

# An examination of genetic diversity and effective population size in Atlantic salmon populations

NATACHA NIKOLIC<sup>1\*</sup>, JAMES R. A. BUTLER<sup>2</sup>, JEAN-LUC BAGLINIÈRE<sup>3</sup>,  
ROBERT LAUGHTON<sup>4</sup>, IAIN A. G. MCMYNN<sup>5</sup> AND CLAUDE CHEVALET<sup>1</sup>

<sup>1</sup> Laboratoire de Génétique Cellulaire (UMR 444), INRA-ENVT, BP 52627, 31326 Castanet Tolosan Cedex, France

<sup>2</sup> CSIRO Sustainable Ecosystems, James Cook University, PO Box 12139, Earlville BC, Cairns, QLD 4870, Australia

<sup>3</sup> INRA-Agrocampus, 65 rue de St Briec CS 84215, 35042 Rennes, France

<sup>4</sup> Spey Fishery Board and Spey Research Trust, 1 Nether Borlum Cottage, Knockando, Aberlour, Morayshire AB38 7SD, UK

<sup>5</sup> Kyle of Sutherland District Salmon Fishery Board, c/o Bell Ingram Estate Office, Bonar Bridge, Sutherland IV24 3EA, UK

(Received 18 May 2009 and in revised form 26 October 2009)

## Summary

Effective population size ( $N_e$ ) is an important parameter in the conservation of genetic diversity. Comparative studies of empirical data that gauge the relative accuracy of  $N_e$  methods are limited, and a better understanding of the limitations and potential of  $N_e$  estimators is needed. This paper investigates genetic diversity and  $N_e$  in four populations of wild anadromous Atlantic salmon (*Salmo salar* L.) in Europe, from the Rivers Oir and Scorff (France) and Spey and Shin (Scotland). We aimed to understand present diversity and historical processes influencing current population structure. Our results showed high genetic diversity for all populations studied, despite their wide range of current effective sizes. To improve understanding of high genetic diversity observed in the populations with low effective size, we developed a model predicting present diversity as a function of past demographic history. This suggested that high genetic diversity could be explained by a bottleneck occurring within recent centuries rather than by gene flow. Previous studies have demonstrated the efficiency of coalescence models to estimate  $N_e$ . Using nine subsets from 37 microsatellite DNA markers from the four salmon populations, we compared three coalescence estimators based on single and dual samples. Comparing  $N_e$  estimates confirmed the efficiency of increasing the number and variability of microsatellite markers. This efficiency was more accentuated for the smaller populations. Analysis with low numbers of neutral markers revealed uneven distributions of allelic frequencies and overestimated short-term  $N_e$ . In addition, we found evidence of artificial stock enhancement using native and non-native origin. We propose estimates of  $N_e$  for the four populations, and their applications for salmon conservation and management are discussed.

## 1. Introduction

The concept of effective population size ( $N_e$ ) was introduced by Wright (1931, 1938) and can be defined as the number of breeding individuals in an idealized population that would show the same amount of variation of allele frequencies under random genetic drift. This parameter is usually smaller than the absolute population size. It is a key parameter in conservation and management because it affects the degree to which a population can respond to selection.  $N_e$  influences

the rate of loss of genetic diversity, the rate of fixation of deleterious alleles and the efficiency of natural selection at maintaining beneficial alleles (Berthier *et al.*, 2002). If  $N_e$  declines too far, the loss of genetic variation resulting from genetic drift may put species or populations at risk of extinction by losing the raw material on which selection can operate.

Demographic data from mating systems or genetic data can be used to infer  $N_e$ . Unfortunately demographic data are generally difficult to collect in many wild populations and life-history information is often unavailable with sufficient precision to make a good estimate of  $N_e$  (Frankham, 1995). Genetic methods

\* Corresponding author. Tel: +33.5.61.28.55.68. Fax: +33.5.61.28.53.08. e-mail: natanicholic@hotmail.com

are more widely used. They are increasingly being developed by applying polymorphic molecular markers to resolve taxonomic problems and describe evolutionary and demographic history of species and populations. Currently, many different methods are available to infer  $N_e$  from genetic data in one or multiple samples. The most widely used estimator consists of measuring the variance of allele frequencies between generations (Tallmon *et al.*, 2004). However, several studies noted that it is often biased towards high values (Luikart *et al.*, 1999; Wang, 2001; Berthier *et al.*, 2002). Phylogenetic methods are more efficient than non-phylogenetic methods (Felsenstein, 1992), primarily due to additional information provided by the tree structure. Genealogical modelling has greatly facilitated the estimation of demographic and mutational parameters using length variants microsatellite data (Chakraborty & Kimmel, 1999; Feldman *et al.*, 1999; King *et al.*, 2000 by Storz & Beaumont, 2002). In the past, most studies of genetic variation had used protein electrophoresis (Guyomard, 1994). However, protein polymorphism has a limited potential for discrimination (Norris *et al.*, 1999).

In this paper, we focus on highly variable microsatellite loci due to their power to detect genetic structures within and among populations. Furthermore, their assumed neutrality avoids the effects of selection, and allows the standard coalescence model to be used to estimate effective population sizes from the distribution of allelic frequencies. Effects of migration are integrated in some models but make the estimation more complicated. In the framework of the coalescence approach, a tree linking the alleles up to their common ancestor describes the relationships among alleles. A coalescence event appears each time two lineages in the tree join into a common ancestor, and the intervals between such events have a distribution that depends on  $N_e$  (Kuhner *et al.*, 1995). The usual approach for estimating this parameter is to compare statistics calculated from empirical datasets with a distribution generated by Monte Carlo simulations of the coalescent process (Wilson & Balding, 1998; Storz & Beaumont, 2002).

Here, we compare estimates of  $N_e$  based on the coalescence model, and the performances of such methods based either on a single sample (MSVAR (Beaumont, 1999) and DIYABC (Cornuet *et al.*, 2008)) or on two samples (TM3; Berthier *et al.*, 2002). Hence, we also compare a likelihood approach (TM3 and MSVAR) and Approximate Bayesian Computation (DIYABC). Comparative studies on real data that gauge the relative performances of  $N_e$  methods are few. Salmonids are an ideal species for assessing population structure's influence on  $N_e$  estimations. They show a propensity for structuring into genetically distinct populations due to the combined effects

of their rearing habitat and their homing behaviour (Stabell, 1984). Here, we use four wild anadromous (i.e. adults migrate from the sea to breed in freshwater) European Atlantic salmon populations from north-west France (Rivers Oir and Scorff) and north-east Scotland (Rivers Spey and Shin) (Fig. 1). Atlantic salmon are subject to many pressures in Europe, including physical barriers to migration, exploitation by net and rod fisheries, pollution, the introduction of non-native salmon stocks, physical degradation of spawning and nursery habitat, and increased marine mortality. During the last 30 years, the decline of wild salmon on both sides of the North Atlantic (Parrish *et al.*, 1998; Jonsson & Jonsson, 2004) has affected populations to differing degrees (Hawkins, 2000). The four populations studied are pressured by different factors and are therefore subject to varying conservation and management strategies. Because their characteristics are well understood (Baglinière & Champigneulle, 1986; Baglinière *et al.*, 2005; Butler, 2004; Butler *et al.*, 2008), and they have large differences in abundance, they provide a useful opportunity to design tools for estimating  $N_e$ .

For each population, fish from two samples have been genotyped with 37 microsatellite markers chosen for their high genetic quality (Nikolic *et al.*, 2009). Based on a general analysis of their present genetic variability and structure, and on different subsets of markers, this paper provides an overview on the performance and sensitivity of three different estimators of  $N_e$  applied to declining populations of different sizes.

## 2. Materials and methods

### (i) Current status and management

Atlantic salmon provide highly valued ecosystem services. They support rod and net fisheries (e.g. Butler *et al.*, 2009), and provide source stock for an aquaculture industry that produces over a hundred million Atlantic salmon, and whose biomass already exceeds that of wild populations (Gross, 1998). However, wild Atlantic salmon are considered an endangered species. As a consequence of contracting range, a decline in abundance and the modification of demographic structure of adult populations (Anonymous, 2001, 2003; Caron & Fontaine, 2003), the species has been placed on the Red List of threatened species in Europe (Porcher & Baglinière, 2001). Recent studies (e.g. McGinnity *et al.*, 2003; Finnegan & Stevens, 2008) have shown that the introduction of non-native salmon stocks can result in the alteration of genetic structure of the recipient populations, posing another potential threat to population fitness.

Atlantic salmon conservation and management is complicated by the species' life history. Juvenile

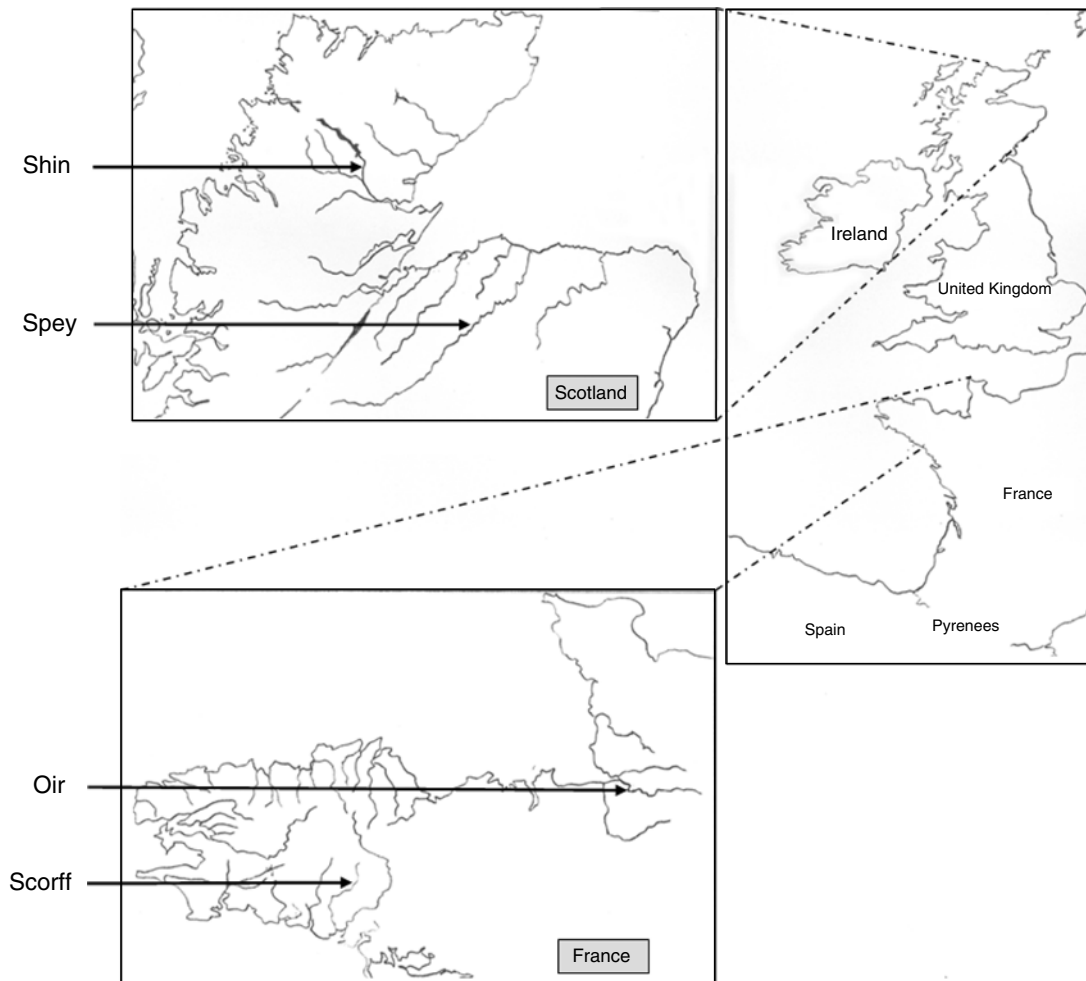


Fig. 1. Geographic locations of the four wild Atlantic salmon populations: Rivers Oir and Scorff in north-west France and the Rivers Shin and Spey in north-east Scotland.

salmon spend 1–4 years in freshwater depending on the growth conditions and the latitudinal position of the river (Baglinière, 1976). They then migrate to the ocean and return as adults to spawn after 1 year ('grilse') or multiple years ('Multi-sea winter' or MSW) (Klemetsen *et al.*, 2003). The Atlantic salmon is a good example of a highly migratory species with complex spatial and temporal patterns demonstrating significant local adaptations (Taylor, 1991), homing behaviour (Hansen & Jonsson, 1994; Saglio, 1994) and reproductive strategies (Fleming, 1996).

