENDIST, NEIGHBOR, CONSENSE, RETREE, and FITCH programs in the PHYLIP version 3.6 software package (Felsenstein 1995). Phylograms were created based on chord distances (D_{CE}) (Cavalli-Sforza and Edwards 1967) for microsatellite and biallelic marker data separately and also for both datasets combined. The reliability of phylograms was estimated by bootstrapping 2000 replicates over loci and the extended majority rule consensus trees were inferred.

Estimation of the Conservation Value of Different Populations

The conservation value of different populations (or groups of populations), which can be defined as the relative contribution of a particular population (or group of populations) to the overall genetic diversity or divergence, was assessed using 2 different approaches. The first method assesses the conservation priorities based on phylogenetic approach as implemented in the CONSERVE v3.2 program (available from http://www.agapow.net/software/conserve/) (Crozier et al. 1999). Shortly, the conservation value GD ("genetic diversity") of 7 different population groupings, based on the evolutionary lineages proposed by Tonteri et al. (2005), was estimated by calculating the proportion of the overall diversity retained in the various population groupings (see Crozier and Kusmierski 1994 for further details). In total, 100 bootstrap trees for both biallelic marker and microsatellite datasets were constructed using the SEQBOOT, GENDIST, and NEIGHBOR programs of the PHYLIP version 3.6 software package (Felsenstein 1995) as described above, and the resulting multiple trees acted as an input file for CONSERVE.

The second method applied to assess the conservation value of a particular population was that proposed by Petit et al. (1998), which estimates the contribution of each population to the total diversity based on the population's own genetic diversity (measured by expected heterozygosity) and its relative differentiation from the remaining populations. This was estimated using the CONTRIB version 1.01 software package (http://www.pierroton.inra.fr/genetics/labo/Software/Contrib/) (Petit et al. 1998), treating biallelic marker and microsatellite data separately. In order to evaluate the congruence of the different conservation parameters estimated for the biallelic and microsatellite loci, the Pearson correlation coefficient between 2 marker types was calculated.

Multilocus Test for Neutrality

A coalescent simulation-based method developed by Beaumont and Nichols (1996) was applied to test for potential signatures of selection acting on the biallelic and microsatellite loci. The method is based on a symmetrical island model of population structure and generates datasets with the mean F_{ST} similar to the empirical distribution. The observed F_{ST} values of each locus are then compared with the distribution of simulated F_{ST} values (based on 20 000 simulated loci) to identify putative outliers deviating from the neutral expectations as implemented in the FDIST2 software (http://www.rubic.rdg.ac.uk/~mab/software.html). As some of the sampled populations in this study (e.g.

nonanadromous populations) were likely to violate the assumption of equal migration rates, analyses were conducted only for the "Atlantic clade" excluding Baltic and nonanadromous populations. Thus, the number of populations was set to 11 in all simulations, and sample sizes were set to 50 (n = 25 individuals) as recommended by Beaumont and Nichols (1996) if the median sample size is >50. The simulations were made assuming an infinite allele mutation model for biallelic marker data and a stepwise mutation model for microsatellite data.

Results

Genetic Diversity, Hardy–Weinberg, and Linkage Equilibrium

The average observed heterozygosity for biallelic loci ranged between 0.04 (Lizhma) to 0.39 (Pulonga), whereas the estimates for microsatellite data were generally higher ranging from 0.29 (Saimaa) to 0.72 (Dee) (Table 1). Overall, the nonanadromous populations exhibited significantly lower levels of diversity than anadromous populations for both biallelic loci (0.13 vs. 0.26; permutation test $P = 0.006$) and microsatellites (0.43 vs. 0.63; permutation test $P = 0.004$).

No deviations from the Hardy–Weinberg equilibrium were observed in any of the populations or loci after the correction for multiple tests. One pair of microsatellites (Ssa197&SLEEN82) and 1 biallelic locus–microsatellite pair (rps24_Indel01&SSOSL483) remained in significant ($P < 0.05$) linkage disequilibrium (LDE) after a Bonferroni correction ($\alpha = 0.00024$). The effects of these linkages are, however, assumed to be minor as significant LDE of these locus combinations only originated from 2 populations instead of being a strong signal for physical linkage in all populations.

A significant correlation between the estimated heterozygosities (H_e) of biallelic markers and microsatellites in the analyzed populations was observed (Pearson's $r = 0.647$, $P = 0.002$) (Figure 2a). This correlation appeared to be largely driven by the general difference in genetic variability between anadromous and nonanadromous populations (Figure 2a, Table 1) as, when analyzed separately, neither of the associations remained significant (anadromous populations: $n = 14$; $r = 0.434$, $P = 0.12$; nonanadromous populations: $n = 7$; $r = 0.03$, $P = 0.95$). Similarly, a significant, positive correlation ($r = 0.438$, $P = 0.001$) between allelic richness estimates for different marker types was observed when all populations were analyzed together (Figure 2b), but again the correlation was not significant when the populations were split based on their life-history strategy.

Genetic Differentiation and IBD

All pairwise comparisons showed significant differentiation for microsatellite markers, whereas 10 out of a total 210 comparisons were nonsignificant for biallelic loci after the Bonferroni correction (Appendix 1). Pairwise F_{ST} estimates

