

Genetic stock identification of Atlantic salmon and its evaluation in a large population complex

Juha-Pekka Vähä, Jaakko Erkinaro, Morten Falkegård, Panu Orell, and Eero Niemelä

Abstract: Addressing biocomplexity in fisheries management is a challenge requiring an ability to differentiate among distinct populations contributing to fisheries. We produced extensive genetic baseline data involving 36 sampling locations and 33 microsatellite markers, which allowed characterization of the genetic structure and diversity in a large Atlantic salmon (*Salmo salar*) population complex of the River Teno system, northernmost Europe. Altogether, we identified 28 hierarchically structured and genetically distinct population segments (global $F_{ST} = 0.065$) corresponding exceptionally well with their geographical locations. An assessment of factors affecting the stock identification accuracy indicated that the identification success is largely defined by the interaction of genetic divergence and the baseline sample sizes. The choice between the two statistical methods tested for performance in genetic stock identification, ONCOR and cBAYES, was not critical, albeit the latter demonstrated slightly higher identification accuracy and lower sensitivity to population composition of the mixture sample. The strong genetic structuring among populations together with a powerful marker system allowed for accurate stock identification of individuals and enabled assessment of stock compositions contributing to mixed-stock fisheries.

Résumé : La question de la biocomplexité dans la gestion des pêches pose un défi qui requiert la capacité de distinguer les différentes populations contribuant aux pêches. Nous avons produit une grande quantité de données génétiques de référence portant sur 36 lieux d'échantillonnage et 33 marqueurs microsatellites, qui rend possible la caractérisation de la structure et de la diversité génétiques dans un grand complexe de populations de saumons atlantiques (*Salmo salar*) du réseau de la rivière Teno, dans l'extrême nord de l'Europe. En tout, nous avons identifié 28 segments de population structurés hiérarchiquement et génétiquement distincts (F_{ST} global = 0,065) qui correspondent exceptionnellement bien à leurs emplacements géographiques. Une évaluation des facteurs ayant une incidence sur l'exactitude de l'identification du stock indique que le succès de l'identification est en bonne partie défini par l'interaction de la divergence génétique et des tailles des échantillons de référence. Le choix entre les deux méthodes statistiques testées pour leur performance en matière d'identification du stock génétique, ONCOR et cBAYES, n'était pas critique, même si la seconde méthode était caractérisée par une exactitude légèrement plus grande de l'identification et qu'elle était moins sensible aux populations qui composaient l'échantillon mixte. La forte structuration génétique au sein des populations, combinée à un puissant système de marqueurs, a permis l'identification exacte du stock des individus et a rendu possible l'évaluation des stocks qui contribuent à des pêches visant des stocks mixtes. [Traduit par la Rédaction]

Introduction

Species diversity supports community stability and productivity (e.g., Tilman et al. 1996) as well as ecosystem functioning and services (Garnfeldt et al. 2008; Isbell et al. 2011), while intraspecific phenotypic and genotypic diversity plays an important role in population persistence and dynamics (Agashe 2009; Bolnick et al. 2011). Biocomplexity in fish populations — in the form of population diversity, life-history variation, and genetic structures — is important for sustainable fisheries (Hilborn et al. 2003; Hutchinson 2008). Population diversity provides resilience to disturbances (e.g., exploitation), which contribute to long-term sustainability, and diverse systems provide more temporally stable ecosystem services; a phenomenon referred to as the “portfolio effect” (Schindler et al. 2010). Strong or selective harvesting can negatively affect abundance and diversity (e.g., Youngson et al. 2003) and requires reductions in harvest to protect less abundant stocks. This can be at odds with socio-economic factors that might point towards continued harvest, to provide food, income, and other social goods. Managers balancing this trade-off between

harvesting and conservation needs face a particularly challenging quest when the exploitation targets a multitude of species or populations.

In mixed-stock fisheries, stocks at poor status might be targeted to an unknown extent, threatening small and vulnerable populations (Nehlsen et al. 1991; ICES 2015). In the absence of detailed, stock-specific knowledge, the precautionary approach would be a substantial reduction or closure of the mixed-stock fishery. Thus, managers need tools that enable stock-specific estimates of exploitation to enable a stock-specific management of mixed-stock fisheries, tailored towards safeguarding vulnerable populations while allowing the continued harvest of healthy stocks (Crozier et al. 2004).

Molecular genetic techniques and the application of genetic markers have not only revolutionized our understanding of population diversity, but genetic aspects have become an integral part of fishery and wildlife management (Mills 2012; Allendorf et al. 2013). Genetic monitoring of the population composition of mixed-stock catches provides a way to ensure a stock-specific

Received 29 December 2015. Accepted 4 July 2016.

J.-P. Vähä,* Kevo Subarctic Research Institute, University of Turku, Finland.
J. Erkinaro, P. Orell, and E. Niemelä. Natural Resources Institute Finland (Luke), POB 413, FI-90014 Oulu, Finland.
M. Falkegård. Norwegian Institute for Nature Research, Tromsø, Norway.
Corresponding author: Jaakko Erkinaro (email: jaakko.erkinaro@luke.fi).

*Present address: Association for Water and Environment of Western Uusimaa, POB 51, FI-08101 Lohja, Finland.

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.elsevier.com/locate/permissions).

management of the mixed-stock fisheries, including the knowledge needed to establish targeted stock-specific regulatory measures tailored towards safeguarding vulnerable populations. With the advent of powerful genetic markers and reduced costs of analyzing large numbers of samples accompanied with the development of tailored statistical methods, genetic stock identification (GSI) is now one of the most successful biological tools available for effective monitoring and stock-specific management actions (e.g., Beacham et al. 2008; Ensing et al. 2013).

Anadromous fish populations undertake feeding migrations between fresh water and open ocean areas and return to spawn in their natal rivers after sexual maturation. Accurate homing to natal rivers (e.g., Quinn 1993) provides the potential for genetic population structuring between and even within river systems (Vähä et al. 2007; Hess et al. 2014). At sea, anadromous fish populations may mix, and harvesting during the migratory stage targets multiple populations resulting in a mixed-stock fishery. GSI has been used to manage Pacific salmon (*Oncorhynchus* spp.) fisheries for three decades (Milner et al. 1985). Ecological applications of GSI on Pacific salmon species cover a variety of spatial and temporal scales from ocean distribution of juveniles and adults (e.g., Beacham et al. 2006; Habicht et al. 2010) to in-river investigations of population of origin among migrating individuals (e.g., Beacham et al. 2008; Hess et al. 2014). GSI is a routine tool for monitoring mixed-stock fisheries of Pacific salmon, providing real-time information on catch stock composition to managers and fishermen (e.g., <http://pacificfishtrax.org>).

In contrast, GSI has been less used in management of Atlantic salmon (*Salmo salar*) mixed-stock fisheries. There are only some marine examples: for example, genetic mixed-stock analyses resolving stock group proportions in the Baltic sea (Koljonen 2006), contributions of North American Atlantic salmon stocks to the West-Greenland salmon fishery (Gauthier-Ouellet et al. 2009), and some coastal studies demonstrating the potential for such approach (e.g., from England (Gilbey et al. 2012), Ireland (Ensing et al. 2013), Canada (Bradbury et al. 2015), St. Pierre & Miquelon of France (Bradbury et al. 2016), and the Barents Sea coast in Norway and Russia (Vähä et al. 2014)). Freshwater examples include lake-run brown trout (*Salmo trutta*; Swatdipong et al. 2013), lake trout (*Salvelinus namaycush*; Northrup et al. 2010), and broad whitefish (*Coregonus nasus*; Harris and Taylor 2010). However, genetic mixed-stock analyses of Atlantic salmon fisheries in large rivers have so far received little or no attention (Vähä et al. 2011).