Table 1 shows the differences in habitat, census population size, fishery pressure, and conservation and management regimes for the Oir, Scorff, Spey and Shin. Census population sizes were estimated by extrapolating from in-river rod catches using known exploitation rates. The Scorff, Shin and Spey are coastal catchments, whereas the Oir is an estuarine tributary of the river Sélune and the most productive area for salmon within the Sélune catchment. Salmon in the Spey are little affected by artificial barriers to migration, whereas the Oir, Scorff and Shin are more

affected. While water quality in the Oir and Scorff is impacted by human pressures (mainly agricultural practices and urban run-off in the estuaries), quality in the Spey and Shin is not acutely affected by anthropogenic activities. The Spey and Shin juvenile populations predominately migrate to sea after 2–4 years, and returning adults include both grilse and MSW fish. In the Oir and Scorff, most juveniles migrate after 1 year, and most returning adults are grilse. Exploitation rates by net and rod fisheries are low (10–20%) for the Shin and Spey, but high (30–50%) for the Oir and Scorff. According to the International Union for the Conservation of Nature's (IUCN, 1994) classification of conservation status, the Spey and Shin populations are of 'minor concern', whereas the Oir and Scorff populations are 'vulnerable'.

There is a history of stock enhancement in all of the rivers. Hatchery-reared juvenile salmon of unknown Scottish origin have been planted in the Scorff during 1973–1979 (Baglinière, 1979). In the Oir, some returning adults might have originated from hatchery-reared smolts sourced from native Sélune broodstock

Table 1. *Biotic and abiotic characteristics of the four studied salmon populations and their catchments*

River	Oir	Scorff	Shin	Spey
Mean annual water temperature fluctuates (°C)	3–19	5–20	1–17	1–20
Drought year	1976, 2003	1976, 2003	1976, 2003	1976, 2003
Catchment area (km <sup>2</sup> )	87	480	494.6	3000
Streamflow (m <sup>3</sup> /s)	0.86	5	15.2	65
River length (km)	19.5	75	27.2	172
Slope (‰)	11	3.6	NA	22
Soil type	Schiste and granite	Schiste and granite	Schiste and gneiss	Schiste and gneiss
Recent pH	7.1–7.8	7	6–7	6–9
Recent nitrate concentrations (mg/l)	High:30	18.4	12.5	27
Number of major artificial barriers	2	16	2	3
Adult salmon population structure 2005	Grilse (90%)	Grilse (90%)	MSW and grilse	MSW and grilse
Adult salmon population structure 1988 (1992 for Shin)	Grilse	MSW and grilse	MSW and grilse	MSW and grilse
Census population size (N) in 2005	130	1000	3000	< 60 000 <sup>a</sup>
Census population size (N) in 1988 (1992 for Shin)	230–260	NA	2000–4000	< 60 000 <sup>a</sup>
Conservation and management policies	1976 Plan Saumon	1976 Plan Saumon	1986 Salmon Act	1986 Salmon Act, 1999 EU Habitats Directive Special Area of Conservation
Conservation status (IUCN 1994)	Vulnerable	Vulnerable	Minor concern	Minor concern
Coastal net fisheries	Present	Present	Present but declining	Present but negligible effect
River rod fisheries	March–July (extension possible until October)	March–July (extension possible until October)	11 January–30 September	11 February–30 September
Estimated total exploitation rates	30–50%	30–50%	10–20%	10–20%
Non-native stock enhancement programmes	Introduction of Sélune smolts, 1990s	Non-native juveniles of unknown Scottish source 1973–1979	Some fish farm escapees	Some fish farm escapees
Introduction of native salmon stocks	None	None	Native juveniles introduced to mitigate hydroelectricity dam impacts since 1950s	Native juveniles introduced since 1970s

<sup>a</sup> Spey samples were taken from an upper catchment sub-stock of unknown size, estimated to be less than the minimum total adult population.  
NA = not available.

in the 1990s. In the Spey, native wild broodstock have been used to source ova and juveniles for enhancing production in areas above impassable obstacles since the 1970s. In the Shin, a hatchery program was established to mitigate the effects of the installation of hydroelectricity dams in the 1950s, based on native broodstock. Escaped farmed salmon occasionally enter the Spey and Shin in small numbers.

The Spey supports one of the largest Atlantic salmon populations in Scotland, with at least 60 000 adults entering the river annually (Butler, 2004). Because of the relatively pristine nature of the river's habitat, and the status of the salmon population, the Spey was designated a Special Area of Conservation under the EU Habitats Directive (Council Directive 92/43/EEC) in 1999, with Atlantic salmon as a

qualifying species. The objectives of special area of conservations are to avoid deterioration of the habitats of qualifying species or significant disturbance to those species, ensuring that the integrity of the site is maintained and it achieves favourable conservation status of the qualifying features (Anonymous, 2000).

#### (ii) *Samples and DNA extraction*

A total of 367 wild adult anadromous Atlantic salmon were sampled during the spawning migration in 2005 and 1988 for the Scorff, Oir and Spey, and in 2005 and 1992 for the Shin. For the Scorff, Oir and Shin, it was assumed that the samples were representative of the annual adult spawning populations, which are small (100–4000 fish; Table 1). However, in the Spey samples were taken from the upper catchment, where spring-running fish are known to originate (Laughton, 1991). Larger populations in rivers such as the Spey are known to contain genetically distinct population units ('sub-stocks'), which differ in the timing of their return migration (Stewart *et al.*, 2002; Jordan *et al.*, 2005). Consequently, it was assumed that the Spey samples were taken from a sub-stock of spring-running fish of unknown size, but of less than 60 000. In all rivers, the individuals came from the same cohort (Table 2) to avoid biases in effective population sizes (Jorde & Ryman, 1996).

Sample sizes from each river ranged between 89 and 96 individuals. Pectoral or caudal fin clips were conserved in 95% ethanol, and scale samples were placed and dried in paper envelopes. All samples were kept at ambient temperature. Genomic DNA from fin and scale samples was extracted by boiling samples in 230  $\mu$ l solution (proteinase K, TE buffer (tris/EDTA) and chelex) at 55 °C for several hours, with a final period of 105 °C for 15 min (Estoup *et al.*, 1996; S. Launey, personal communication). After one night at 4 °C, the supernatant was diluted in 200  $\mu$ l of chelex and then stored at –20 °C.

#### (iii) *Genotyping*

In the present study, a set of 37 salmon microsatellite loci from the Salmon Genome project (<http://www.salmongenome.no/cgi-bin/sgp.cgi>) previously identified by Nikolic *et al.* (2009) were screened in all the individuals with the M13 labelling method (Schuelke, 2000). PCR for all 367 individuals was handled in 10  $\mu$ l within a 384 plate (TECAN 200): 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M dNTPs, 0.1  $\mu$ M forward primer, 0.15  $\mu$ M reverse primer, 0.15  $\mu$ M of M13-Fluo, 25–50 ng DNA and 0.5 unit Taq DNA polymerase. Precisions on the primers and amplification conditions for each marker are given in Nikolic *et al.* (2009). A 2  $\mu$ l volume of PCR product was added to 8  $\mu$ l of deionized formamide and the internal size

standard GENESCAN-400HD Rox (Applied Biosystems). Individual genotypes were obtained using ABI 3730 multi-capillary sequencer. Fluorescent DNA fragments size data were labelled by Genescan Analysis Software v3.7 (Applied Biosystems) to assign individuals by Genotyper 3.7 NT software (Applied Biosystems). From 10 random replicates, we evaluated genotyping error as relatively common (2.16%). From the initial set of 37 microsatellite markers, subsets of 28, 20, 10 and 5 markers have been selected according to their highest (H+ subsets) or lowest (H– subsets) observed heterozygosity, as estimated from the 367 individuals (Nikolic *et al.*, 2009).

#### (iv) *Historical processes*

##### (a) *Genetic diversity*

The genetic structure of the populations was assessed from the complete genotypic dataset involving the 37 markers. Genetic diversity parameters were estimated using GENETIX software version 4.05.2 (Belkhir *et al.*, 1998) and GenAlEx 6 software (Peakall & Smouse, 2006). Mean number of alleles per locus, actual heterozygosity and unbiased expected heterozygosity were calculated using the Hardy–Weinberg equilibrium (Nall, Hobs and Hnb, respectively). The inbreeding coefficient  $F_{IS}$  per population was estimated with 10 000 bootstraps by GENETIX software.

We developed a coalescent model to predict final genetic diversity when population size has undergone past changes, DemoDivMS (available at <https://qgp.jouy.inra.fr/>). Based on the Stepwise Mutation Model, this allows the present diversity at a microsatellite marker locus to be predicted as a function of present and past population size and mutation rate. Analytical calculations provide 'Pk distributions' (Shriver *et al.*, 1997), i.e. the expected frequency of pairs of alleles that are alike in state, (i.e. the expected homozygosity at the locus) and the frequencies of pairs of alleles with any given difference of the numbers of the microsatellite motif (Chevalet & Nikolic, 2010). This program includes an extension allowing for a continuous gene flow from a stable large population.