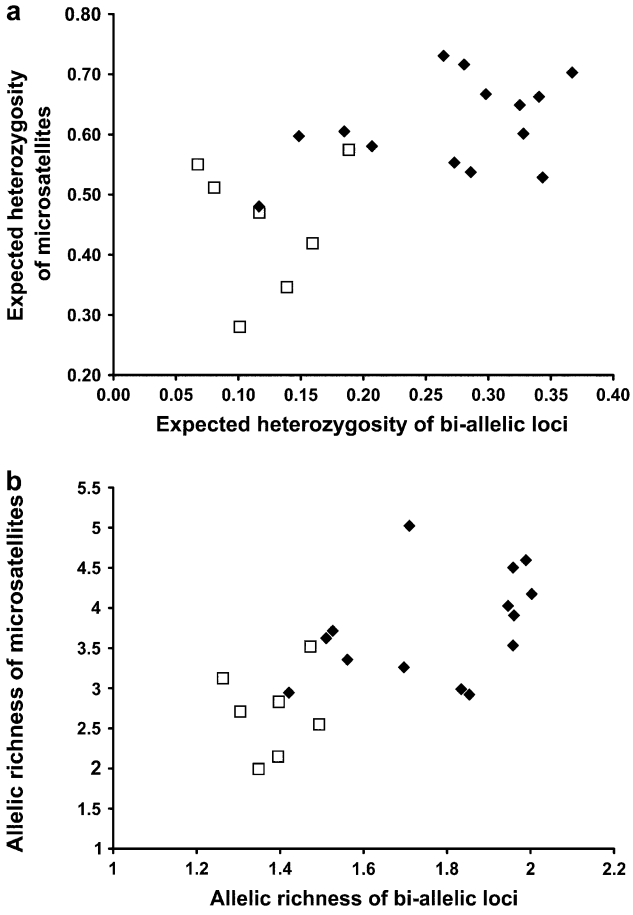

Figure 2. Positive association between seven biallelic loci and 14 microsatellites in (**a**) heterozygosity (*He*) (Pearson's $r = 0.647$, $P = 0.002$) and (**b**) allelic richness ($r = 0.662$, $P = 0.001$) estimates from 21 Atlantic salmon populations. Filled diamonds and open squares represent the anadromous and nonanadromous populations, respectively.

calculated for biallelic and microsatellite markers showed a highly significant, positive correlation (Mantel's $r_{xy} = 0.652$, $P = 0.001$, $r^2 = 0.425$) (Figure 3, Appendix 1). Significant correlations were also observed when the populations were split based on their life-history strategy (anadromous populations: $r_{xy} = 0.509$, $P = 0.002$, $r^2 = 0.259$; nonanadromous populations: $r_{xy} = 0.833$, $P = 0.001$, $r^2 = 0.694$; Figure 3). As expected, consistently higher values of F_{ST} were observed for the biallelic marker data than for the microsatellite loci (e.g., Hedrick 2005). For example, the highest pairwise F_{ST} estimates for microsatellite loci were close to 0.5 (average 0.22), whereas for biallelic loci F_{ST} sometimes exceeded 0.7(average 0.30) (Figure 3). Maximum G'_{ST} estimates exceeded 0.7 for the both marker types, whereas the average G'_{ST} was 0.37 and 0.49 for biallelic and microsatellite loci, respectively. Similarly to pairwise F_{ST} estimates, standardized G'_{ST} estimates were highly correlated between the 2 different

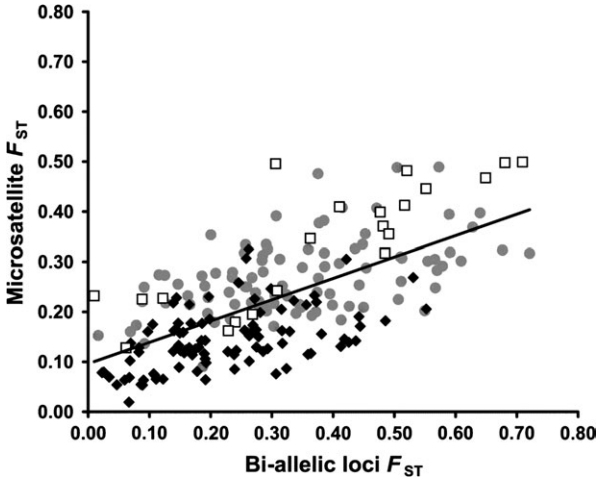

Figure 3. Pairwise F_{ST} estimates based on 7 biallelic loci versus 14 microsatellites (Mantel's $r_{xy} = 0.652$, $P = 0.001$, $r^2 = 0.425$). Filled diamonds, open squares, and gray circles represent pairwise population comparisons among anadromous, nonanadromous, and between anadromous and nonanadromous populations, respectively.

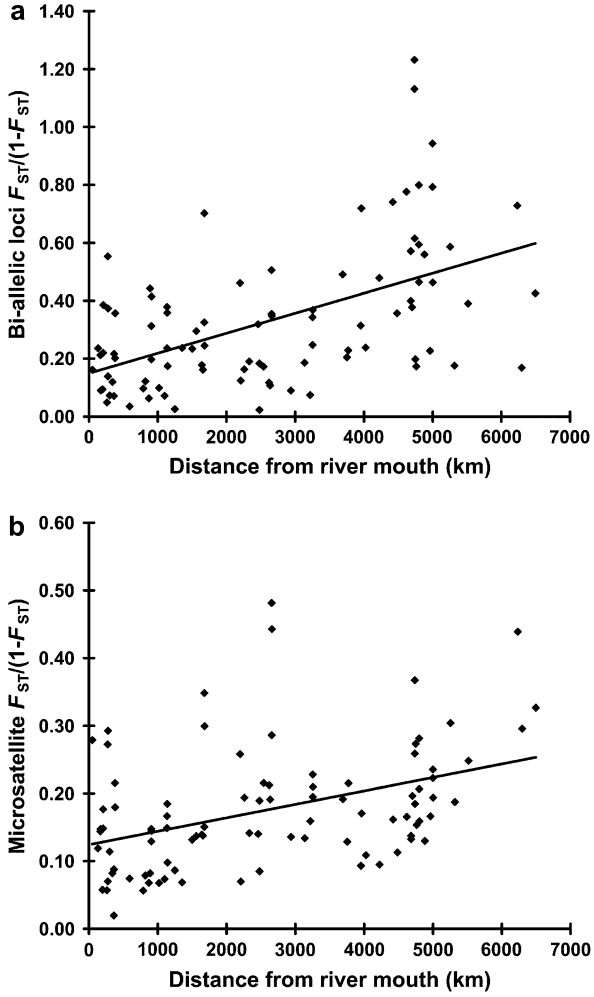

Figure 4. The observed IBD pattern for anadromous Atlantic salmon populations based on (**a**) seven biallelic loci (Mantel's $r_{xy} = 0.527$, $P = 0.004$, $r^2 = 0.278$) and (**b**) 14 microsatellites ($r_{xy} = 0.410$, $P = 0.002$, $r^2 = 0.168$). Note the variable scale of the *y*-axis between the 2 graphs.

marker types ($r_{xy} = 0.543$, $P = 0.001$, $r^2 = 0.294$), and thus only F_{ST} values were used subsequently. An association between genetic divergence (F_{ST}) and geographical distances revealed a strikingly similar IBD signal based on both biallelic marker and microsatellite data (Figure 4a and b). A significant correlation ($r_{xy} = 0.527$, $P = 0.004$, $r^2 = 0.278$) was observed between the estimated genetic and geographic distances for biallelic marker data when all anadromous populations were analyzed (Figure 4a), and the results were highly similar for microsatellite data ($r_{xy} = 0.410$, $P = 0.002$, $r^2 = 0.168$; Figure 4b).