The River Teno (Tana in Norwegian) in the northernmost Europe is one of the few remaining large river systems that still support multiple and abundant wild Atlantic salmon populations (Niemelä et al. 2006; Vähä et al. 2007). Genetic studies have revealed a structured population complex consisting of a number of demographically independent, genetically distinct, and temporally stable population segments (Vähä et al. 2007, 2008). Genetic assessment of systematically collected salmon scale samples from the mixed-stock fishery of the Teno main stem can provide a fine-resolution estimation of the origin of the captured salmon (Vähä et al. 2011). In the present study, we set out to (i) complement the baseline data on salmon population structure of the Teno complex, (ii) assess factors affecting the GSI success, and (iii) investigate the feasibility of using GSI to monitor the harvest of multiple salmon populations in the mixed-stock fishery of the River Teno main stem.

Material and methods

Baseline data

The examined genetic baseline consisted of (1) previously described baseline data from 12 tributaries ($n = 1076$; Vähä et al. 2007,

2008, 2011), which all were supplemented by (2) a new collection of samples ($n = 866$), (3) samples from 15 previously unsampled tributaries ($n = 770$), and (4) samples from eight different parts of the Teno main stem ($n = 611$: TMS TB, AK, PI, SI, GJ, KO, YK, OUT) and from the Inarijoki ($n = 71$) (Fig. 1). The old baseline data were collected by sampling adult salmon catches (Vähä et al. 2007), while the new data consisted of juvenile samples. F1

Juvenile salmon were electrofished and sampled for adipose fin tissue. Only one individual of both major juvenile salmon age groups (age-0, age-1) were sampled at one site, and the next pair of fish were sampled 50–100 m apart, depending on the area, to minimize the probability of sampling siblings. License for sample collection was issued by the County Governor of Finnmark, Norway, and the Center for Economic Development, Transport and the Environment in Finland. Individual tissue samples were stored in 96% ethanol. In total, the updated baseline data consisted of 3323 salmon from 36 sites within the Teno river system (Fig. 1; also see online supplementary material, Table S1¹). Owing to the temporal stability of the genetic structure within the Teno system (Vähä et al. 2008), temporally replicated samples were pooled for each site. Fn2

Test samples

To test the performance of the new baseline, we used test samples of known origin. Two sets of test samples (TS) were used: (i) late season samples of adult salmon from five areas of the main stem (TS 1–5; Fig. 1) and (ii) juvenile samples collected from spawning grounds at the lowermost part of the River Utsjoki, a major tributary of the Teno (TS 6; Fig. 1). The first set comprised samples from the old baseline data (Vähä et al. 2007), which were not included in the current baseline. While these individuals were sampled in late August after migration period and assumed to originate from that area, the samples may include transient fish from nearby populations (Vähä et al. 2007; Erkinaro et al. 2010).

Microsatellite analysis

Total genomic DNA was extracted from scale or fin tissue as previously described (Vähä et al. 2007). In the present study, we used genotypes of 33 microsatellite markers, of which 32 were described in detail by Vähä et al. (2007, 2011). In addition, locus Sssp3016 (Paterson et al. 2004) was added to the marker panel. Primer sequences for amplifying alternative MHC I locus amplicons (Grimholt et al. 2002) were as follows: forward: GAAGGTGCTGAA-GAGGAACGTC and reverse: GTTTCATTACCACAAGCCCGCTC.

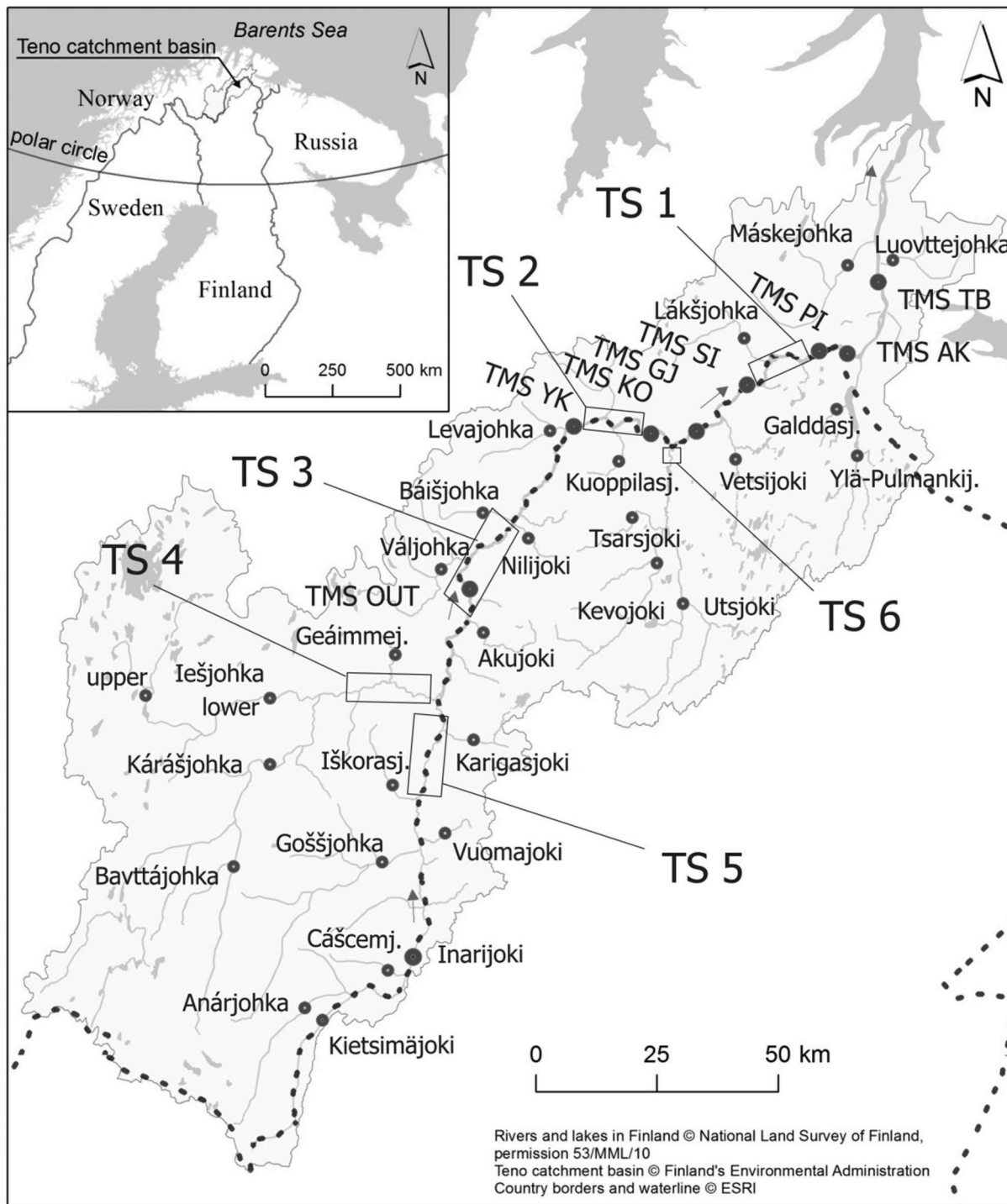
All microsatellite loci were amplified by multiplex PCR slightly modified from Vähä et al. (2011), and details are available from the authors upon request. Varying volumes of the PCR-amplified products were pooled and electrophoresis was then performed on ABI 3130xl (Applied Biosystems). Electropherograms were inspected and allele scoring performed with GeneMapper version 4.0 (Applied Biosystems) followed by manual corrections. Electropherograms and allele scores were reviewed by two people independently.