For the smaller populations (Oir and Scorff), we simulated the evolution of small populations (200–1000 individuals), derived from a larger population (10 000–50 000 effective size) 2000 to 4000 generations ago. We built a scenario in three steps assuming known ancestral and current effective sizes. Between the origin and the present day, we investigated the occurrence of a bottleneck, and checked various values of the effective sizes before and during the bottleneck, and of the times when population size changed and of the rate of immigration ( $Nm$  between



Table 2. Demographic and genetic parameters of the four studied salmon populations

Rivers	Oir	Scorff	Shin	Spey	Oir	Scorff	Shin	Spey
Names (Fig. 4)	Oir88	Scorff88	Shin92	Spey88	Oir	Scorff	Shin	Spey
Sampling year	1988	1988	1992	1988	2005	2005	2005	2005
Cohort year	1985	1985	1989	1985	2003	2003	2003	2003
Adults sampled	47	45	48	48	48	48	41	42
<i>FIS</i>	0.014	0.047	-0.005	0.038	0.003	0.001	0.005	0.001
95% interval	[-0.022-0.024]	[-0.006-0.058]	[-0.038-0.006]	[-0.003-0.049]	[-0.033-0.017]	[-0.035-0.015]	[-0.033-0.016]	[-0.030-0.006]
<i>Hnb</i>	0.80	0.79	0.78	0.84	0.81	0.80	0.78	0.83
<i>Hobs</i>	0.79	0.76	0.78	0.80	0.81	0.80	0.78	0.82
<i>Nall</i>	11	11	13	15	13	13	11	15
Mean <i>FIS</i>			0.021				0.001	
95% interval			[0.005-0.037]				[-0.012-0.014]	
Mean <i>FST</i>			0.055				0.049	
95% interval			[0.048-0.063]				[0.043-0.055]	
Mean <i>FIT</i>			0.075				0.050	
95% interval			[0.059-0.091]				[0.038-0.062]	
<i>Nm</i>			3.58				3.43	
Ancestral <i>Ne</i> (a)	49 015	54 883	47 941	54 780	50 106	64 152	54 682	64 416
95% interval	[8201-2 96 625]	[9230-3 28 417]	[8000-2 79 413]	[9099-3 30 348]	[8510-2 93 430]	[10 595-3 86 467]	[9273-3 21 886]	[9912-3 72 382]
Ancestral <i>Ne</i> (b)	18 186	29 643	18 776	14 500	29 995	27 235	29 995	13 371
95% interval	[4926-35 466]	[910-39 664]	[798-42 136]	[2404-40 061]	[6547-35 247]	[826-40 713]	[942-39 172]	[1963-42 608]
<i>Tf</i> (a)	10 038	11 475	7945	19 374	6243	10 985	7717	14 245
95% interval	[1580-65 963]	[1852-73 502]	[1275-50 047]	[2387-1 96 931]	[1014-39 214]	[1588-79 286]	[1245-49 084]	[1982-1 21 566]
<i>Tf</i> (b)			12 876				13 652	
95% interval			[1076-29 448]				[1252-31 704]	
Growth rate (g)	-307	-127	-247	-251	-284	-234	-209	-236
Current <i>Ne</i> (a)	501	831	598	12 503	383	689	304	7344
95% interval	[63-3767]	[113-5899]	[81-4384]	[1940-80 556]	[51-2968]	[81-5549]	[36-2468]	[1076-50 364]
Current <i>Ne</i> (b)	45	1165	1840	9596	100	1174	1842	9417
95% interval	[1-485]	[100-1798]	[140-3758]	[958-18 052]	[12-661]	[102-1791]	[144-3764]	[963-18 049]
Populations	Oir88/2005	Scorff88/2005	Shin92/2005	Spey88/2005				
Current <i>Ne</i> (c)	196	212	3424	10 082				
95% interval	[135-283]	[152-322]	[1237-4833]	[548-19 045]				

(*FIS*, *FST*, *FIT*) Wright's *F*-statistics; (*Nm*) number of migrants per generation with the private allele method of Barton & Slatkin (1986); (*Hnb*) unbiased expected heterozygosity; (*Hobs*) observed heterozygosity; (*Nall*) mean numbers of alleles per locus; (*Tf*) estimation of time (in years) since population started to decline with the interval at 95%; (*Ancestral Ne*) estimation of median ancestral population size with the interval at 95%; (*Current Ne*) estimation of median current population size with the interval at 95%; (a) MSVAR method; (b) DIYABC method; (c) TM3 method.

0 and 6). We assumed the Single Step Mutation Model and a mutation rate of  $3 \times 10^{-4}$  and  $9 \times 10^{-4}$ .

(b) *Structure and gene flow*

Geographic and temporal genetic differentiation were estimated with  $R_{ST}$  (Slatkin, 1995) by GenAlEx software and with  $F_{ST}$  according to Weir & Cockerham (1984) by Fstat 2.9.4 software (Goudet, 1995). Statistical significance was calculated from 10 000 permutations. Genetic distances (Nei, 1978) were derived from allele frequencies. Ninety-five percent confidence intervals of the mean  $F_{ST}$ ,  $F_{IS}$  and  $F_{IT}$  estimates were obtained by bootstrapping (1000 replicates) over loci by GENETIX software. Partition of genetic variance by Euclidean genetic distances among and within populations was calculated according to analysis of molecular variance (AMOVA) resulting in  $R_{ST}$  estimations by 9999 permutations. The repartition of individual populations was graphically represented using factorial corresponding analysis (FCA) with GENETIX software.

Average effective numbers of migrants per generation ( $Nm$ ) were derived using the four salmon populations sampled in 2005 and 1988/1992 by applying the private allele method (Barton & Slatkin, 1986) using GENEPOP v3.4 (Raymond & Rousset, 1995), and from  $F_{ST}$  according to the relationship  $Nm = (1 - F_{ST}) / (4 * F_{ST})$  using GENETIX v 4.03. To give a better indication of the migration rate per population, we used IM software (Nielsen & Wakeley, 2001; Hey & Nielsen, 2004) between populations within each country (Oir-Scorff and Shin-Spey), with 500 000 burning steps and 5 000 000 records period. Genetic assignment was performed using a Bayesian method (Rannala & Mountain, 1997) as in GENECLASS 2.0 (Piry *et al.*, 2004), and using the STRUCTURE software (Pritchard *et al.*, 2000) with 20 000 iterations and a burn-in of 5000.

(c) *Evolutionary scenarios*

Past evolution of populations was analysed using four algorithms and the 37 available markers using LAMARC (version 2.1, <http://evolution.gs.washington.edu/lamarc>, Kuhner, 2006), MSVAR (Beaumont, 1999), BOTTLENECK (Piry *et al.*, 1999) and DIYABC (Cornuet *et al.*, 2008). The average mutation rate over loci ( $u$ ), ancestral time ( $Tf$ ) and ancestral  $Ne$  were estimated using MSVAR, running  $2 \times 10^7$  steps per Markov chain Monte Carlo (MCMC) chain and DIYABC with 500 000 simulations. LAMARC was run only with the set of 37 markers to derive the growth rate ( $g$ ) and provide the global evolution of populations (growing if  $g$  is positive, shrinking if  $g$  is negative). The  $g$  was determined by running 10 initial chains of 1000 steps each,

discarding the first 500 and for 2 final chains of 10 000 steps discarding the first 1000. The recent effective size decline was tested using BOTTLENECK with the Two Phase Step Mutation model with 10 000 iterations. Evolutionary scenarios were compared using DIYABC software to derive the most likely outcomes.

(v) *Estimators of current effective population sizes*

(a) *Current  $Ne$  from short-term temporal analysis*

In each population, a temporal method was used to estimate the harmonic mean of current  $Ne$  between samples. The method was based on short-term allelic frequency changes between sampling periods (Oir, Scorff and Spey 1988 versus 2005; Shin 1992 versus 2005). Four generations were assumed to have elapsed between the 1988 and 2005 samples, and three generations between the 1992 and 2005 samples. The TM3 temporal method (Berthier *et al.*, 2002) was chosen for its higher efficiency compared to the classical  $F$ -statistic estimator (Nei & Tajima, 1981; Waples, 1989), because it shows a narrower credible interval (CI) and greater accuracy when genetic drift is strong (Berthier *et al.*, 2002). The method was run using the eight subsets of markers and all the 37 markers with 500 000 iterations. Bayesian priors on the maximum size were based on demographic population data.

(b) *Current  $Ne$  from long-term analysis*

Two methods, DIYABC and MSVAR, were used to assess the current  $Ne$  from the distribution of alleles and coalescence processes. Long-term current  $Ne$  refers to the analysis with these two methods. For each population, both samples (1988/1992 and 2005) were analysed separately. DIYABC is an Approximate Bayesian Computation (ABC) that simulates data sets from priors, and then only data sets that are closest to the observed set are retained. The parameter values used to simulate these selected data provide an approximate posterior distribution by local linear regression. A second difference is the simulation of coalescence. Traditionally, it has been assumed that  $Ne$  is large enough to discount the probability that two or more coalescence events occurred in the same generation. However, population size can be very small and multiple coalescence events can occur in the same generation. The alternative is to reconstruct the lineages one generation at a time. DIYABC swaps between these two algorithms according to  $Ne$ . The last algorithm is taken when the effective size is very small or when the first algorithm overestimates the number of lineages (Cornuet *et al.*, 2008).

MSVAR version 1.3 was run assuming a linear trend of population size, and setting as prior a starting

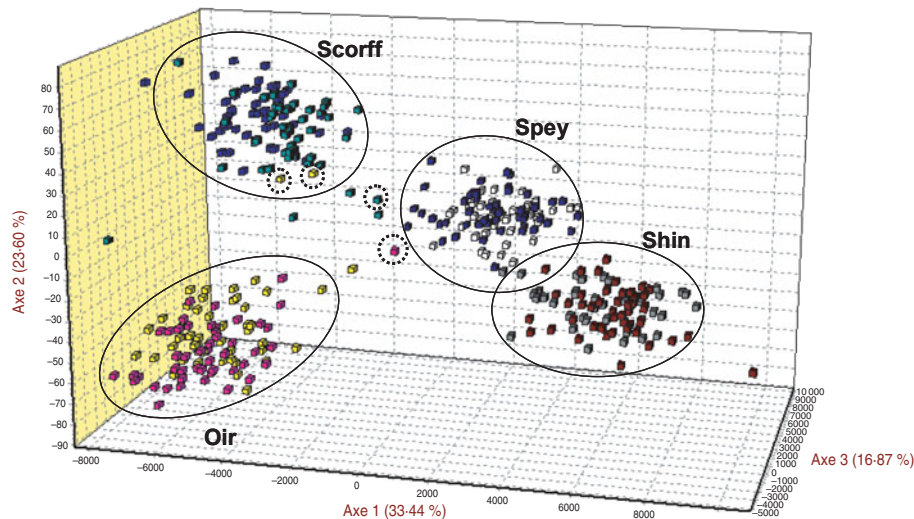


Fig. 2. Factorial correspondence analysis of 367 wild salmon from the four populations (GENETIX software version 4.05.2, Belkhir *et al.*, 1998). The doped circles represent the migrants detected by GENECLASS 2.0 (Piry *et al.*, 2004).

value compatible with demographic data. We assumed linear evolution because the average changes in population over long periods are more likely to be linear than exponential (Beaumont, 1999). As for TM3, a Bayesian prior was set on the maximum size for DIYABC. Also, an  $N_e$  of 50 000 (which could represent a maximum ancestral size) was set for all four populations in order to evaluate the influence of prior information on posterior distributions.

### 3. Results

#### (i) Historical processes

##### (a) Genetic diversity

Inbreeding coefficient ( $F_{IS}$ ) quantifies the difference between observed and expected heterozygosity and evaluates the reduction of heterozygosity due to non-panmictic reproduction (Weir & Cockerham, 1984).  $F_{IS}$  was very low in all populations (Table 2) with the highest values for the Shin in 2005 and the Oir in 1988. Average numbers of alleles per locus ( $N_{all}$ ) and observed heterozygosity ( $H_{obs}$ ) across the four Atlantic salmon populations for the two sampling periods (1988/1992 and 2005), ranged from 11 (Oir in 1988 and Scorff in 1988) to 15 (Spey in 1988) and from 0.76 (Scorff) to 0.82 (Spey), respectively (Table 2).  $N_{all}$  was stable for Spey, increased for the Oir and Scorff and decreased for the Shin (Table 2). Unbiased expected heterozygosity ( $H_{nb}$ ) were very close to  $H_{obs}$  in all samples.  $H_{obs}$  levels across the 37 microsatellites markers were very high with a maximum at 0.93 (SsaD144, BHMS331 and BHMS230) and few loci showed a significant deviation from Hardy–Weinberg equilibrium (Nikolic *et al.*, 2009).