Phylogenetic Relationships Between Populations

The same main population groupings were generally identified using either biallelic or microsatellite markers (Figure 5a and b). Based on biallelic marker data alone, however, the bootstrap support for different groupings was relatively poor with an average bootstrap support 34% per node and only 4 nodes supported by bootstrap values higher than 50% (Figure 5a). In contrast, the average bootstrap support for the microsatellite data was considerably higher (55%) with 9 nodes exhibiting bootstrap values higher than 50% (Figure 5b). Combining both biallelic marker data and microsatellite data only marginally improved the overall resolution with the average bootstrap value per node of 59% (Figure 5c). However, it did result in some changes to the structure and branch lengths of the tree. For example, the bootstrap support for the grouping of the Saimaa population with the nonanadromous White Sea populations Kamennoe and Pisto increased from 45% to 53%, and for the node grouping of all Baltic Sea basin populations together, the inclusion of biallelic loci increased the bootstrap support from 31% to 51% and improved the basal branch structure within the group (Figure 5c).

Population Prioritization for Conservation

Population prioritization estimates for 7 different geographical groupings using the genetic distance–based method implemented in CONSERVE program showed highly correlated estimates (Pearson's $r = 0.964$, $P = 0.0005$) between the biallelic marker and microsatellite datasets (Figure 6). Based on biallelic marker data, the White Sea, Barents Sea, and Atlantic Ocean grouping retained the highest proportion of the total variation (85%), whereas the proportion was 68% for microsatellites (Figure 6). The lowest estimates (ca. 13% with both marker types) were obtained for two nonanadromous populations (Kamennoe and Pisto): interestingly, adding a third anadromous population,

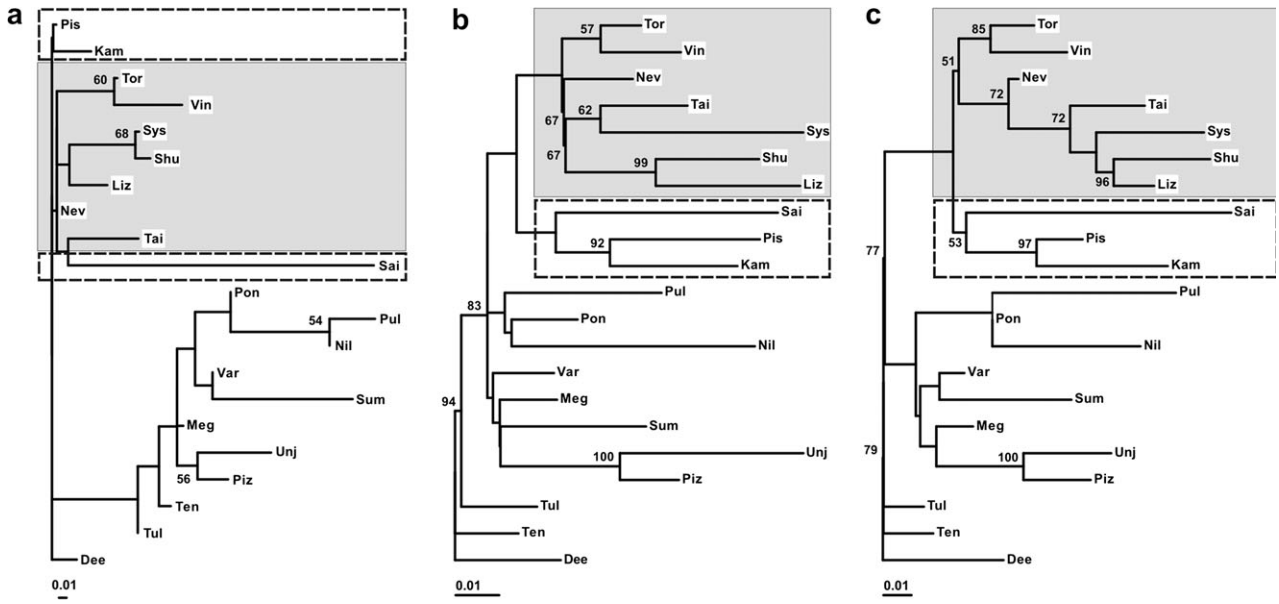

Figure 5. D_{CE} neighbor-joining phylograms for (**a**) biallelic marker data, (**b**) microsatellite data, or (**c**) the combined microsatellite and biallelic loci datasets. Numbers next to branches indicate bootstrap support over 2000 replicates; only values over 50% are shown. The gray area indicates Baltic Sea basin, the dashed framed area contains nonanadromous White Sea and Saimaa populations. The remaining populations originate from the White and Barents Seas or the Atlantic Ocean (see Table 1). The lengths of some of the shortest branches have been slightly increased for presentation purposes.

Saimaa, increased the conservation value more than 3-fold (to 48%) in biallelic markers and 2-fold (to 27%) in microsatellites (Figure 6).

Similarly, the second population prioritization method for conservation implemented in CONTRIB program showed a congruent pattern between biallelic marker and microsatellite datasets, both when conservation contribution estimates were based on genetic diversity (Pearson's $r = 0.422$, $P = 0.001$; Figure 7a) or population differentiation ($r = 0.477$; $P = 0.03$; Figure 7b). Interestingly, however, when the populations' overall contributions to the total diversity were assessed (sum of the diversity and divergence contributions), there was no correlation ($r = 0.01$, $P = 0.98$) between the 2 marker types (Figure 7c).