Microsatellite variability

All 33 microsatellite loci examined were polymorphic in all 36 baseline samples, displaying 527 alleles in total (4–33 per locus). Tests of conformance to Hardy–Weinberg equilibrium (HWE) expectations over all samples indicated that loci MHC I ($p < 0.001$) and Sssp2201 ($p < 0.01$) departed significantly from expectation. In locus MHC I, the departure from HWE was caused by the presence of a null allele; when a subset of individuals ($n = 2064$) from all baseline samples was amplified with an alternative set of primers, the null allele was detected with a frequency varying from 0.01 to 0.12 in 14 of the samples. An excess of homozygotes in locus Sssp2201 may arise from large allele dropouts, since the locus is

¹Supplementary data are available with the article through the journal Web site at <http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2015-0606>.

Fig. 1. Map of the Teno river system and its location in northernmost Europe. Locations of the sampled baseline populations are indicated with circles (small circles — tributary samples; large circles — main stem samples). Rectangles indicate sites where test samples (TS 1–6) were collected.



highly variable (33 alleles) with long amplicons (243–387 base pairs).

Over all loci, excluding MHC I and Sssp2201, only the Anárjohka sample deviated significantly from conformity to HW proportions, expressing an excess of homozygotes (Bonferroni correction applied; $\alpha = 0.05$). Deficiency of heterozygotes was most likely due to allelic dropouts stemming from low-quality DNA extracts (inferred from peak intensity of electropherograms) from archived juvenile scales collected in 1996.

Regardless, all loci and samples were retained in the baseline, as several previous studies have indicated that inclusion of additional marker data despite aberrations add to the accuracy of genetic stock identification (Beacham et al. 2006; Carlsson 2008; Griffiths et al. 2010).

Statistical analysis

Genetic differentiation between samples and deviations from HWE within and across loci for each locality and globally were

tested using F statistics of Weir and Cockerham (1984), with significance levels calculated with a randomization procedure (3300 permutations) as implemented in FSTAT version 2.9.3 (Goudet 2001). FreeNA (Chapuis and Estoup 2007) was used to assess the presence of null alleles and their effect on the global and pairwise F_{ST} estimates as well as on the tree topology.

Genetic relationships among samples were estimated as Cavalli-Sforza and Edwards' Dc (Cavalli-Sforza and Edwards 1967) genetic distance using PowerMarker version 3.25 (Liu and Muse 2005), and the genetic relationships were visualized by the neighbor-joining method (Saitou and Nei 1987) in SplitsTree4 version 4.11.3 (Huson and Bryant 2006). Robustness of clustering was evaluated by calculating split support values by bootstrapping 10 000 times over loci using PHYLIP (Felsenstein 2005).

The Bayesian clustering approach of STRUCTURE version 2.3.2. (Pritchard et al. 2000) with the correlated allele frequency model (Falush et al. 2003) was used to create plots of ancestry (i.e., admixture coefficients (Q)) for evaluating genetic clustering of individuals (Fig. S1'). Pairwise and global F_{ST} values as well as the variance components among groups of populations were calculated using ARLEQUIN version 3.5.1 (Excoffier et al. 2005). Population-specific F_{ST} values were calculated using GESTE version 2.0 (Foll and Gaggiotti 2006).

For testing isolation-by-distance patterns, Mantel tests were performed using PASSaGE 2 (Rosenberg and Anderson 2011). Geographical distances (km) between sampling localities were plotted against the estimates of $F_{ST}/(1 - F_{ST})$ (Rousset 1997), and the significance was tested through a randomization with 10 000 permutations of the data.

The effects of genetic divergence and sample size of the baseline population on GSI success were analyzed using generalized linear mixed models (GLMMs). Identification events per trial from analyses of known-origin mixtures constructed by resampling individuals (see below) were used as the response variable, whereas population-specific F_{ST} and baseline sample size were set as fixed effects and treated as continuous variables. Analyses were performed using the GLIMMIX procedure of SAS version 9.3 (SAS Institute Inc.) with a logit link function and a binomial error term.

GSI methods

GSI performance was tested using ONCOR (Kalinowski et al. 2007) and cBAYES (Neaves et al. 2005). ONCOR implements the method of Rannala and Mountain (1997), which uses an equal probability Dirichlet density as the prior for the allele frequencies at a locus, assigning a frequency of $1/(n + 1)$ to an allele not found in a population. The prior densities updated with the observed baseline data give the posterior probability densities of allele frequencies.

cBAYES (Neaves et al. 2005) implements the Bayesian method of Pella and Masuda (2001). The prior distribution of alleles at a locus follows the mean of the allele frequencies over all stocks, and posterior distributions of the baseline allele frequencies are the product of priors and the observed allele frequencies. Shrinking the observed values toward central values takes advantage of the genetic similarity of populations and is thought to minimize estimation error in allele frequencies. Further, the allele frequencies of mixture individuals assigned to a baseline population, at each MCMC step, are used to update the baseline allele frequencies.

All analyses with cBAYES were performed using eight independent chains of 100 000 iterations starting from four random stocks. The number of replications was increased if diagnostics indicated convergence issues (shrink factor by population > 1.2).

Power analysis

The discriminatory power of the baseline for stock identification was first evaluated through simulation procedures implemented in the program ONCOR, which has been shown to

produce essentially unbiased estimates of GSI accuracy (Anderson et al. 2008).

A second power test analyzed known-origin mixtures constructed by resampling individuals without replacement from each of the baseline samples. Our purpose was to compare the performance of the GSI methods when the data included missing genotypes and potential genotyping errors, typical of microsatellite data. Each mixture sample (12 in total) was composed of ~10% randomly chosen individuals from each baseline sample ($n = 3816$) and was analyzed against remaining baseline samples.

Both methods above use the estimated stock composition of the mixture sample, which might cause bias. Resampled baseline test individuals were therefore analyzed jointly with varying combinations of test samples (as described above) to assess the sensitivity of the power estimates to mixture sample composition. To evaluate the effect of mixture sample composition while retaining full baseline data, mixture samples with varying combinations of test samples were constructed and analyzed and their effect on individual assignment patterns assessed.

Throughout the paper, the proportion of correctly identified individuals is referred to as the correct identification rate, applicable only for samples with known origin. The proportion of correctly identified individuals of all individuals assigned to a specific baseline population is referred to as the correct assignment rate. Efficiency was defined as the proportion of individuals in a group that were correctly identified, and accuracy was defined as the proportion of an identified group that truly belongs to that category (see Vähä and Primmer 2006 for details).