Our calculation of the expected evolution of diversity under variable population sizes by DemoDivMs

(section 2.4.1) was qualitatively explained by a recent bottleneck 25–100 generations ago, assuming an effective size of 2500–5000 before the bottleneck. Even though results on differentiation and assignment results showed weak migration, we checked its impact on smaller populations. From IM values (0.004 and 0.005), the mean number  $Nm$  of immigrants in a population of size 200 is about 1. Accounting for possible migration was modelled by assuming immigrants come from a large population in equilibrium that is representative of the ancestral population of Atlantic salmon. The analysis indicates that the high observed heterozygosity cannot be maintained in the smallest populations ( $N_e = 200$  or 1000) by gene flow alone, except if it would take place at a higher rate ( $Nm > 5$ ) than observed. Assuming a recent bottleneck seems necessary to account for the observed diversity, even if the joint effects of genetic drift and gene flow are required, in such a way that the higher the migration rate is, the more ancient the bottleneck must be assumed. This suggested that this population was derived from a larger one that underwent a recent bottleneck.

##### (b) Structure and gene flow

Temporal genetic differentiations ( $R_{ST}$  and  $F_{ST}$ ) were significant ( $P < 0.05$ ) only for the Scorff which had the highest values (0.022 and 0.009), indicating that this population underwent a significant genetic change during this period (Appendix A). The lowest values of  $R_{ST}$  and  $F_{ST}$  between samples were found for the Shin (0 and 0.001) (Appendix A). Similar trends were observed with  $Nei$  distance. We have not presented the pairwise  $R_{ST}$  in Appendix A because they were of similar magnitude to the  $F_{ST}$ . In spite of the low  $F_{ST}$  values, the populations studied were well



Table 3. Pairwise numbers of migrant ( $N_m$ ) from  $F_{ST}$  of Weir & Cockerham (1984) by GENETIX software on the superior half matrix and number of migrants using private alleles (Barton & Slatkin, 1986) by GENEPOP software on the inferior half matrix

Sampling Years	Years Populations	2005 Oir	1988 Oir	2005 Scorff	1988 Scorff	2005 Shin	1992 Shin	2005 Spey	1988 Spey
2005	Oir	0		6.51	6.10	3.38	3.39	6.43	6.64
1988	Oir		0	4.93	4.68	2.67	2.62	5.08	5.26
2005	Scorff	2.68	1.83	0		2.96	2.99	5.38	6.29
1988	Scorff	2.46	1.67		0	3.01	3.23	6.47	8.07
2005	Shin	1.97	1.25	1.81	1.20	0		7.08	6.01
1992	Shin	2.21	1.40	1.93	1.92		0	6.71	5.87
2005	Spey	2.79	2.17	2.81	2.09	2.91	3.62	0	
1988	Spey	2.83	2.58	3.11	3.10	4.00	3.79		0

differentiated by the factorial correspondence analysis (Fig. 2) and using the unsupervised Bayesian approach in STRUCTURE. The set of 367 individuals was clearly split into four clusters indicating the presence of mixed stock introduction in the Oir 2005 sample and non-native introduction in the 1988 Scorff sample, which was absent in the 2005 sample.

An AMOVA (based on Euclidean  $R_{ST}$  distances) across the four populations for the two sampling times revealed that the largest proportion of variation (90%) was found within populations. AMOVA between samples from the same populations showed the highest variance (2%) for the Scorff and lowest (0%) for the Shin, while the Oir and Spey were 1%.

The pairwise  $N_m$  estimated from the private allele method of Barton & Slatkin (1986) were lower than the pairwise from  $F_{ST}$  of Weir & Cockerham (1984) (Table 3) with the highest (8.07) between Spey 1988 and Scorff 1988. The migration rates by IM were at 0.0030 and 0.0038 for Oir, 0.0047 and 0.038 for Scorff, 0.0026 and 0.0025 for Shin and 0.005 and 0.0067 for Spey, respectively, in 2005 and past samples. Four individuals were assigned outside their population of origin with GENECLASS 2.0: two from the Oir in 2005 were assigned to the Scorff 2005 population (probability ( $P$ ) were equal to 0 and 0.002), one individual from the Oir in 1988 was assigned to the Spey 1988 population ( $P=0$ ), and one individual from the Scorff in 1988 was assigned to the Spey 1988 population ( $P=0.001$ ). These four migrants were all males and can be seen in the factorial correspondence analysis (Fig. 2, dotted circles).

### (c) Evolutionary scenarios

Average mutation rates for the 37 microsatellites were estimated at  $3 \times 10^{-4}$  by MSVAR and slightly higher by DIYABC at  $9 \times 10^{-4}$ . The first result is concordant with previous studies on *Salmo salar* by O'Reilly *et al.* (1998). A negative posterior distribution of  $\log_{10}(r)$  values (with  $r$  equal  $Current\ Ne/Ancestral\ Ne$ ) was

revealed by MSVAR 0.4 and a negative growth rate ( $g$ ) for overall populations by Lamarc software, which is consistent with a decline in effective sizes (Table 2). The various tests of heterozygosity deficit proposed in BOTTLENECK (Sign test and Wilcoxon test, TPM model) suggested a significant departure from constant population size in all populations.

Given an assumed generation time of 3–4 years, the estimated time since decline ( $T_f$ ) ranged from 8000 to 20 000 years ago according to MSVAR 1.3, and around 13 000 years ago according to DIYABC. This was consistent for the four populations and for both sampling times (Table 2). The ancestral  $N_e$  values calculated by MSVAR 1.2 were approximately the same for all populations (approximately 50 000 individuals) suggesting a common ancestor. DIYABC revealed a smaller ancestral population size of 13 000–30 000 (Table 2) and two equally likely scenarios in which the populations were separated from a common ancestor by a star or cascade process from the southern (Scorff) to the northern population (Shin).

### (ii) Estimators of effective population sizes

#### (a) Markers' number and heterozygosity variation

An increase in the markers' number and polymorphism led to a decrease in the variance of posterior distribution of current  $N_e$  estimates in almost all methods. While clearly shown for all populations with the MSVAR method, this phenomenon was less visible using TM3 and DIYABC methods for larger populations, and was absent for the Spey, the largest (*c* and *d*, Figs 4 and 6).

Using MSVAR (Fig. 3) the same trend was observed for the four rivers, showing an increase of mean values and a decrease of variances when the number and polymorphism of markers was increased. The TM3 (Fig. 4) method provided unexpected profiles with the subsets of five and ten loci (5mH–,

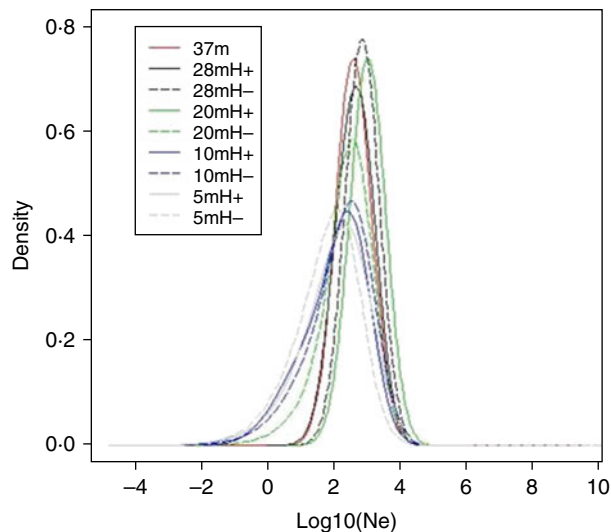


Fig. 3. Posterior distribution of current effective population size estimators ( $\log_{10}$  scale), for river Oir (2005 sampling), according to MSVAR. Curves refer to the eight subsets of markers ( $m$ ) according to their lower heterozygosity ( $H^-$ ) in dotted line and their higher heterozygosity ( $H^+$ ) in solid line. The red curve refers to the full set of 37 markers. Similar results were obtained with other samples and rivers.

5  $mH^+$ , 10  $mH^-$ , 10  $mH^+$ ) for Oir and the subsets of five (5  $mH^-$ ) for Scorff, generated large estimates of population size (Fig. 4A and B). DIY ABC also generated large estimates for the Oir population size with the subset of low number of markers (Fig. 5A).

(b) *Estimation when setting priors on the mean current  $N_e$*

Regarding the CIs and variance, short-term current  $N_e$  estimates using TM3 were more accurate for smaller populations ( $\leq 1000$ ) (Fig. 6A and B) than for larger populations ( $\geq 3000$ ) (Fig. 6C and D). Long-term current  $N_e$  estimates using both MSVAR and DIYABC were very consistent between sampling years, but their variances and 95% confidence intervals remained high for all samples. For example, the Spey estimate derived by MSVAR was 1076–50 364. Overall, current  $N_e$  for the four populations in 2005 were estimated to be 383 (Oir), 689 (Scorff), 304 (Shin) and 7344 (Spey) using MSVAR, and 100 (Oir), 1174 (Scorff), 1842 (Shin) and 9417 (Spey) using DIYABC (Table 2). These values are nearer to the census population sizes ( $N$ ) for the smaller populations (Oir and Scorff) than for the larger populations (Shin and Spey). However, if the variance around the median of  $N_e$  was considered the estimation of the effective size was near census size for all populations.

Considering the differences in effective size values between methods, and taking into account higher ratios of  $N_e/N$  for smaller salmonid fish (Palstra & Ruzzante, 2008), TM3 underestimated the Scorff

population size (Fig. 6B) and MSVAR underestimated the Shin population size (Fig. 6C).

(c) *Estimation when setting priors on the maximum current  $N_e$*

Changing the prior maximum size to 50 000 had an effect on point estimates of current  $N_e$  so that for all populations when running DIYABC, mean values were increased and the upper 95% limits of confidence intervals approached this maximum. On the contrary no effect was seen with TM3 for the smaller populations (Oir and Scorff). Setting the starting value at 50 000 using MSVAR had no appreciable effect on point estimates of current  $N_e$  in any population.

#### 4. Discussion

Measurements of effective population size ( $N_e$ ) are of importance in conservation and management because they give an overview on the evolution of genetic diversity. Effective size determines the rate at which genetic diversity is lost in the population by genetic drift (Franklin, 1980). The most genetically diverse populations are assumed to be 'fitter' (Ligoxygakis, 2001). A genetically viable population possesses the evolutionary legacy of the species and the genetic variability on which future evolutionary potential depends (Dodson *et al.*, 1998). Overfishing, blockage of migratory routes by hydroelectric dams and destruction of spawning habitat have severely depleted many wild salmon stocks (Waples, 1990). Humans have altered natural ecosystems for many thousands of years, but the magnitude and rate of these changes have increased dramatically since the industrial revolution. Over recent decades, Atlantic salmon populations have declined or have been extirpated in many parts of its ancestral range (Parrish *et al.*, 1998; Knockaert, 2006) and measurements of  $N_e$  are few. Genetic estimates of effective size have never been conducted on the Oir, Scorff or Spey populations. Consuegra *et al.* (2005) investigated  $N_e$  in the Shin, but based their estimate on four markers and did not provide estimates. These populations have management regimes and legislation in place to minimize the risks of further declines, but so far effective population size has not been implemented as a tool in their management strategies. Furthermore, the diversity of the rivers' habitats makes them interesting case studies of the effects of contrasting environments on their genetic structure (Whitlock & McCauley, 1999; Leberg, 2005; Wang, 2005).