Tests for Selection

The majority of the biallelic loci fell within the 95% confidence limits of the simulated neutral distribution (Figure 8), indicating that it is not necessary to invoke a selective hypothesis to explain the observed levels of genetic differentiation. Interestingly, the *MHC* marker that was selected a priori as a potential outlier candidate did not deviate from the simulated neutral expectations (Figure 8). However, 1 biallelic locus *(EF1a)* out of 8 deviated ($P = 0.008$) from the simulated neutral distribution. Of the 14 microsatellites, 3 loci, Ssa412 ($P = 0.001$), Ssa422 ($P = 0.01$), and SLEE184 ($P = 0.009$) were pointed out as putative outliers falling outside the 95% confidence interval (Figure 8).

Discussion

In this study, we utilized a limited number of biallelic loci and compared their efficiency and resolution with 14 microsatellite markers for estimating various population genetic parameters in Atlantic salmon. Interestingly, the 6 biallelic loci together with 1 *GH1* haplotype (a total of 16 alleles) used here performed relatively similarly in most of the analyses when compared with a dataset consisting of 14 microsatellite (in total of 237 alleles) markers (Tonteri et al. 2005).

Heterozygosity estimates were positively correlated between the 2 types of markers (Figure 2). A weak correlation between SNP and microsatellite markers was also observed earlier in a population level study of Atlantic salmon (Rengmark et al. 2006) but not at the individual level in wolves (Seddon et al. 2005). The observed correlation between the markers indicates that even a rather limited number of biallelic loci can provide useful information about the general level of genetic diversity in salmon populations assessed. However, when the populations were split based on their life-history strategy, the correlation between 2 marker types was nonsignificant (Figure 2). This may be due small number of populations analysed but might also suggests that while major differences in genetic variability can be easily detected using both marker types, more subtle differences in diversity might be more difficult to detect using only a small number of biallelic markers.

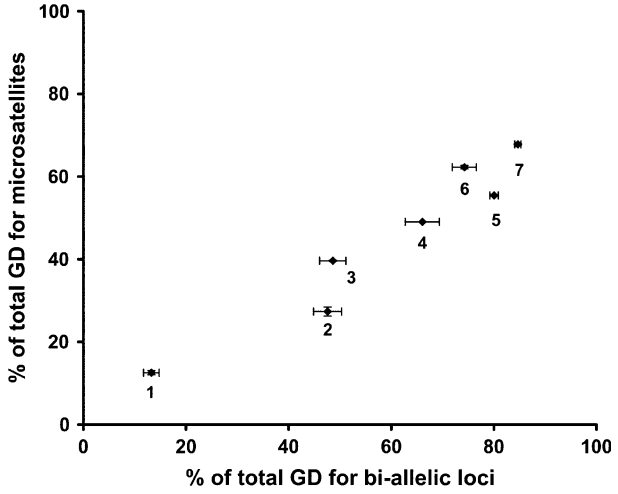

Figure 6. Conservation priority values estimated for 7 different geographical groupings using 7 biallelic markers and 14 microsatellites (Pearson's $r = 0.969$, $P = 0.0003$). Error bars represent the 95% confidence intervals based on 100 bootstrap trees. Numbering of the groups: 1) Kamennoe and Pisto 2) Kammennoe, Pisto and Saimaa; 3) Baltic Sea without Saimaa; 4) Baltic Sea with Saimaa; 5) White and Barents Seas and the Atlantic Ocean without Kamennoe and Pisto; 6) Baltic Sea, Kamennoe and Pisto; and 7) White and Barents Seas and the Atlantic Ocean with Kamennoe and Pisto. See also Table 1 and Figure 1.

Pairwise genetic divergence estimates, measured as F_{ST}, showed highly correlated patterns between biallelic and microsatellite markers (Figure 3). A further notable finding in the performance of a small set of biallelic loci was the congruency between biallelic and microsatellite markers in the isolation-by-distance signals observed in the anadromous populations (Figure 4a and b). This indicates that both marker types adequately reflect the underlying evolutionary forces, such as migration and drift, which influence the population genetic structure of Atlantic salmon. In contrast to estimates of genetic diversity, significant correlations between the genetic divergence values estimated for biallelic markers and microsatellites could still be observed when populations were split based on their life-history strategy. This implies that a fewer biallelic loci may be required to estimate genetic distance than are required to accurately estimate genetic diversity.

Although the biallelic markers provided in general congruent phylogenetic signal compared with microsatellites, the biallelic marker data alone were not sufficient to give strong support for clustering of European salmon populations (Figure 5a), probably as a result of the limited number of loci and the low number of alleles per locus (see Kalinowski 2005). The microsatellite NJ phylogram (Figure 5b) was almost identical to the original (Tonteri et al. 2005), with a minor reduction in the bootstrap values of some main nodes, possibly due to a smaller number of analyzed