Results

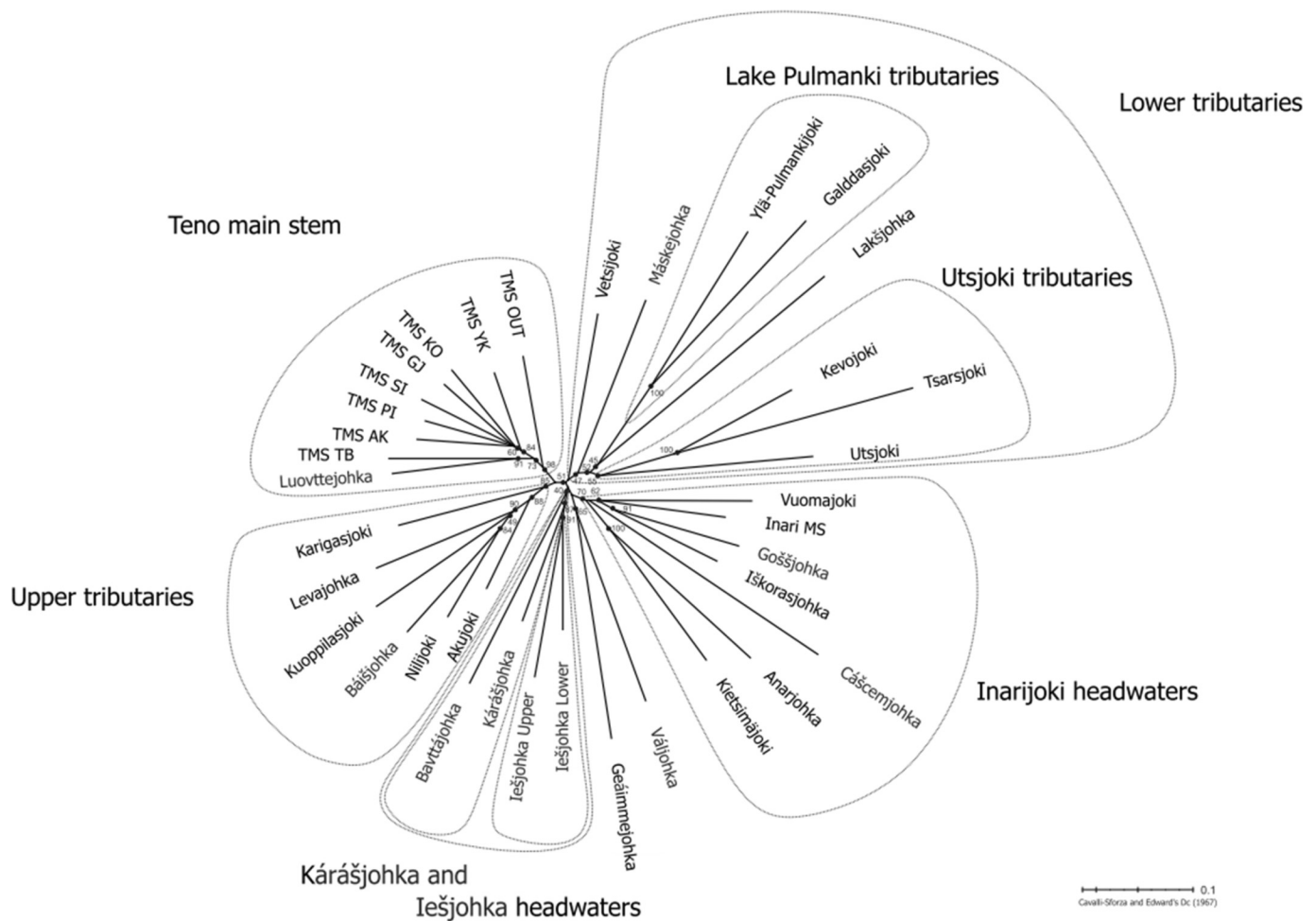
Genetic structure of Teno salmon

Overall, genetic structuring within the Teno salmon population complex was strong and highly significant (global $F_{ST} = 0.065$, $p < 0.001$). All pairwise comparisons of genetic differentiation were statistically highly significant ($p < 0.001$) apart from those among five Teno main stem samples (AK, PI, SI, GJ, and KO) and between Teno main stem YK and GJ (Fig. 1; $F_{ST} = 0.002$, $p = 0.024$; Bonferroni-adjusted p value for $\alpha = 0.05$ was 0.022 across 630 correlated tests, Pearson's $r = 0.87$). Estimates of genetic differentiation were affected by null alleles (mean $d = 0.0007$, paired t test $t = 21.9$, $df = 629$, $p < 0.001$), but this did not have a significant effect on the overall pattern (Pearson's $r = 0.999$, $p < 0.001$) or the overall level of differentiation (mean pairwise F_{ST} 0.058 versus 0.057 for null allele corrected F_{ST} ; $t = 0.32$, $p = 0.75$).

Neighbor-joining analysis based on genetic distance with and without null allele correction (Cavalli-Sforza and Edwards 1967) provided the same tree topology (Fig. 2). In general, patterns of genetic relationships among samples corresponded well with their geographical locations. There was a clear distinction between the Teno main stem and the headwater river systems, as well as between groups of tributaries draining to lower and upper parts of the main stem. Analysis of molecular variance (AMOVA), made in accordance with these geographical groupings, apportioned 1.2% of the total genetic variation among groups and 5.6% between populations within groups. Variation among populations was highest in the "lower tributaries" group (10%) and lowest in the "main stem" group (0.64%; Table 1).

Closer inspection of the genetic structuring revealed less distinct patterns within the Inarijoki headwater group and Teno river main stem. In addition to very low levels of genetic differentiation between Iškorasjohka and Goššjohka ($F_{ST} = 0.008$) as well as between Vuomajoki and Inarijoki ($F_{ST} = 0.008$) within the Inarijoki basin, genetic clustering of individuals with the program STRUCTURE suggested no significant divergence between these samples (results not shown in detail). Similarly, in the Teno main stem, model-based clustering of individuals without sample location information did not support explicit population boundaries.

Fig. 2. Unrooted neighbor-joining phylogram based on Cavalli-Sforza and Edwards' genetic distances among samples collected from 36 sites within the Teno river system. Bootstrap values shown are in percentage of 10 000 replicates.



However, despite low levels of genetic structure, a signal of isolation by distance was observed from Tana Bru to Outakoski (Mantel's $r_{xy} = 0.638$, $p = 0.010$) as well as from Teno main stem AK to YK (Mantel's $r_{xy} = 0.58$, $p = 0.026$; see Fig. 1 for sample locations).

The observed strong clustering of geographically close populations and the configuration of regional groups of populations allowed their proper use as reporting groups in the mixed-stock analyses. For subsequent power analyses, baseline samples from Teno main stem AK, PI, SI, GJ, and KO were pooled and referred to as the Teno main stem lower. This pooled area along with the Teno main stem at TB, YK, and OUT and all the tributary samples were tested separately for the power of resolution.

Power analyses

An analysis of simulated single-stock mixture samples provided the first reference point for evaluating the resolving power of the baseline data. Across 32 baseline samples, mean identification rate was 91% ($\pm 12\%$) with 18 sites showing >95% identification accuracy, implying a high-resolution GSI power within the Teno river system (Table 1).