Before discussing the results on estimations of  $N_e$  obtained about the four salmon populations and their implication for conservation and management, it is useful to recall some features of these populations and to present the results on historical processes obtained

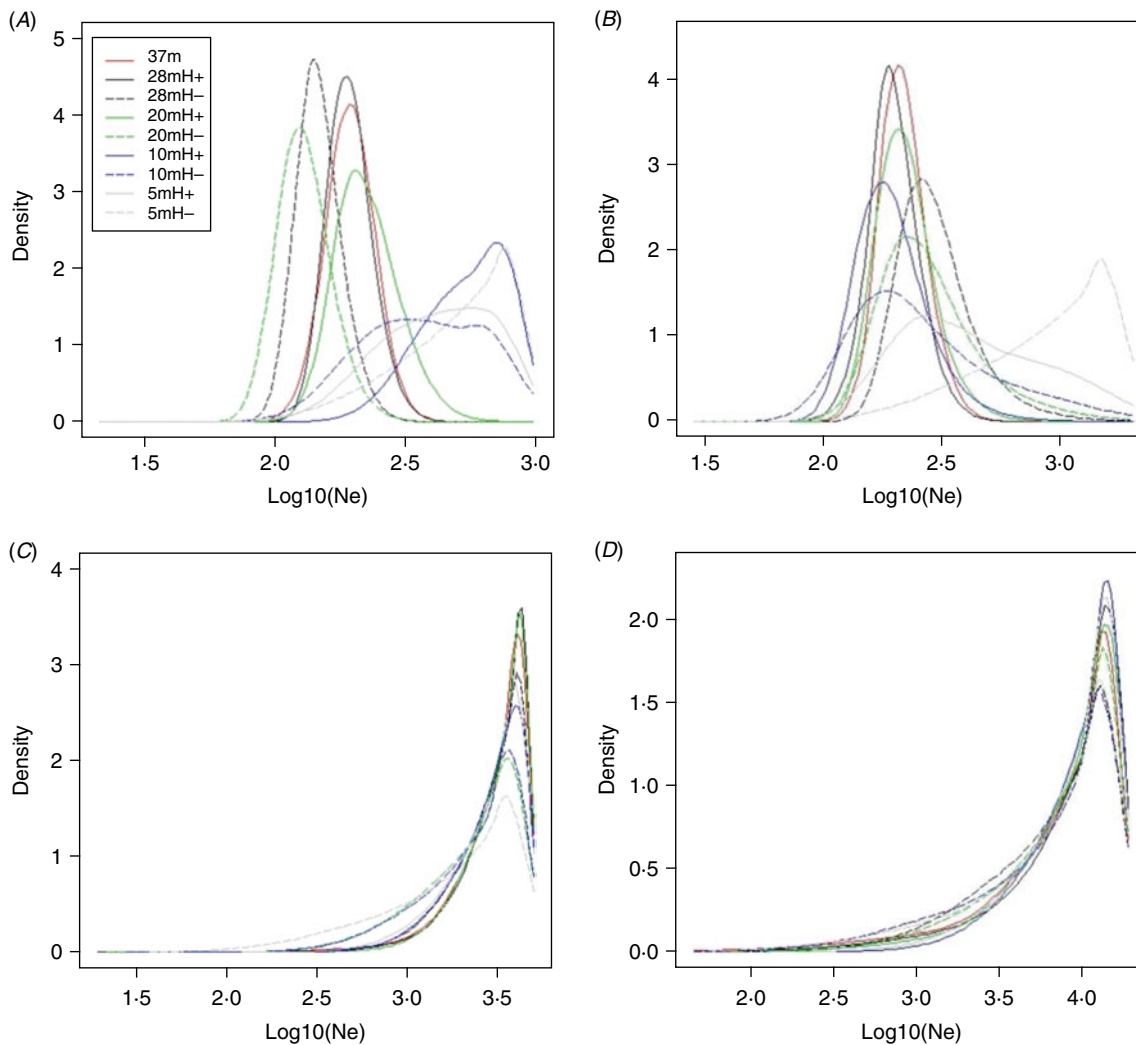


Fig. 4. Posterior distribution of current effective population size estimators ( $\log_{10}$  scale) according to TM3, for the populations Oir (A), Scorff (B), Shin (C) and Spey (D). Curves refer to the eight subsets of markers (m) according to their lower heterozygosity (H $-$ ) in dotted line and their higher heterozygosity (H $+$ ) in solid line. The red curve refers to the full set of 37 markers. Similar results were obtained with other samples and rivers. See the legend in the corner of (A).

by genetic analysis. Atlantic salmon populations are highly structured with strong genetic differentiation between regions and among rivers (Ståhl, 1987; Nielsen *et al.*, 1999; King *et al.*, 2001), with weaker differentiation within river systems (Nielsen *et al.*, 1999; Garant *et al.*, 2000; Primmer *et al.*, 2006; Vaha *et al.*, 2007, 2008). This differentiation is partly explained by high fidelity to natal rivers (Stabell, 1984). Geographic genetic differences observed reflect the effects of underlying evolutionary forces, such as drift. Temporal (between samples from the same population) genetic differentiation ( $R_{ST}$ ,  $F_{ST}$  and  $N_e$  distances) was significantly higher for the Scorff and lower for the Shin. These differences are explained in the Scorff by the introduction of non-native juveniles of unknown Scottish origin in the 1980s. If the Spey was the source, this last point could explain the highest values of numbers of migrants ( $Nm$ ) between the Spey and Scorff in 1988. The evidence of reduced

genetic variability in the Shin and of low temporal difference is probably from the result of long-term artificial stock enhancement program using fish of native (Shin) origin employed to mitigate the blocking of freshwater habitat by hydroelectricity dams in the 1950s. This native stock enhancement may explain the low temporal genetic distance. An effect of stock enhancement in the Oir is suggested by the STRUCTURE analysis. Significant differentiation among the Scorff samples, but not among the Oir samples can be explained by a brief and genetically distant gene flow into the Scorff from a Scottish source, whereas gene flow into the Oir has been more continuous and from a genetically similar French source, the Sélune. For the Spey, gene flow can be negligible given its large  $N_e$ . The clear differentiation between rivers and the high rate of assignment of individuals to their river of origin confirm that the populations maintained their originality.

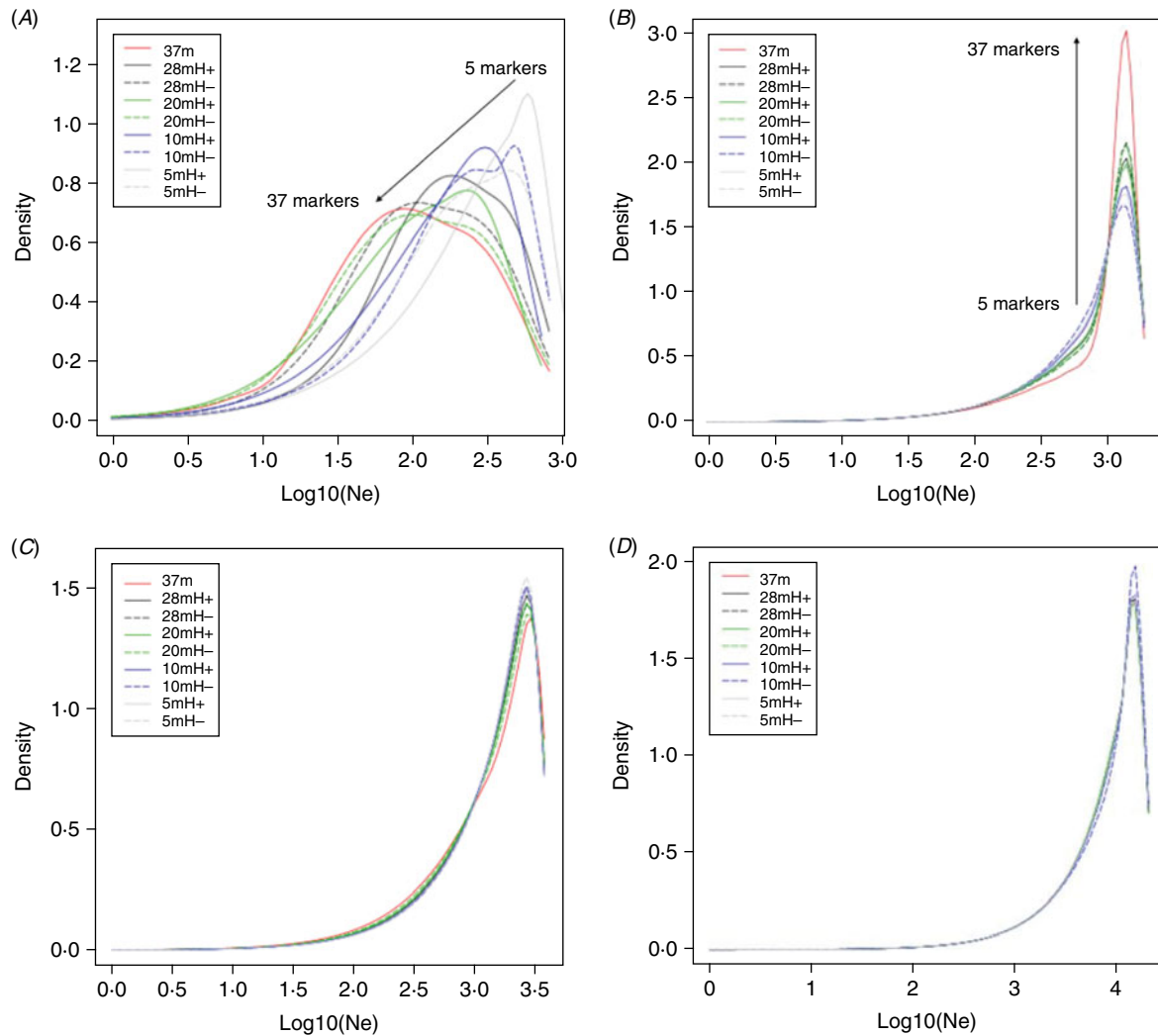


Fig. 5. Posterior distributions of current effective population size estimators ( $\log_{10}$  scale) according to the DIYABC method for the populations Oir (A), Scorff (B), Shin (C) and Spey (D) in 2005. Similar results have been observed in the past samples (1988/1992). Curves refer to the eight subsets of markers (m) according to their lower heterozygosity (H<sup>-</sup>) in dotted line and their higher heterozygosity (H<sup>+</sup>) in solid line. The red curve refers to the full set of 37 markers. Similar results were obtained with other samples and rivers. See the legend in the left corner.