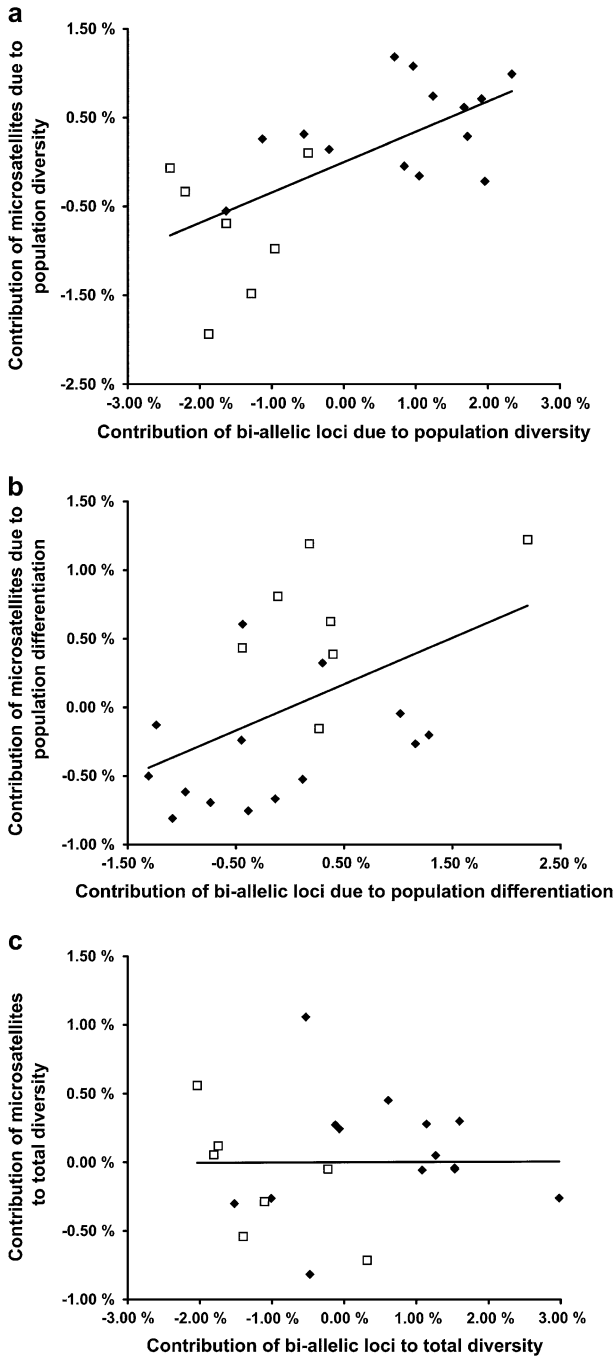

Figure 7. Comparison of 3 contribution estimates: (**a**) diversity component, (**b**) differentiation component, and (**c**) the total gene diversity for the analyzed populations based on data from 7 biallelic markers versus 14 microsatellites. Filled diamonds and open squares represent the anadromous and nonanadromous populations, respectively. Note the variable scales of the axes between the charts.

populations and individuals. However, when the biallelic marker data were combined with the microsatellite data for phylogenetic analyses, some changes in population groupings and slight increases in bootstrap support for some main

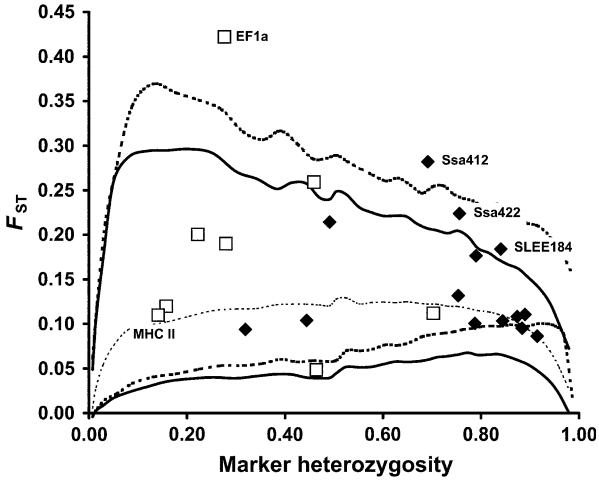

Figure 8. Estimated F_{ST} values plotted against heterozygosities for 8 biallelic loci and 14 microsatellite loci in Atlantic salmon populations from the Barents and White Seas and the Atlantic Ocean. Open squares and filled diamonds represent the biallelic loci and microsatellites, respectively. The simulated median (faint line), 0.95 confidential limit (dashed lines for biallelic loci and solid lines for microsatellites) values are also plotted based on expected F_{ST} value 0.17 and 0.14 for biallelic loci and microsatellite data, respectively.

nodes of the NJ phylogram were observed (Figure 5c). One noteworthy change in clustering was that the Saimaa population grouped together with the nonanadromous White Sea populations Kamennoe and Pisto, as opposed to clustering with Baltic Sea basin populations as in the study of Tonteri et al. (2005), where allozymes were combined with microsatellite markers. This indicates the relatively similar genetic distance signals of biallelic and microsatellite loci which somewhat differed from allozyme data.

Ascertainment bias is a further issue that needs to be considered when using biallelic markers such as SNPs for population genetic analyses as it may introduce a systematic bias in estimates of variation within and between populations (Kuhner et al. 2000; Nielsen 2000). In this study, such a bias may potentially have resulted in an overestimation of F_{ST} values in biallelic marker data as only a limited number of loci were used. However, this is assumed to have a minor role in the observed correlations between the signals from the 2 different marker types compared.

The genetic diversity content of biallelic and microsatellite loci from a conservation point of view was compared in this study, and the congruence between results obtained from different molecular markers was high for broad-scale comparisons (Figures 6 and 7a–c). Recently, there has been special interest in assessing genetic diversity using markers directly targeted at specific genes or gene families and encouraging increased use of gene-targeted markers in ecological and conservation studies (van Tienderen et al. 2002). The biallelic loci used in this study can be referred to as gene-targeted markers as they were identified around the coding regions of some functional genes (Ryynänen and Primmer 2004a, 2006). When the genetic diversity value was estimated for different groupings of salmon populations using either biallelic or microsatellite loci (Figure 6), the nonanadromous and Baltic Sea populations seemed to contain a relatively small proportion of the overall diversity. This may reflect the overall lower genetic diversity observed within the Baltic and nonanadromous populations compared with the Atlantic ones (Vasemägi, Nilsson, et al. 2005; Tonteri et al. 2007) However, when the nonanadromous and Baltic Sea populations were grouped together, their conservation priority estimate based on both marker types was relatively similar to that observed in the geographically more diverse anadromous White and Barents Seas and the Atlantic Ocean grouping (Figure 6). This suggests that different geographical lineages of Atlantic salmon populations hold unique allele combinations, which should be considered highly valuable for conservation purposes.

Each population's priority for conservation purposes was also assessed individually, and the split diversity and divergence contribution estimates correlated positively between different genetic markers (Figure 7a and b). Petit et al. (1998) suggested that both population divergence and diversity should be considered in a conservation context because they both contribute to total diversity. However, these total diversity estimates were not correlated between marker systems (Figure 7c), which may indicate that there is a difference in the impact of genetic divergence and genetic diversity components on the total diversity in microsatellites and biallelic markers. Furthermore, when populations were assessed based on their life history, no correlations existed between marker types for either genetic divergence or genetic diversity contribution and as a consequence, different populations would be prioritized for conservation depending on the marker type and conservation criteria applied. This emphasizes that further research is required before making general conclusions regarding which markers and molecular diversity/divergence criteria are most appropriate for developing conservation guidelines.