Analyzing simulated mixtures with equal proportions from all 32 baseline samples provided a pattern similar to 100% simulations (Pearson's $r = 0.975$, $p < 0.001$; Fig. 3). Despite a lower identification rate ($82\% \pm 19\%$) compared with single-stock simulations, 15 baseline samples showed >95% correct identification rates. The correct assignment rate was slightly higher ($85\% \pm 19\%$) and higher than 95% for 16 sites.

ONCOR and cBAYES performed equally well in identifying the stock of origin for individuals in mixtures constructed by resampling (83.9% and 84.8%, respectively). The mean site-specific identification rates (ONCOR: $77\% \pm 20\%$ and cBAYES: $79\% \pm 21\%$) and the correct assignment rates (ONCOR: $84\% \pm 18\%$ and cBAYES: $83\% \pm 18\%$) were slightly lower than results from simulated mixtures (Fig. 3).

An assessment of factors affecting stock identification accuracy of the two methods indicated that the success rate is defined by the interaction of genetic divergence and baseline sample size (Fig. 4; Table S2¹). The effect of baseline sample size is large when the population divergence is low ($F_{ST} < 0.05$), while increasing sample size beyond 150 benefits only very little with highly diverged populations ($F_{ST} > 0.07$). The models predict cBAYES to require smaller sample size than ONCOR at any population divergence level for the same identification success.

There was a clear tendency for misassignments to occur between nearby tributaries or locations; as a corollary, higher stock identification success to regions or groups of nearby rivers was achieved (Table 1). For example, in equal proportion simulations, 74% of the incorrectly identified TMS YK individuals were assigned to TMS lower, 13% to TMS OUT, and 3% to TMS TB. Thus, despite low (36%) identification success to exact site of origin, 95% of TMS YK individuals were successfully identified to originate from the Teno main stem.

Table 1. Genetic differentiation of baseline samples and results from single and multisample simulations in ONCOR.

Baseline sample	F_{ST}	F_{ST} within	100% simulation	Equal prop. ID rate	Largest misassignment to	Stock contribution	ID rate/stock contribution
Iešjohka headwaters							
Iešjohka Upper	0.035		92%	79%	Iešjohka Lower	12%	82%
Iešjohka Lower	0.020		98%	87%	Kárášjohka	3%	112%
Kárášjohka headwaters							
Kárášjohka	0.027	3%	100%	95%	Iešjohka Lower	2%	114%
Bavttájohka	0.067		99%	96%	Kárášjohka	2%	96%
Geáimmejohka	0.101		100%	100%			100%
Inarijoki headwaters							
Kietsimäjoki	0.072		99%	96%	Inarijoki MS	3%	98%
Anárjohka	0.078		100%	98%	Inarijoki MS	1%	99%
Inarijoki MS	0.022		92%	79%	Inarijoki MS	6%	144%
Cáčsemjohka	0.106	3.6%	97%	92%	Inarijoki MS	7%	92%
Vuomajoki	0.032		67%	47%	Inarijoki MS	23%	53%
Goššjohka	0.034		83%	64%	Iškorasjohka	12%	80%
Iškorasjohka	0.029		79%	61%	Inarijoki MS	17%	95%
Upper tributaries							
Váljohka	0.075		100%	99%	Kárášjohka	0.3%	100%
Karigasjoki	0.032		96%	85%	Akujoki	4%	104%
Akujoki	0.021		73%	49%	Nilijoki	19%	83%
Báišjohka	0.058	4.23%	87%	72%	Nilijoki	19%	83%
Nilijoki	0.033		78%	63%	Akujoki	19%	114%
Levajohka	0.038		93%	83%	Nilijoki	5%	93%
Kuoppilasjoki	0.067		99%	97%	Nilijoki	2%	102%
Lower tributaries							
Tsarsjoki	0.221		100%	100%			100%
Kevojoki	0.083		100%	100%			100%
Utsjoki	0.094		100%	100%			100%
Vetsijoki	0.038	10%	100%	98%	TMS Lower	1%	101%
Lakšjohka	0.169		100%	100%			100%
Ylä-Pulmankijoki	0.136		100%	100%			100%
Galldasjoki	0.187		100%	100%			100%
Máskejohka	0.042		100%	98%	TMS Lower	0.3%	100%
Teno main stem							
TMS OUT	0.012		75%	56%	TMS YK	20%	93%
TMS YK	0.008		54%	36%	TMS Lower	48%	108%
TMS Lower ^a	0.010	0.64%	93%	79%	TMS YK	15%	214%
TMS TB	0.024		84%	63%	TMS Lower	23%	79%
Luovttiejohka	0.024		77%	58%	TMS Lower	23%	65%

Note: Column abbreviations: F_{ST} , population-specific estimate of genetic differentiation (Foll and Gaggiotti 2006); F_{ST} within, percentage of variation among populations within group; 100% simulation, discriminatory power of the baseline for stock identification as predicted from single-stock simulations in ONCOR (Kalinowski et al. 2007); Equal prop. ID rate, proportion of correctly identified individuals in equal-proportion, multisample simulations in ONCOR; Largest misassignment to, baseline stock with largest proportion of incorrectly identified individuals; Stock contribution, estimated contribution of stock to equal-proportion, multisample simulation mixtures; ID rate/stock contributions, proportion of correctly identified individuals and estimated stock contributions to a regional group. Inarijoki MS = Inarijoki main stem.

^aBaseline samples from Teno main stem AK, PI, SI, GJ, and KO were pooled and referred to as the Teno main stem lower; see text.

While estimated contributions to mixtures were relatively accurate ($\pm 10\%$) for the majority of sites, the tendency of some stocks to receive largest proportions of the misidentified samples implied potential for biased estimates (Table 1). For example, in the equal proportion mixtures, the contribution of TMS lower was overestimated by a factor of 2.1 and Inarijoki main stem by 1.4. Again, since misassignments occurred largely within regions, significantly more accurate estimates were obtained at the regional level (1.12 and 0.95 for the Teno main stem and the Inarijoki region, respectively).

Analysis of test samples

The estimated origins of test samples were to a large extent according to expectations for both ONCOR and cBAYES (Table 2). Both methods correctly estimated all but one (127/128) individual salmon (TS 6) to originate from the Teno main stem lower population. On the other hand, TS 3 from TMS OUT fishery appeared to include a high proportion of transient individuals originating from the Inarijoki region. The most significant difference in assignment patterns between the methods was observed for TS 4,

which was presumed to include individuals originating from populations within the Inarijoki region, including the Karigasjoki tributary. While cBAYES assigned 97% of the test individuals to expected local populations, the same was true only for 88% of the samples by using ONCOR.

Analyses of mixture samples containing varying proportions of individuals from TS 4 (Inarijoki region) and 5 (Kárášjohka-Iešjohka region) and from a pool of TS 1, 2, and 6 (Teno main stem) indicated that cBAYES performed better than ONCOR in estimating the region of origin (Table 3). cBAYES identified more than 90% of the individuals from the headwater tributaries in all mixture assemblies, while the identification success with ONCOR varied from 78% to 92%. ONCOR systematically assigned lower than expected proportions of individuals to the Inarijoki region and higher than expected to Teno main stem. The effect of this bias is best illustrated when no Teno main stem samples were expected present in a mixture assembly; main stem contribution was then estimated at 7% with ONCOR and 1% with cBAYES.