The four populations also seem subject to a global decrease in wild salmon stocks. The genetic analysis revealed a negative growth rate and a bottleneck for all populations. Whereas allelic richness seemed to decrease in the Shin, increase in the Oir and Scorff and be stable in the Spey, genetic analysis on heterozygosity demonstrated a high and stable variability over time with no inbreeding for all samples. These results are concordant with previous studies that revealed high genetic diversity on wild populations of Atlantic salmon from Ireland and Norway with *Nall* and *Hobs* of 17.8 and 0.70, respectively, with 17 microsatellites (Norris *et al.*, 1999) and from Scotland for the Shin population (sampled in 1987–1990) with a *Hobs* of 0.71 with 12 microsatellites (King *et al.*, 2001) and a bit lower with 4 microsatellites (Consuegra *et al.*, 2005). The rate of loss of genetic diversity via genetic drift is higher in small populations and, in the

absence of migration, this rate is expected to rise as the effective size decreases (Frankham *et al.*, 2002). Taking account of the low gene flow detected by IM for the smaller populations (0.003–0.004 for Oir and 0.004–0.005 for Scorff), we tested the bottleneck hypothesis by our new model DemoDivMs. Regarding the high genetic diversity (75–81%) detected in smaller populations (Oir and Scorff), we used it to evaluate the period when the bottleneck occurred. Our results suggest a common ancestor with a large size (10 000–50 000 effective individuals) dating back to the last glaciations, representing 2000 to 4000 generations. We assumed a scenario with an ancestral large population at great genetic diversity in equilibrium that splits into several populations. Trials with our method (using a mutation rate of  $3 \times 10^{-4}$  to  $9 \times 10^{-4}$  and the Single Step Mutation model for microsatellites) suggested that observed small



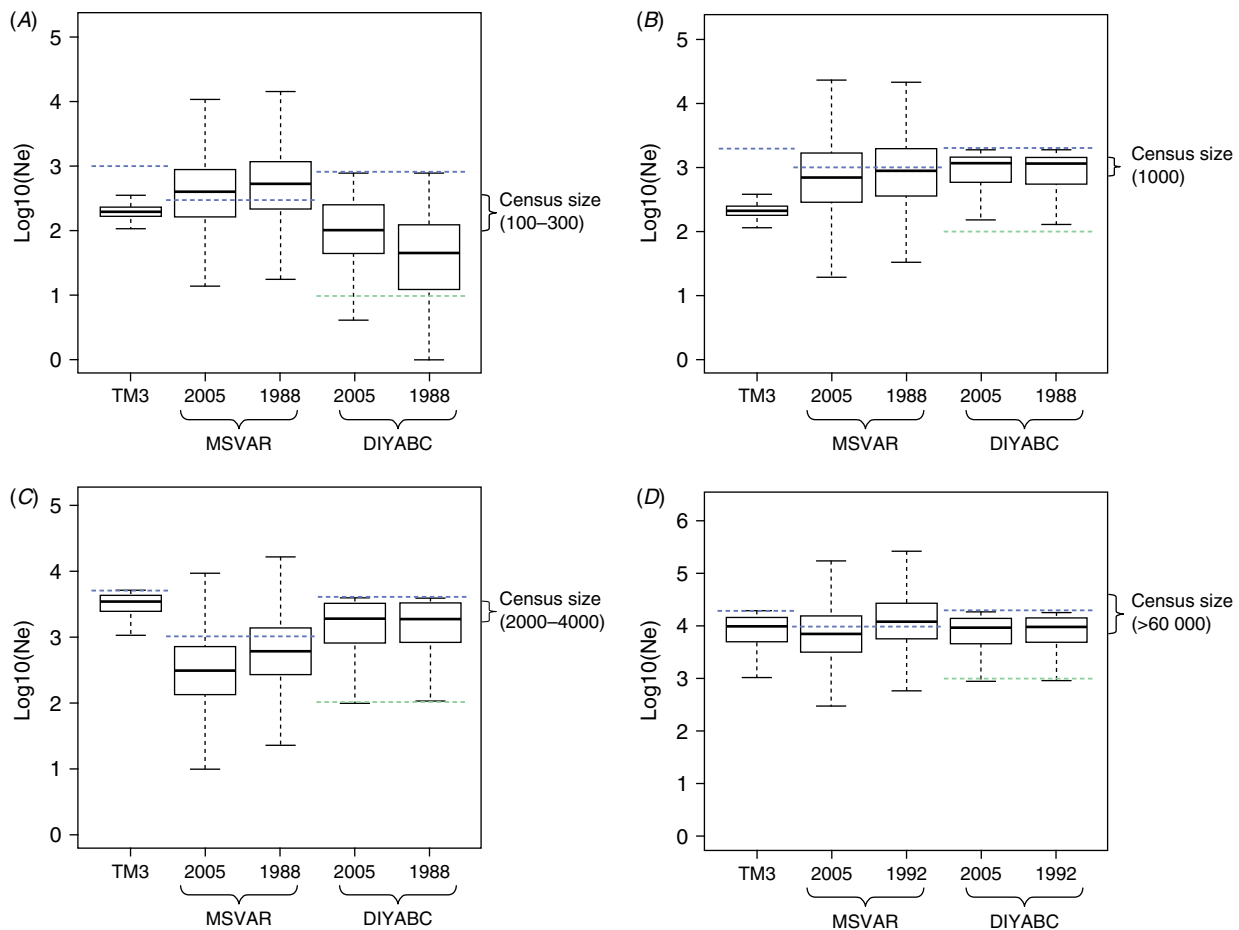


Fig. 6. Summary statistics of effective population size estimators ( $\log_{10}$  scale) from the different methods (TM3, MSVAR and DIYABC) for the four populations: Oir (A), Scorff (B), Shin (C) and Spey (D). Census size is mentioned on the right y-axis. Median (horizontal black line), variance (box) and confidence interval at 95% (dotted black line on both sides of boxes) are given for each method. Priors are plotted as follows: Blue dotted lines: highest current  $N_e$  value for TM3 and DIYABC, starting value for MSVAR. Green dotted lines: lowest current  $N_e$  value for DIYABC.

populations (200–1000 effective size) could show the high observed heterozygosity if they derived from larger populations of size 2500–5000 that experienced a recent bottleneck 25–100 generations ago (i.e. 100–400 years), the bottleneck being all the more recent as gene flow occurs at a lower rate. Noteworthy, neither gene flow nor bottleneck alone could be considered as a single explanation, both phenomena seem necessary to explain the current status of these populations. However, the hypothesis of precocious parr contributing to the genetic variability in salmonid (García-Vázquez *et al.*, 2001; Jones & Hutchings, 2001, 2002) cannot be neglected in view of the high proportions of fertilization of eggs by parr in Oir and Scorff (Baglinière & Maisse, 1985; Baglinière *et al.*, 1993).

Given the importance of identifying  $N_e$  for populations in decline, we compared recent coalescent methods to assess their sensitivity to the number of polymorphic markers, to the priors used in Bayesian approaches and to the structure of populations. Using different subsets of markers, we showed that methods providing estimates of current  $N_e$  may be more or

less efficient. The efficiency of estimators is improved when the number of markers and their variability are increased. We chose markers which were independent, but violation of the independence among markers' loci should not affect the accuracy of estimation of current  $N_e$  (Wang & Whitlock, 2003). The CIs and variance decrease by increasing the number and variability of markers with the three methods. However, this behaviour depended on populations, except for MSVAR. MSVAR results were improved on all populations by increasing the number of polymorphic markers (Fig. 4). The difference of accuracy between sets of markers (from low number and variability markers to high number and variability markers) was less accentuated for larger populations with TM3 and DIYABC (Figs 5 and 6). These methods presented more complex behaviour. They overestimate current  $N_e$  in smaller populations (Oir and Scorff) when using few markers. It appears that the used marker sets corresponded to uneven distributions of allelic frequencies. The sets corresponded to five and ten markers at low (H–) and high (H+) heterozygosity.

The sets at low heterozygosity possess alleles at high frequencies around 60–80%. The sets at high heterozygosity (H+) did not show alleles with high frequency (>50%), but show uneven distributions between samples from the same population. The absence of alleles at intermediate frequency and high frequency of a single allele may lead to an overestimation of current  $N_e$  (Waples, 1989). MSVAR did not seem to be impacted by the uneven distribution of allele frequencies in the smaller populations.

The incidence of Bayesian priors was weak using MSVAR, but quite strong using DIYABC for all populations, and TM3 for the largest populations (Spey and Shin). Hence, at least for these practical methods, Bayesian priors should be applied using reliable field data to avoid the derivation of inaccurate predictions.

Results from the set of 37 microsatellite markers demonstrated different sensitivity to the size of populations. The CIs and variance were greater for the larger populations (Shin and Spey) using TM3 (short-term  $N_e$ ), large and homogeneous on all populations using MSVAR (long-term  $N_e$ ) and heterogeneous using DIYABC (long-term  $N_e$ ). The long-term current  $N_e$  estimates were strongly correlated between samples from the same population but CIs were larger for MSVAR than DIYABC except for the smallest population (Oir). Additional analysis using the classical temporal approach based on  $F$ -statistics (see equation (9) in Waples, 1989) from NeEstimator v 1.3 (Peel *et al.*, 2004) demonstrated a high uncertainty in larger populations. The imprecision of temporal methods in large populations seems to occur with classical and coalescent models due to weak genetic drift. Berthier *et al.* (2002) compared TM3 with the classical  $F$ -statistical-based  $N_e$  estimator and showed narrower CIs and greater accuracy with TM3 when genetic drift was strong, but there was no improvement with weak genetic drift. Large CIs and variance observed in larger populations with TM3 can be explained by the ratio of sample size ( $S$ ) to  $N_e$ . Too small a sample size may lead to insufficient extraction of genetic information in large populations. As pointed out by Nei & Tajima (1981), precision increases with the ratio  $S/N_e$ . Waples (1989) demonstrated that a low ratio would translate into larger CIs, which means that populations with small  $N_e$  are most effectively studied. Furthermore, the effects of genetic drift in large populations might be swamped by sampling error using the temporal method, and if  $tS/N_e > \sqrt{2}$ , with  $t$  the sampling interval, increasing the number of loci (or alleles) has a greater effect on precision than increasing  $t$  or  $S$  (Waples, 1989). This means that given the effective size to be estimated, minimum numbers of generations and sample sizes are required to achieve reasonable precision, unless a large number of loci can be surveyed.

The long-term methods (MSVAR and DIYABC) assumed that selection and migration were unimportant in changing population allelic frequencies relative to genetic drift. For the short-term current  $N_e$  estimates (TM3), all systematic forces (mutation, selection and migration) are assumed to be absent. The assumption of no mutation is reasonable because our sampling periods of 3–4 generations are sufficiently short to discount the effects of mutations. It may be reasonable to neglect the effects of selection because selection on most markers is unlikely to be strong enough to cause substantial changes in their frequencies (Wang & Whitlock, 2003). Moreover, the loci used in our study have been tested to detect potential selection (Nikolic *et al.*, 2009) and can be considered a reliable panel. In the context of this study, even if stock enhancement programs (based on non-native or native source stocks in Scorff and Shin, respectively) were associated with low values of  $Nm$ , they seem affected the temporal genetic differentiation, with stronger  $R_{ST}$ ,  $F_{ST}$  and  $N_e$  distances in the Scorff and lower in the Shin. Since temporal estimators (TM3) interpreted allele frequency changes as due to genetic drift, they led to decreased estimated effective size in the Scorff, yielding a lower value than the long-term  $N_e$ . The effective size estimates derived from short-term methods on the Shin was higher than with the long-term methods because native reintroductions helped reduce the temporal genetic differences. Using model simulations of infinite source populations sending a constant supply of migrants with constant allele frequency Wang & Whitlock (2003) revealed that migration changes allele frequency more in the short term and less in the long term, leading to under and overestimation of current  $N_e$ , respectively. Regarding our results, introduction seems to have modified effective size estimators in different ways depending on the source. By comparing the effective size values between methods and by taking into account the results of higher ratio  $N_e/N$  on smaller salmonid fish by Palstra & Ruzzante (2008), TM3 seemed to underestimate the Scorff population size and MSVAR to underestimate the Shin population size. A native introduction, or from a genetically close source, could lead to underestimations of current  $N_e$  with long-term methods. A pulse of genetically different migrants coming between samples will underestimate current  $N_e$  with short-term methods. Even with small values of  $Nm$ , the temporal methods seemed more sensitive than expected to weak gene flow. In cases when artificial or natural gene flow is suspected, we suggest that both short-term and long-term estimates of population size should be evaluated.