Some indications of potential signatures of selection in the analyzed biallelic and microsatellite loci were detected among 11 Atlantic salmon populations from the Barents and White Seas and the Atlantic Ocean. Three microsatellite loci identified as potential outliers here (Figure 8) were not the same as those reported earlier by (Vasemägi, Nilsson, et al. 2005), although the same microsatellite loci were included in both studies. Hence, it is possible that because the present study covered different set of populations compared with (Vasemägi, Nilsson, et al. 2005), the analysis of outlier loci identified different loci that might be affected by selection. However, it is also worth noting that the results here are based on only 1 F_{ST} neutrality test (Beaumont and Nichols 1996), and it has been suggested that potential outliers should be confirmed with at least 2 or more tests that are based on different models (e.g., Storz et al. 2004; Vasemägi, Nilsson, et al. 2005). Hence, the putative outlier status of these loci should be treated with some caution.

The locus named *EF1a*, whose primer sequences (Table 3) were designed using the published sequence of the translation elongation factor 1 alpha subunit in zebrafish (Gao et al. 1997), was the only biallelic locus identified as a putative outlier (Figure 8). The sequences of the 2 mRNAs coding for Atlantic salmon *EF1* paralogous genes have in fact been studied in terms of potential reference genes in expression studies (Olsvik et al. 2005). It should be noted, however, that as no clear sequence homology was found between the locus used here and the salmon *EF1a* genes (Ryynänen and Primmer 2006), the association of this biallelic marker with the corresponding gene in salmon is currently unclear. Therefore, further studies are warranted in order to determine the true homology of this gene sequence in salmon and whether the effects of selection might have implications for its use as a reference gene in expression studies.

Interestingly, the *MHC II beta* locus, which was selected as an a priori candidate for selection, did not deviate from the simulated neutral expectations (Figure 8), despite the fact that the locus has been identified as an outlier in several population genetic studies of Atlantic salmon (Landry and Bernatchez 2001; Vasemägi, Gross, et al. 2005) and has also been demonstrated to directly affect the survival in several disease trials (Langefors et al. 2001; Lohm et al. 2002) and during empirical experiments in natural environment (de Eyto et al. 2007).

To conclude, the main limitation on the use of biallelic markers in ecology and population genetics of nonmodel species continues to be the lack of markers, and this, together with the low information content compared with microsatellites (Kalinowski 2002), hinders their application in nonmodel organisms. The emphasis of this study was to evaluate the usefulness of a small set of biallelic markers compared with microsatellite data, especially from a population genetics and conservation genetics perspective in Atlantic salmon. The study demonstrates that a relatively small set of biallelic markers can be sufficient for obtaining generally concordant results in most of the analyses compared with microsatellites, although estimates of genetic distance are generally more concordant than genetic diversity. However, we emphasize that the use of a much higher number of loci is preferable, especially when the genetic differences between populations are subtle or individual multilocus genotype-based analysis is performed.

Funding

The Finnish Academy, the Finnish Graduate School of Population Genetics, University of Helsinki; Finnish Ministry of Agriculture and Forestry.

Acknowledgments

We would like to thank Paula Lehtonen and Leena Laaksonen for their excellent technical assistance and Alexey Veselov, Sergey Titov, Alexander

Appendix 1. Pairwise F_{ST} estimates based on 7 biallelic loci (below diagonal) and 14 microsatellites (above diagonal) of the studied Atlantic salmon populations (the estimates, which were not significantly different from 0 after the Bonferroni correction, are indicated as boldface)