Fig. 3. Stock-specific identification and assignment rate estimates from simulated equal proportion mixtures with ONCOR (grey line) and from resampled baseline mixtures with ONCOR (solid black line) and cBAYES (dashed black line). Horizontal lines show mean rates over all stock estimates. HW = large headwater tributaries.

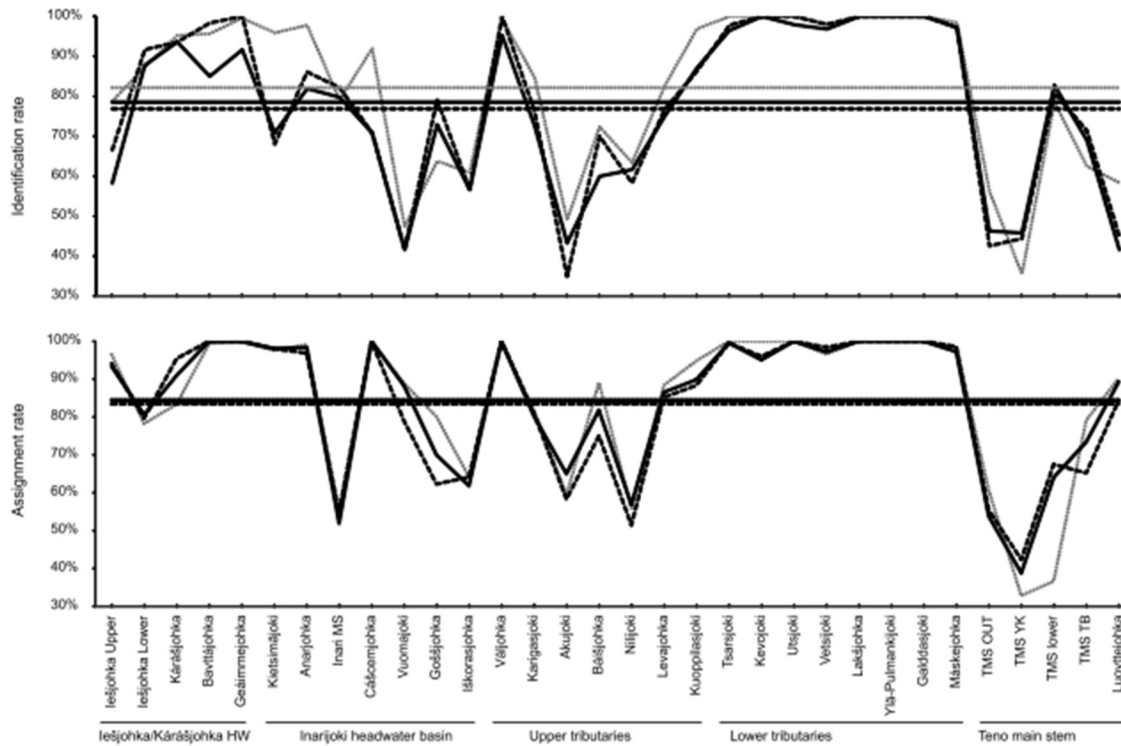
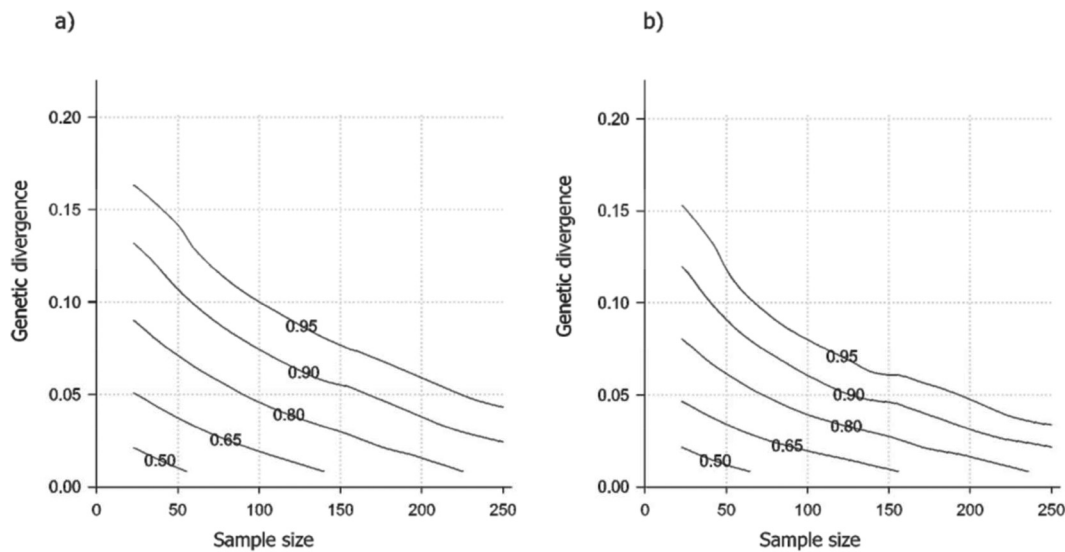


Fig. 4. Contour plot views of GLMM model predictions for identification success using (a) ONCOR or (b) cBAYES versus genetic differentiation and baseline sample size.



A comparison of individual stock estimates indicated that results for both ONCOR and cBAYES were affected by mixture sample composition. Joint analyses of resampled mixture samples and different combinations of real test samples obtained the same stock estimate for 97.7% ($\pm 0.5\%$) of individuals with ONCOR and for 96% ($\pm 1.7\%$) of individuals with cBAYES. For the test samples, the pattern was the opposite and more significant. For example, when test samples from the headwater regions were analyzed separately and in combination with varying proportions of the Teno main stem samples, ONCOR provided the same stock estimate only for 91% of individuals, while 96% of the estimates were unaffected with cBAYES.

GSI of mixed-stock fishery samples

The present study was based on an improved, complemented baseline that allows for fine-scale assignment of individual salmon to 28 genetic stocks, while the original baseline, used in a study assessing run timing of one-sea-winter Atlantic salmon (Vähä et al. 2011), contained only 14 baseline stocks. A reanalysis of the previous study showed that regional genetic stock estimates remained unchanged for 83% of the individuals. The largest changes were observed for the Teno main stem (20%) and Inarijoki regions (31%), which were expected given the now more thorough sampling and better understanding of the genetic structure in

Table 2. Estimated origins for the adult (TS 1–5) and juvenile (TS 6) salmon test individuals collected in different parts of the Teno river system (see Fig. 1 for locations).