In spite of these challenges, the current  $N_e$  estimates presented here are the first to be proposed for the French and the larger Scottish (Spey) studied

populations, and provide potentially useful tool for their management. Efforts to establish lower population size thresholds have suggested current  $N_e = 500$  to maintain sufficient evolutionary potential (Franklin, 1980; Lande, 1988; Franklin & Frankham, 1998; Lynch & Lande, 1998). Estimating current  $N_e$  is thus an important tool in the assessment of genetic variability (Mace & Lande, 1991). For the Oir, current  $N_e$  estimates of 383 and 100, compared with a 2005 estimated adult population of 130, suggesting that the river's IUCN (1994) conservation status of 'vulnerable' is still justified. The same is true for the Scorff where current  $N_e$  estimates are 689 and 1174, compared to a 2005 adult population of 1000. For the Shin, the 2005 adult population of 3000 exceeds the current  $N_e$  estimates of 304 and 1842. Similarly on the Spey current  $N_e$  estimates of 7344 and 9417 are probably exceeded by the 2005 population. In both cases, the IUCN (1994) status of 'minor concern' remains valid.

The discrepancies between current  $N_e$  estimators across the salmon populations contrasted here raise the question of which methods provide best estimates of  $N_e$ . All coalescent methods are potentially useful, but they may be biased because of their assumptions, particularly regarding migration. A short-term coalescent-based estimate, such as TM3, seems better suited to anadromous salmon populations than long-term estimates if the population is not large, if it does not undergo identified gene flow and if a high number of highly polymorphic markers are available. CIs become wider as current  $N_e$  values become larger. In the case of the long-term coalescence estimator, with a minimum sample of 50 individuals as recommended by Waples (1989), increasing the number of sampled individuals ( $S$ ) may marginally improve the results because the variance of the estimator of  $\theta$  ( $4N_e\mu$ ) decays slowly, at a rate proportional to  $1/\log(S)$  (Deonier *et al.*, 2005). Hence, for larger populations, such as the Shin (up to 4000 individuals) and Spey (up to 60 000 individuals), the 37 markers may not be sufficient to accurately estimate  $N_e$ , and it would make little difference to increase sample sizes. For the short-term estimator, increasing sample sizes may be useful for larger populations. Using simulations, Ovenden *et al.* (2007) estimated that 2000 sampled individuals are required to get a reliable estimate of an effective size of 8000 (Palstra & Ruzzante, 2008). Another way to decrease CIs for short-term estimates (TM3) in the Shin and Spey populations would be to increase the number of generations between samples. Nevertheless, the amount of information is no longer simply proportional to sampling interval  $t$ , and too large a  $t$  may decrease the power of the method (Wang & Whitlock, 2003). Usually, a short  $t$  is recommended (two to four generations) but if the migration rate is low,  $t$  could be larger to increase the estimation

precision (Wang & Whitlock, 2003). Finally, the temporal method (TM3) analysis is not equally efficient for all population sizes and other methods of estimating effective population size are necessary if current  $N_e$  is very large such as in the Shin and Spey, and if the influence of migration cannot be ignored. Regardless, the importance of selecting the appropriate tool for estimating current  $N_e$  is important for the conservation of wild salmon populations.

## 5. Conclusion

Despite their small size and declining status, French populations still show high levels of genetic diversity, similar to those found in the larger populations. Our results and coalescent simulations, where populations are assumed to derive from a common ancestor, suggest that a recent bottleneck rather than gene flow explains the high genetic diversity found in all studied populations. However, we cannot exclude the possibility that precocious parr also contribute to genetic variability. Concerning the sensitivity of methods, the results raise the importance of number and variability of neutral markers, of the Bayesian priors and of the structure of populations. Large populations require higher numbers of markers for long-term coalescence estimators and larger sample sizes for short-term coalescence estimators. Even in the case of low migration, establishing a conservation program should rely on both short-term and long-term estimates of population size.

## 6. Perspectives

Because variability of effective population size is a main factor determining the risk of extinction (Waples, 2002), fluctuating population size is an important consideration for evolution and conservation. Some models have been recently developed (Drummond *et al.*, 2005; Heled & Drummond, 2008) to calculate the fluctuation of current  $N_e$  from the most recent ancestor. Although restricted for the moment to sequence data (mitochondrial or viral DNA) and large intervals of time, extension to microsatellite markers could help provide information on more recent history.

The authors wish to thank N. Jeannot, F. Marchand, G. Evanno and D. Azam (INRA Rennes, France); E. Prévost (INRA St Pée sur Nivelle, France); S. Burns, J. Woods, J. Reid, D. Ferguson and S. Grant (Spey Fishery Board Research Office, Scotland); D. Knox and E. Verspoor (Fisheries Research Services Perthshire, U.K.) and J. Stoddart (Kyle of Sutherland District Salmon Fishery Board, U.K.) for providing support collection of scale and fin samples. Genotyping were done at the Toulouse Genopole Platform (<http://www.genotoul.fr/>) with the help of K. Fève and J. Riquet that we fully thank. This work was financial supported by INRA, Laboratoire de Génétique

Cellulaire (UMR 444), INRA-ENVT, BP 52627, 31326 Castanet Tolosan Cedex, France.

## References

- Anonymous (2000). Revised Guidance Updating Scottish Office Circular No. 6/1995. Nature Conservation: Implementation in Scotland of EC Directives on the Conservation of Natural Habitats and of Wild Flora and Fauna and the Conservation of Wild Birds (The Habitats and Bird Directives), p. 96. Edinburgh: Scottish Executive.
- Anonymous (2001). Report of the Working Group on the Atlantic salmon. ICES CM2001/ACFM 15, p. 199.
- Anonymous (2003). Report of the working group on the Atlantic salmon. ICES CM2001/ACFM, 19, p. 297.
- Baglinière, J. L. (1976). Etude des populations de Saumon atlantique (*Salmo salar* L. 1766) en Bretagne-Basse-Normandie. 1 – Caractéristiques des smolts de la Rivière Ellé. *Annales d'Hydrobiologie* 7, 141–158.
- Baglinière, J. L. (1979). Production de juvéniles de Saumon atlantique (*Salmo salar* L.) sur quatre affluents du Scorff, rivière de Bretagne-Sud. *Annales de Limnologie* 3, 347–366.
- Baglinière, J. L. & Champigneulle, A. (1986). Populations estimates of juvenile Atlantic salmon (*Salmo salar*) as indices of smolt production in the River Scorff, Brittany. *Journal of Fish Biology* 29, 467–482.
- Baglinière, J. L. & Maisse, G. (1985). Precocious maturation and smoltification in wild Atlantic salmon in the armoricain Massif, France. *Aquaculture* 45, 249–263.
- Baglinière, J. L., Maisse, G. & Nihouarn, A. (1993). Comparison of two methods of estimating Atlantic salmon, *Salmo salar*, wild smolt production. In *Production of Juvenile Atlantic salmon, Salmo salar, in Natural Waters* (ed. R. J. Gibson & R. E. Cutting). Canadian Special Publication in Fisheries and Aquatic Sciences 118, 189–201.
- Baglinière, J. L., Marchand, F. & Vauclin, V. (2005). Interannual changes in recruitment of Atlantic salmon (*Salmo salar*) population in the River Oir (Lower Normandy, France): relationship with spawners and in-stream habitat. *Journal of Marine Science* 62, 695–707.
- Barton, N. H. & Slatkin, M. (1986). A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409–415.
- Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics* 153, 2013–2029.
- Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. & Bonhomme, F. (1998). GENETIX, Logiciel sous Windows™ pour la Génétique des Populations, Laboratoire Génome et Populations, CNRS UPR 9060. Montpellier, France: Université de Montpellier II.
- Berthier, P., Beaumont, M. A., Cornuet, J. M. & Luikart, G. (2002). Likelihood-based estimate of the effective population size using temporal changes in allele frequency: a genealogical approach. *Genetics* 160, 741–751.
- Butler, J. R. A. (2004). Moray Firth Seal Management Plan: A Pilot Project for Managing Interactions Between Seals and Salmon in Scotland, pp. 64. Aberlour, Moray-shire: Spey Fishery Board.
- Butler, J. R. A., Middlemas, S. J., McKelvey, S. A., McMyn, I., Leyshon, B., Walker, I., Thompson, P. M., Boyd, I. L., Duck, C., Armstrong, J. D., Graham, I. M. & Baxter, J. M. (2008). The Moray Firth Seal Management Plan: an adaptive framework for balancing the conservation of seals, salmon, fisheries and wildlife tourism in the UK. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18, 1025–1038.
- Butler, J. R. A., Radford, A., Riddington, G. & Laughton, R. (2009). Evaluating an ecosystem service provided by Atlantic salmon and other fish species in the River Spey, Scotland: the economic impact of recreational rod fisheries. *Fisheries Research* 96, 259–266.
- Caron, F. & Fontaine, P. M. (2003). L'état des stocks de Saumon atlantique au Québec en 2002. Société de la faune et des parcs du Québec. Direction sur la Faune, pp. 48.
- Chakraborty, R. & Kimmel, M. (1999). Statistics of microsatellite loci: estimation of mutation rate and pattern of population expansion. In *Microsatellites: Evolution and Applications* (ed. D. B. Goldstein & C. Schlötterer), pp. 139–150. Oxford, UK: Oxford University Press.
- Chevalet, C. & Nikolic, N. (2010). Distribution of coalescence times and distances between microsatellite alleles with changing effective population size. *Theoretical Population Biology*, in press.
- Consuegra, S., Verspoor, E., Knox, D. & De Leaniz, C. G. (2005). Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. *Conservation Genetics* 6, 823–842.
- Cornuet, J.-M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J.-M., Balding, D. J., Guillemaud, T. & Estoup, A. (2008). Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24, 2713–2719.
- Deonier, R. C., Waterman, M. S. & Tavaré, S. (2005). Computational Genome Analysis: An Introduction, pp. 403–404. New York: Springer.
- Dodson, J. J., Gibson, R. J., Cunjak, R. A., Friedland, K. D., Garcia de Leaniz, C., Gross, M. R., Newbury, R., Nielsen, J. L., Power, M. E. & Roy, S. (1998). Elements in the development of conservation plans for Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences* 55, 312–323.
- Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22, 1185–1192.
- Estoup, A., Largiadèr, C.-R., Perrot, E. & Chourrout, D. (1996). One-tube rapid DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology* 5, 295–298.
- Feldman, M. W., Kumm, J. & Pritchard, J. (1999). Mutation and migration in models of microsatellite evolution. In *Microsatellites: Evolution and Applications* (ed. D. B. Goldstein & C. Schlötterer), pp. 98–115. Oxford, UK: Oxford University Press.
- Felsenstein, J. (1992). Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genetical Research* 59, 139–147.
- Finnegan, A. K. & Stevens, J. R. (2008). Assessing the long-term genetic impact of historical stocking events on contemporary populations of Atlantic salmon, *Salmo salar*. *Fisheries Management and Ecology* 15, 315–326.
- Fleming, I. A. (1996). Reproductive strategies of Atlantic salmon: Ecology and evolution. *Reviews of Fish Biology and Fisheries* 6, 379–416.
- Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics* 29, 305–327.