Pop.	Dee	Kam	Liz	Meg	Nev	Nil	Sai	Pis	Piz	Pon	Pul	Shu	Sum	Sys	Tai	Ten	Tor	Tul	Unj	Var	Vin
Dee		0.2731	0.2367	0.1196	0.1624	0.1857	0.3247	0.2205	0.1771	0.1373	0.1734	0.2168	0.1631	0.2694	0.2150	0.0654	0.1230	0.0784	0.2147	0.1181	0.2052
Kam	0.1250		0.4126	0.2740	0.3246	0.3771	0.4957	0.2255	0.3822	0.3340	0.4070	0.4459	0.3173	0.4819	0.3465	0.2684	0.3014	0.2494	0.4758	0.2789	0.3972
Liz	0.3056	0.5168		0.1928	0.1365	0.3186	0.4674	0.3709	0.2393	0.2319	0.2877	0.1282	0.3015	0.2273	0.1797	0.2096	0.2136	0.2089	0.3028	0.2362	0.2481
Meg	0.0828	0.1156	0.3654		0.1014	0.1772	0.3177	0.2178	0.0643	0.0760	0.1523	0.2002	0.0808	0.2347	0.1712	0.0732	0.1170	0.0694	0.1240	0.0540	0.1391
Nev	0.1402	0.3595	0.0920	0.2630		0.2197	0.3071	0.2385	0.1578	0.1331	0.1372	0.1604	0.1713	0.1528	0.0905	0.1140	0.0891	0.0981	0.2283	0.1208	0.0760
Nil	0.1984	0.3770	0.5910	0.1678	0.3727		0.4892	0.3126	0.2583	0.1286	0.2181	0.3048	0.2264	0.3353	0.2538	0.1428	0.1906	0.1285	0.3250	0.1292	0.2687
Sai	0.2931	0.3061	0.6490	0.2550	0.5123	0.5728		0.3555	0.4079	0.3558	0.3945	0.4992	0.3689	0.4979	0.3173	0.2907	0.3233	0.2800	0.4882	0.3262	0.3169
Pis	0.0883	**0.0882**	0.4815	0.1265	0.2734	0.2851	0.4911		0.2888	0.2554	0.2672	0.3990	0.2502	0.4095	0.2426	0.1969	0.2093	0.1733	0.3914	0.2148	0.2828
Piz	0.1860	0.3850	0.2303	0.1916	0.1496	0.2456	0.4150	0.3498		0.1222	0.2305	0.2504	0.1309	0.2717	0.1853	0.1211	0.1989	0.1163	0.0796	0.1206	0.2332
Pon	0.0694	0.2570	0.3248	0.1069	0.1477	**0.0829**	0.4526	0.1467	0.1510		0.1259	0.2212	0.1062	0.2517	0.2017	0.0686	0.1426	0.0636	0.1751	0.0546	0.1642
Pul	0.2691	0.4712	0.4539	0.2631	0.3172	0.1390	0.5895	0.3759	0.1969	0.1752		0.2968	0.2142	0.3052	0.2112	0.1559	0.1624	0.1264	0.3070	0.1502	0.1559
Shu	0.3034	0.5516	**0.0618**	0.3720	**0.0690**	0.5661	0.7092	0.4769	0.2564	0.2787	0.4100		0.2916	0.2323	0.1952	0.2191	0.2325	0.2150	0.3347	0.2394	0.2610
Sum	0.2555	0.3856	0.6090	0.1780	0.4442	0.2719	0.6280	0.2593	0.4125	0.1909	0.3562	0.5786		0.3097	0.2249	0.1296	0.1822	0.1144	0.2225	0.0655	0.2057
Sys	0.2351	0.5204	**0.1220**	0.3235	**0.0164**	0.4482	0.6805	0.4103	0.1858	0.1905	0.3130	**0.0098**	0.5099		0.1623	0.2286	0.2324	0.2114	0.3535	0.2518	0.2902
Tai	0.2394	0.3627	0.2402	0.3031	0.1871	0.4426	0.4847	0.3086	0.2321	0.2913	0.4264	0.2683	0.5065	0.2288		0.1789	0.1839	0.1582	0.2772	0.1970	0.2023
Ten	0.1109	0.2674	0.2904	0.1087	0.1700	0.1911	0.3865	0.1950	0.1390	0.0672	0.2640	0.2666	0.2748	0.1958	0.2065		0.0852	0.0193	0.1605	0.0635	0.1608
Tor	0.2418	0.5547	0.3988	0.3637	0.1484	0.4422	0.6765	0.4496	0.2808	0.1850	0.3165	0.3051	0.4854	**0.1627**	0.4131	0.2391		0.0866	0.2463	0.1149	0.1024
Tul	0.0227	0.0914	0.3475	**0.0343**	0.1925	0.1650	0.2857	0.0785	0.1896	0.0596	0.2931	0.3448	0.2383	0.2796	0.2621	0.0666	0.3238		0.1592	0.0535	0.1457
Unj	0.1652	0.3753	0.2839	0.1599	0.1442	0.2620	0.5045	0.3070	**0.0255**	0.1053	0.2580	0.2908	0.3359	0.2004	0.2082	0.0973	0.2985	0.1550		0.1773	0.3052
Var	0.1568	0.2358	0.3947	0.0468	0.2856	0.1804	0.4359	0.1660	0.2280	0.0863	0.2783	0.3739	0.1223	0.3169	0.3393	0.0904	0.3589	0.0888	0.1471		0.1421
Vin	0.3156	0.6400	0.5670	0.4255	0.3068	0.5308	0.7209	0.5697	0.3697	0.2742	0.3810	0.5111	0.5520	0.3861	0.5497	0.3292	0.0684	0.4184	0.4215	0.4370	

Zubchenko, Svjatoslav Kaluzhin, Igor Bakhmet, Jorma Piironen, and John Taggart for supplying salmon tissue samples. Miika Tapio gave valuable assistance with the CONTRIB analyses.

References

Allendorf FW, Thorgaard GH. 1984. Tetraploidy and the evolution of salmonid fishes. In: Turner BJ, editor. The evolutionary genetics of fishes. New York: Plenum Press. p. 1–53.

Anderson EC, Garza JC. 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics. 172:2567–2582.

Bailey-Wilson J, Almasy L, de Andrade M, Bailey J, Bickeboller H, Cordell H, Daw EW, Goldin L, Goode E, Gray-McGuire C, et al. 2005. Genetic Analysis Workshop 14: microsatellite and single-nucleotide polymorphism marker loci for genome-wide scans. BMC Genet. 6:S1.

Beaumont MA, Nichols RA. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B Biol Sci. 263:1619–1626.

Britten RJ, Rowen L, Williams J, Cameron RA. 2003. Majority of divergence between closely related DNA samples is due to indels. Proc Natl Acad Sci USA. 100:4661–4665.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol. 18:249–256.

Budowle B. 2004. SNP typing strategies. Forensic Sci Int. 146 (Suppl.):S139–S142.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. Am J Hum Genet. 19:233–257.

Chen X, Sullivan P. 2003. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenomics J. 3:77–96.

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, et al. 2004. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. Plant J. 37: 778–786.

Crozier RH, Agapow P-M, Pedersen K. 1999. Towards complete biodiversity assessment: an evaluation of the subterranean bacterial communities in the Oklo region of the sole surviving natural nuclear reactor. FEMS Microbiol Ecol. 28:325–334.

Crozier RH, Kusmierski RM. 1994. Genetic distances and the settings of conservation priorities. In: Loeschcke V, Tomiuk J, Jain SK, editors. Conservation genetics. Basel (Switzerland): Birkhäuser Verlag. p. 227–237.

de Eyto E, McGinnity P, Consuegra S, Coughlan J, Tufto J, Farrell K, Megens H-J, Jordan W, Cross T, Stet RJM. 2007. Natural selection acts on Atlantic salmon major histocompatibility (MH) variability in the wild. Proc R Soc Lond B Biol Sci. 274:861–869.

Felsenstein J. 1995. Department of Genome Sciences and Department of Biology. Phylip (Phylogeney Inference Package) version 3.573c. Seattle (WA): University of Washington.

Gao D, Li Z, Murphy T, Sauerbier W. 1997. Structure and transcription of the gene for translation elongation factor 1 subunit alpha of zebrafish (*Danio rerio*). Biochim Biophys Acta. 1350:1–5.

Glaubitz JC, Rhodes OE, Dewoody JA. 2003. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Mol Ecol. 12:1039–1047.

Goudet J. 2001. Lausanne University, Lausanne, Switzerland. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). [Internet]. [cited 2006 Jan]. Available from http://www2.unil.ch/popgen/softwares/fstat.html.