	Test sample 1		Test sample 2		Test sample 3		Test sample 4		Test sample 5		Test sample 6	
	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES
1. Iešjohka Upper									2	3		
2. Iešjohka Lower	2	2			3	2	5	4	78	76		
3. Kárášjohka	1	1	1	1	3	2			76	80		
4. Bavttájohka									4	4		
5. Geáimmejohka					1	1			13	13		
6. Kietsimäjoki					3	3	7	19				
7. Anarjohka							5	5				
8. Inari MS	7	10	3	7	20	18	67	54				
9. Cáščemjohka							3					
10. Vuomajoki			1		4	7	1	4				
11. Goššjohka	3	1	2	2	15	26	73	125				
12. Iskurasjoki	2	2	1		8	4	45	15				
13. Karigasjoki					2		12	12				
14. Váljohka					1	1						
15. Akujoki					3	5						
16. Báišjohka												
17. Nilijoki												
18. Levajohka			1									
19. Kuoppilasjoki												
20. Tsarsjoki											1	1
21. Kevojoki	1	1										
22. Utsjoki	1											
23. Vetsijoki							3	1				
24. Lakšjohka	1	1										
25. Ylä-Pulmankijoki												
26. Galldasjoki												
27. Máskejohka									2	1		
28. TMS Outakoski	4	4	1	1	10	5	11					
29. TMS YK	6	4	4	6	26	25	7		6	4		
30. TMS lower	69	71	54	52			1	2			127	127
31. TMS TB	1	1	5	4								
32. Luovttejohka							1					

Note: Numbers within a grey-shaded area are according to a priori expectations.

Table 3. Effect of mixture sample stock composition to estimated stock proportions in ONCOR and cBAYES.

	Assembly of mixture sample: INA/KAR-IES/TMS							
	57%/43%/0%		40%/40%/20%		33%/33%/33%		20%/20%/60%	
	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES	ONCOR	cBAYES
INA								
Identified	84%	95%	81%	94%	81%	94%	78%	90%
Assigned	98%	99%	93%	93%	88%	88%	77%	77%
Estimated contribution	49%	55%	35%	41%	30%	36%	20%	24%
KAR-IES								
Identified	92%	97%	91%	94%	89%	92%	88%	91%
Assigned	92%	95%	92%	95%	91%	95%	90%	94%
Estimated contribution	43%	44%	40%	40%	33%	32%	20%	20%
TMS								
Identified	NA	NA	81%	79%	84%	84%	89%	89%
Assigned	NA	NA	68%	86%	79%	90%	91%	94%
Estimated contribution	7%	1%	24%	19%	35%	31%	58%	56%

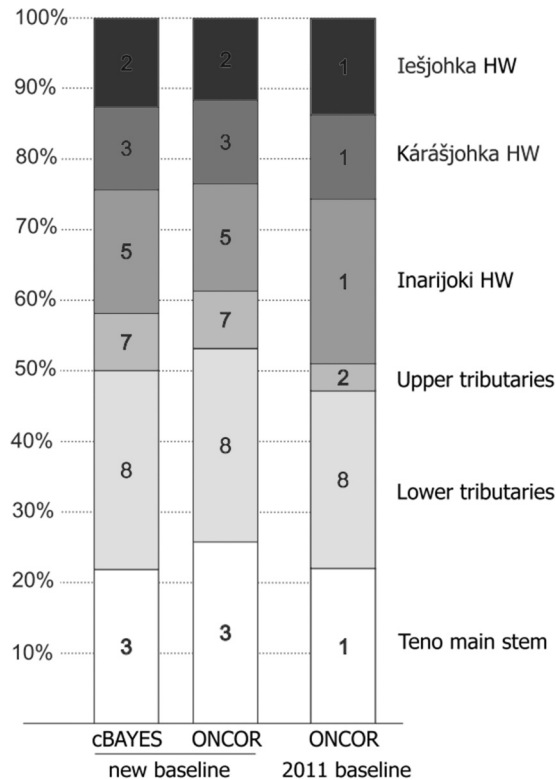
Note: "Assembly of mixture sample" denotes the stock composition of the mixture sample: INA = Inarijoki, KAR-IES = Kárášjohka-Iešjohka, TMS = Teno main stem. "Identified" denotes the correctly identified samples from each region. "Assigned" denotes the correctly identified individuals of all assigned to a region.

these areas. The largest relative change in the estimated contribution to the mixed-stock fishery was observed for the upper main stem tributaries (from 3.9% to 8.1%). The observed tendency of ONCOR to assign lower than expected proportion of individuals to Inarijoki region and higher than expected to Teno main stem was evident also in the apportionment of the mixed-stock fishery samples (Fig. 5).

Discussion

In this study, we established a comprehensive genetic baseline reflecting the Atlantic salmon population structure of the River Teno, with an assessment of factors affecting the performance of GSI within a population complex displaying large genetic variation. In general, the large genetic variation among populations within the river system coupled with the powerful marker system

Fig. 5. Estimated mixture proportions of mixed-stock fishery samples used by Vähä et al. (2011) as inferred with cBAYES and ONCOR applying the new, improved baseline in comparison with the original 2011 baseline. Numbers in boxes refer to number of baseline populations within a region. Size of the box refers to the estimated contribution of a regional group to the total sample. HW = large headwater tributary.



allowed individual salmon from mixed-stock samples to be accurately identified to the population of origin. Furthermore, the strong coherence of geographical and genetic relationships of populations allowed defining rational units for monitoring and management of mixed-stock fisheries. Below, we discuss the main points relevant to the accuracy of the genetic stock estimates and the applicability of GSI in fisheries management.

Genetic structure

Populations within the Teno river system display large variation in their degree of differentiation and diversity, and the improved baseline revealed a substantially improved geographical pattern from the earlier structure (Vähä et al. 2007). This was facilitated by including several previously unsampled tributary populations, but also by replacing adult salmon samples from the Teno main stem and the Inarjoki by juvenile samples in the baseline data. In addition, the more dense sampling coverage allowed clarification of the genetic structure among different areas of the main stem and discrimination between the Teno main stem and the Inarjoki.

A robust genetic clustering of populations was highly coherent with geographical locations as illustrated by the neighbor-joining analyses. Broadly, local populations constituted five distinct groups of populations, wherein more local and strong clusters were evident. This illustrates a strong hierarchical substructuring of Teno salmon and necessitates adjusting management strategies accordingly. Altogether, after excluding three small tributaries that likely do not have self-sustaining populations (Vuomajoki, Iškorasjohka, and Luovttejohka) and treating the middle part of the main stem (between the large rapid sections Alaköngäs

(TMS AK) and Yläköngäs (TMS YK)) as one, we identified 28 genetically distinct demes within the Teno river system, which is clearly more than the 16 demes identified using the earlier baseline (Vähä et al. 2007).

The level of genetic structuring observed within the Teno river system (mean pairwise $F_{ST} = 0.058$) exceeds those generally reported for Atlantic salmon elsewhere (e.g., 0.02 in the River Moy, Ireland (Dillane et al. 2008) and 0.014 in the River Varzuga, Russian Kola Peninsula (Primmer et al. 2006)). The genetic structure of Teno salmon is also high compared with salmon populations within larger geographical areas (e.g., an F_{ST} value of 0.02 for salmon populations in Ireland (Ensing et al. 2013), 0.03 in Scotland (Gilbey et al. 2012), and 0.04 across southern Europe (Griffiths et al. 2010)). The only geographic region with higher reported genetic divergence is the Baltic Sea area with F_{ST} at 0.12 (Koljonen 2006). As noted by Dionne et al. (2009), the overall level of genetic structuring and the distribution of genetic variation between local populations observed within the Teno river system is special in contrast with many other large river systems (cf. Primmer et al. 2006). The underlying reasons for this are unknown, but are likely linked with postglacial colonization history of the Teno salmon (cf. Verspoor et al. 2007) and catchment-specific characteristics (Vähä et al. 2007; Ozerov et al. 2012). The observed high level of genetic differentiation among populations provides an excellent basis for deploying GSI as a tool enabling stock-specific management of the Teno mixed-stock fishery.