- Frankham, R., Ballou, J. D. & Briscoe, D. A. (2002). *Introduction to Conservation Genetics*. Cambridge, UK: Cambridge University Press.
- Franklin, I. R. (1980). Evolutionary change in small populations. In *Conservation Biology: An Evolutionary–Ecological Perspective* (ed. M. E. Soule & B. A. Wilcox), pp. 135–149. Sunderland, MA: Sinauer Associates.
- Franklin, I. R. & Frankham, R. (1998). How large must populations be to retain evolutionary potential? *Animal Conservation* **1**, 69–73.
- Garant, D., Dodson, J. J. & Bernatchez, L. (2000). Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar* L.). *Molecular Ecology* **9**, 615–628.
- García-Vázquez, E., Moran, P., Martínez, J. L., Pérez, J., de Gaudemar, B. & Beall, E. (2001). Alternative mating strategies in Atlantic salmon and brown trout. *Journal of Heredity* **92**, 146–149.
- Goudet, J. (1995). Fstat version 1.2: a computer program to calculate *F*-statistics. *Journal of Heredity* **86**, 485–486.
- Guyomard, R. (1994). La diversité génétique des populations de saumon atlantique. In *Le saumon atlantique: Biologie et gestion de la ressource* (ed. J. C. Gueguen et P. Prouzet), pp. 141–151. Brest, France: IFREMER.
- Gross, M. R. (1998). One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 131–144.
- Hansen, L. P. & Jonsson, B. (1994). Homing of Atlantic salmon: effects of juvenile learning on transplanted post-spawners. *Animal Behavior* **47**, 220–222.
- Hawkins, A. D. (2000). Problems facing salmon in the sea – summing up. In *The Ocean Life of Atlantic salmon: Environmental and Biological Factors Influencing Survival* (ed. D. H. Mills), pp. 211–222. Oxford: Fishing News Books.
- Heled, J. & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* **8**, 289.
- Hey, J. & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747–760.
- IUCN (International Union for Conservation of Nature) (1994). IUCN Red List Categories. Gland, Switzerland: IUCN, Species Survival Commission.
- Jones, M. W. & Hutchings, J. A. (2001). The influence of male parr body size and mate competition on fertilization success and effective population size in Atlantic salmon. *Heredity* **86**, 675–684.
- Jones, M. W. & Hutchings, J. A. (2002). Individual variation in Atlantic salmon fertilization success: implications for effective population size. *Ecological Applications* **12**, 184–193.
- Jonsson, B. & Jonsson, N. (2004). Factors affecting marine production of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences* **61**, 2369–2383.
- Jordan, W. C., Cross, T. F., Crozier, W. W., Ferguson, A., Galvin, P., Hurrell, R. H., McGinnity, P., Martin, S. A. M., Moffett, I. J. J., Price, D. J., Youngson, A. F. & Verspoor, E. (2005). Allozyme variation in Atlantic salmon from the British Isles: associations with geography and the environment. *Journal of Zoology* **65**(Suppl. A), 146–168.
- Jorde, P. & Ryman, N. (1996). Demographic genetics of brown trout (*Salmo trutta*) and estimation of effective population size from temporal change in allele frequencies. *Genetics* **143**, 1369–1381.
- King, J. P., Kimmel, M. & Chakraborty, R. (2000). A power analysis of microsatellite-based statistics for inferring past population growth. *Molecular Biology and Evolution* **17**, 1859–1868.
- King, T. L., Kalinowski, S. T., Schill, W. B., Spidle, A. P. & Lubinski, B. A. (2001). Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology* **10**, 807–821.
- Klemetsen, A., Amundsen, P. A., Dempson, J. B., Jonsson, B., Jonsson, N., O’Connell, M. F. & Mortensen, E. (2003). Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): a review of aspects of their life histories. *Ecology of Freshwater Fish* **12**, 1–59.
- Knockaert, C. (2006). Salmonidés d’aquaculture. De la Production à la Consommation, pp. 327. Versailles, France: Quae Editions.
- Kuhner, M. K. (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics Applications Notes* **22**, 768–770.
- Kuhner, M. K., Yamato, J. & Felsenstein, J. (1995). Estimating effective population size and mutation rate from sequence data using Metropolis–Hastings sampling. *Genetics* **140**, 1421–1430.
- Lande, R. (1988). Genetics and demography in biological conservation. *Science* **241**, 1455–1460.
- Laughton, R. (1991). The movements of adult Atlantic salmon (*Salmo salar* L.) in the River Spey, as determined by radio telemetry during 1988–1989. Scottish Fisheries Research Report 50.
- Leberg, P. (2005). Genetic approaches for estimating the effective size of populations. *Journal of Wildlife Management* **69**, 1385–1399.
- Ligoxygakis, P. (2001). Genetic diversity and conservation efforts. *Trends in Genetics* **17**, 442.
- Luikart, G. H., Cornuet, J. M. & Allendorf, F. W. (1999). Temporal changes in allele frequencies provide estimators of population bottleneck size. *Conservation Biology* **13**, 523–530.
- Lynch, M. & Lande, R. (1998). The critical effective size for a genetically secure population. *Animal Conservation* **1**, 70–72.
- Mace, G. M. & Lande, R. (1991). Assessing extinction threats – toward a reevaluation of IUCN threatened species categories. *Conservation Biology* **5**, 148–157.
- McGinnity, P., Prodohl, P., Ferguson, A., Hynes, R., O’Maoileidigh, N., Baker, N., Cotter, C., O’Hea, B., Cooke, D., Rogan, G., Taggart, J. B. & Cross, T. (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon *Salmo salar* as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London B* **270**, 2443–2450.
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590.
- Nei, M. & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics* **98**, 625–640.
- Nielsen, E. E., Hansen, M. M. & Loeschcke, V. (1999). Genetic variation in time and space: microsatellite analysis of extinct and extant populations of Atlantic salmon. *Evolution* **53**, 261–268.
- Nielsen, R. & Wakeley, J. (2001). Distinguishing migration from isolation. A Markov Chain Monte Carlo approach. *Genetics* **158**, 885–96.
- Nikolic, N., Fève, K., Chevalet, C., Høyheim, B. & Riquet, J. (2009). A set of 37 microsatellite DNA markers for genetic diversity and structure analysis of Atlantic salmon

- Salmo salar* populations. *Journal of Fish Biology* **74**, 458–466.
- Norris, A. T., Bradley, D. G. & Cunningham, E. P. (1999). Microsatellite genetic variation between and within farmed Atlantic salmon (*Salmo salar*) populations. *Aquaculture* **180**, 247–264.
- O'Reilly, P. T., Herbinger, C. & Wright, J. M. (1998). Analysis of parentage determination in Atlantic salmon (*Salmo salar*) using microsatellites. *Animal Genetics* **29**, 363–370.
- Ovenden, J. R., Peel, D., Street, R., Courtney, A. J., Hoyle, S. D., Peel, S. L. & Podlich, H. (2007). The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Molecular Ecology* **16**, 127–138.
- Palstra, F. P. & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology* **17**, 3428–3447.
- Parrish, D. L., Behnke, J., Gephard, S. R., McCormick, S. D. & Reeves, G. H. (1998). Why aren't there more Atlantic salmon (*Salmo salar*)? *Canadian Journal of Fisheries and Aquatic Sciences* **55**(Suppl. 1), 281–287.
- Peakall, R. & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295.
- Porcher, J. P. & Baglinière, J. L. (2001). Le Saumon atlantique. In *Atlas des poissons d'eau douce de France* (ed. P. Keith et J. Allardi (coords)). *Patrimoines Naturels* **47**, 240–243.
- Peel, D., Ovenden, J. R. & Peel, S. L. (2004). NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.
- Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L. & Estoup, A. (2004). GeneClass2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–539.
- Piry, S., Luikart, G. & Cornuet, J.-M. (1999). Bottleneck: a program for detecting recent effective population size reductions from allele data frequencies. *Journal of Heredity* **90**, 502–503.
- Primmer, C. R., Veselov, A. J., Zubchenko, A., Poututkin, A., Bakmet, I. & Koskinen, M. T. (2006). Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. *Molecular Ecology* **15**, 653–66.
- Pritchard, J. K., Stephens, M. & Donnelly, P. J. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.
- Rannala, B. & Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the USA* **94**, 9197–9221.
- Raymond, M. & Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–249.
- Saglio, P. (1994). Le retour aux sites de frai ou “homing”: les mécanismes chimiosensoriels. In *Le saumon atlantique* (ed. J. C. Gueguen and P. Prouzet), pp. 101–122. Brest, France: IFREMER.
- Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* **18**, 233–234.
- Shriver, M. D., Jin, L., Ferrell, R. E. & Deka, R. (1997). Microsatellite data support an early population expansion in Africa. *Genome Research* **7**, 586–591.
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 1463.
- Stabell, O. B. (1984). Homing and olfaction in Salmonids – a critical review with special reference to the Atlantic salmon. *Biological Reviews of the Cambridge Philosophical Society* **59**, 333–388.
- Ståhl, G. (1987). Genetic population structure of Atlantic salmon. In *Population Genetics and Fisheries Management* (ed. N. Ryman and F. M. Utter), pp. 121–140. Seattle, WA: University of Washington Press.
- Stewart, D. C., Smith, G. W. & Youngson, A. F. (2002). Tributary-specific variation in timing of return of adult Atlantic salmon (*Salmo salar*) to freshwater has a genetic component. *Canadian Journal of Fisheries and Aquatic Science* **59**, 276–281.
- Storz, J. F. & Beaumont, M. A. (2002). Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution* **56**, 154–166.
- Tallmon, D. A., Luikart, G. & Beaumont, M. A. (2004). Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* **167**, 977–988.
- Taylor, E. B. (1991). A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. *Aquaculture* **98**, 185–207.
- Vaha, J. P., Erkinaro, J., Niemela, E. & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology* **16**, 2638–2654.
- Vaha, J. P., Erkinaro, J., Niemela, E. & Primmer, C. R. (2008). Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. *Evolutionary Applications* **1**, 137–154.
- Wang, J. & Whitlock, M. C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429–446.
- Wang, J. L. (2001). A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical Research* **78**, 243–257.
- Wang, J. L. (2005). Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences* **360**, 1395–1409.
- Waples, R. S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379–391.
- Waples, R. S. (1990). Conservation genetics of Pacific salmon. III. Estimating effective population size. *Journal of Heredity* **81**, 277–289.
- Waples, R. S. (2002). Effective size of fluctuating salmon populations. *Genetics* **161**, 783–791.
- Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.
- Whitlock, M. C. & McCauley, D. E. (1999). Indirect measures of gene flow and migration:  $F_{st} = 1/(4Nm + 1)$ . *Heredity* **82**, 117–125.
- Wilson, I. J. & Balding, D. J. (1998). Genealogical inference from microsatellite data. *Genetics* **150**, 499–510.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.
- Wright, S. (1938). Size of population and breeding structure in relation to evolution. *Science* **87**, 430–431.