Hayes BJ, Nilsen K, Berg PR, Grindflek E, Lien S. 2007. SNP detection exploiting multiple sources of redundancy in large EST collections improves validation rates. Bioinformatics. 23:1692–1693.

Hedrick PW. 2005. A standardized genetic differentiation measure. Evolution. 59:1633–1638.

Kalinowski ST. 2002. How many alleles per locus should be used to estimate genetic distances? Heredity. 88:62–65.

Kalinowski ST. 2005. Do polymorphic loci require large sample sizes to estimate genetic distances? Heredity. 94:33–36.

Kruglyak L. 1997. The use of a genetic map of biallelic markers in linkage studies. Nat Genet. 17:21–24.

Kuhner MK, Beerli P, Yamato J, Felsenstein J. 2000. Usefulness of single nucleotide polymorphism data for estimating population parameters. Genetics. 156:439–447.

Kwok P-Y. 2001. Methods for genotyping single nucleotide polymorphisms. Annu Rev Genomics Hum Genet. 2:235–258.

Landry C, Bernatchez L. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). Mol Ecol. 10:2525–2539.

Langefors Å Lohm J, Grahn M, Andersen Ø von Schantz T. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. Proc R Soc Lond B Biol Sci. 268:479–485.

Lohm J, Grahn M, Langefors Å A, Andersen Ø , Storset A, von Schantz T. 2002. Experimental evidence for major histocompatibility complex-allele-specific resistance to a bacterial infection. Proc R Soc Lond B Biol Sci. 269:2029–2033.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet. 4:981–994.

Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution. 60:2399–2402.

Morin PA, Luikart G, Wayne RK. 2004. the SNP workshop group SNPs in ecology, evolution and conservation. Trends Ecol Evol. 19:208–216.

Nielsen R. 2000. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. Genetics. 154:931–942.

Olsvik PA, Lie KK, Jordal A-EO, Nilsen TO, Hordvik I. 2005. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. BMC Mol Biol. 6:21.

Orsini L, Pajunen M, Hanski I, Savilahti H. 2007. SNP discovery by mismatch-targeting of Mu transposition. Nucleic Acids Res. 35:e44

Park SDE. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. [PhD dissertation]. Dublin, Ireland: University of Dublin.

Peakall R, Smouse PE. 2005. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 6:288–295.

Petit RJ, El Mousadik A, Pons O. 1998. Identifying populations for conservation on the basis of genetic markers. Conserv Biol. 12:844–855.

Ranade K, Chang M-S, Ting C-T, Pei D, Hsiao C-F, Olivier M, Pesich R, Hebert J, Chen Y-DI, Dzau VJ, et al. 2001. High-throughput genotyping with single nucleotide polymorphisms. Genome Res. 11:1262–1268.

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

Rengmark AH, Slettan A, Skaala O, Lie O, Lingaas F. 2006. Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. Aquaculture. 253:229–237.

Rice WR. 1989. Analyzing tables of statistical tests. Evolution. 43:223–225.

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

Ryynänen HJ, Primmer CR. 2004a. Distribution of genetic variation in the growth hormone 1 gene in Atlantic salmon (*Salmo salar*) populations from Europe and North America. Mol Ecol. 13:3857–3869.

Ryynänen HJ, Primmer CR. 2004b. Primers for sequence characterization and polymorphism detection in the Atlantic salmon (*Salmo salar*) growth hormone 1 (*GH1*) gene. Mol Ecol Notes. 4:664–667.

Ryynänen HJ, Primmer CR. 2006. Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (*Salmo salar*) and other salmonid fishes. BMC Genomics. 7:192.

Saetre G-P, Borge T, Lindroos K, Haavie J, Sheldon BC, Primmer CR, Syvänen A-C. 2003. Sex chromosome evolution and speciation in Ficedula flycatchers. Proc R Soc Lond B Biol Sci. 270:53–59.

Saino N, Primmer CR, Ellegren H, Moller AP. 1997. An experimental study of paternity and tail ornamentation in the barn swallow (*Hirundo rustica*). Evolution. 51:562–570.

Schlötterer C. 2004. The evolution of molecular markers—just a matter of fashion? Nat Rev Genet. 5:63–69.

Seddon JM, Parker HG, Ostrander EA, Ellegren H. 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. Mol Ecol. 14:503–511.

Smith CT, Elfstrom CM, Seeb LW, Seeb JE. 2005. Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Mol Ecol. 14:4193–4203.

Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction. Am J Hum Genet. 73:1162–1169.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 68:978–989.

Storz JF, Payseur BA, Nachman MW. 2004. Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. Mol Biol Evol. 21:1800–1811.

Syvänen A-C. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. Nat Rev Genet. 2:930–942.

Tonteri A, Titov S, Veselov A, Zubchenko A, Koskinen MT, Lesbarrères D, Kaluzchin S, Bakhmet I, Lumme J, Primmer CR. 2005. Phylogeography of anadromous and non-anadromous Atlantic salmon (*Salmo salar*) from northern Europe. Ann Zool Fenn. 42:1–22.

Tonteri A, Veselov A. Je, Titov S, Lumme J, Primmer CR. 2007. The effect of migratory behaviour on genetic diversity and population divergence: a comparison of anadromous and freshwater Atlantic salmon (*Salmo salar* L.). J Fish Biol. 70:381–398.

van Tienderen PH, de Haan AA, van der Linden CG, Vosman B. 2002. Biodiversity assessment using markers for ecologically important traits. Trends Ecol Evol. 17:577–582.

Vasemägi A, Gross R, Paaver T, Koljonen M-L, Säisä M, Nilsson J. 2005. Analysis of gene associated tandem repeat markers in Atlantic salmon (*Salmo salar* L.) populations: implications for restoration and conservation in the Baltic Sea. Conserv Genet. 6:385–397.

Vasemägi A, Nilsson J, Primmer CR. 2005. Expressed sequence tag-linked microsatellites as a source of gene associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Mol Biol Evol. 22:1067–1076.

Vignal A, Milana D, SanCristobala M, Eggenb A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol. 34:275–305.

Received October 3, 2006
Accepted September 21, 2007

Corresponding Editor: Lisa Seeb