Components of GSI success

Conforming to the expectations, the overall accuracy of GSI, estimated on simulated single sample mixtures, was high ($91\% \pm 12\%$), and 18 of the 32 baseline stocks demonstrated $>95\%$ accuracy. Illustrating the importance of genetic distinctiveness to successful stock estimation, 16 of the 17 populations with lowest pairwise genetic distance above 0.015 had assignment success higher than 95%. Unfortunately, the general scarcity of within-river studies and the lack of a GSI study of similar resolution do not allow for proper comparison, but the levels of accuracy in wider geographical surveys of Atlantic salmon stocks are commonly 20% or lower (e.g., Griffiths et al. 2010; Gilbey et al. 2012). In general, 90% accuracy levels are obtained only after defining larger regional groups (Ensing et al. 2013; Gauthier-Ouellet et al. 2009). Pacific salmon generally display higher genetic differentiation than Atlantic salmon ($F_{ST} = 0.06\text{--}0.10$; Beacham et al. 2011 and references therein). Accordingly, similar or even higher levels of GSI accuracy have been reported for Pacific salmon populations both within river systems (e.g., Beacham et al. 2004) and over wider geographical areas (e.g., Beacham et al. 2006). These accuracy levels are similar to those reported in the present study.

In our study, GSI accuracy of a population depended on both the level of genetic differentiation and the baseline sample size (Fig. 4). Genetically less distinct populations required larger baseline sample sizes than more diverged populations, a result in accordance with the findings by Beacham et al. (2011). Our mixed model for assignment success implied an optimal baseline sample size of >100 individuals for highly diverged populations (population-specific $F_{ST} > 0.07$) and >200 for less diverged populations. Mean sample size per population in the current Teno baseline was 103 (range 23–318; highest in combined Teno main stem population; Table S1), and the model estimates imply that stock estimation accuracy could markedly benefit from increasing sample size. Increasing the baseline sample size to ~ 200 individuals would nearly double the identification rate in some populations. A population such as the Akujoki ($n = 53$, $F_{ST} = 0.021$, identification rate $\sim 50\%$) may present accuracy as low as 60% of the theoretical maximum determined by the employed set of markers (see Beacham et al. 2011). On the other hand, the Teno main stem lower population with high baseline sample size ($n = 318$) showed higher assign-

ment success than expected by the level of genetic differentiation alone ($F_{ST} = 0.01$, identification rate $\sim 80\%$).

Stock estimation accuracy is critically dependent on the information content of the employed genetic markers (Banks and Jacobson 2004). While some of the markers we employed showed tendency for large allele dropout with low-quality DNA extracts or presence of null alleles, inclusion of loci deviating from the Hardy-Weinberg proportions in the baseline data have been proven not to hamper GSI (Carlsson 2008; Griffiths et al. 2010). GSI accuracy generally increases with number of loci and number of alleles per locus (Kalinowski 2004; Beacham et al. 2006). In this respect, our baseline data, which included genotypes from 33 microsatellite markers displaying more than 500 alleles in total, is large. The number of markers applied in the present study is roughly a double compared with common GSI studies, albeit depending on application and aimed resolution, even very low number of markers may be sufficient (e.g., seven employed by Ensing et al. (2013)).

Because of many practical reasons (e.g., variability, accessibility, and availability of statistical methods), microsatellites have been the marker of choice for GSI studies for a long time. Recent developments in single-nucleotide polymorphism (SNP) techniques are, however, making SNP markers more attractive also for studies where genetic information from large numbers of samples is required. Recent studies have shown SNP markers to significantly improve stock identification success in GSI studies, especially when the discrimination power with microsatellites is initially low (Ozerov et al. 2013; Moore et al. 2014). However, despite the promising results, the costs for analyzing adequate number of SNP markers is still high for GSI studies when thousands or tens of thousands of samples are to be analyzed (see Moore et al. 2014). Microsatellite markers remain a cost-effective choice, especially in systems where genetic structure and the suite of applied loci provide adequate and sufficient resolution for the requirements.

Resolution power of GSI within Teno river system

In the absence of samples with known origin that are not included in the baseline data, the accuracy of GSI is commonly estimated through simulations on equal proportion mixtures. This approach provided lower level of stock identification accuracy ($82\% \pm 19\%$) than simulated single-sample mixtures ($91\% \pm 12\%$). In accordance with the large variation in genetic differentiation and sample sizes of the baseline populations, there was a large variation among populations in the accuracy estimates, ranging from 36% for Teno main stem (TMS YK) individuals to 100% for seven different tributaries. Overall, the lowest GSI accuracies were observed for the Teno main stem and Inarijoki populations as well as for the Nilijoki and Akujoki tributaries located in the upper main stem area. It is, however, important to note that in all cases misidentified individuals were assigned to neighboring populations within a region.

In the Teno main stem, high proportions of misassignments were observed between Teno main stem lower and TMS YK and between TMS OUT and TMS YK. However, treating the main stem as a single reporting unit resulted in high identification accuracy (94.4%). Similarly, there were substantial proportions of misassignments among some of the tributaries within the Inarijoki headwater system, but the identification rate of Inarijoki salmon at the group level was 92%, with only 2.7% of individuals being misassigned from other regions. A third region, where a group level identification might more appropriate was the upper tributary group where the identification rate was 90%. Evidently, owing to high proportions of misassignments among some of the tributaries (Table 1), interpretation of stock- or tributary-specific estimates within these regions deserves caution. As discussed above, increasing the baseline sample size to ~ 200 for individual tributaries will likely allow better population-specific identifica-

tion within the upper tributaries and the Inarijoki system. Nevertheless, with the current baseline samples and the applied marker system, Teno salmon can be assigned to at least 17 geographically and genetically distinct groups with more than 90% identification rate.

Comparison of the methods

The selected statistical method may affect the accuracy of stock estimates and define the power of resolution. In accordance with Araujo et al. (2014), but in contrast with Griffiths et al. (2010), we found only small inconclusive differences between statistical methods in our simulation and resampling analyses. However, even though stock estimates were expected to be affected by the population composition of the mixture sample (since both ONCOR and cBAYES utilize this information), the most salient difference between the two methods was detected in their sensitivity to population composition of the mixture sample. At the stock level, cBAYES appeared to be slightly more sensitive than ONCOR when all stocks were present in the mixture sample. However, the changes appeared largely among populations within region, and ONCOR appeared significantly more sensitive at the region level. ONCOR significantly overestimated the contribution of Teno main stem stocks, as illustrated with mixtures including only headwater samples in which case the Teno main stem contribution was estimated at 8% by ONCOR. This outcome, in addition to slightly higher overall accuracies, indicates a preference for using cBAYES over ONCOR when analyzing mixed-stock fishery samples, an observation in agreement with Griffiths et al. (2010) but in contrast with Araujo et al. (2014). Notwithstanding, the sensitivity of stock estimates to the population composition of mixture sample warrants dividing large mixture samples into subsets composed, as far as possible, of single stocks. For example, large mixture samples can be divided into subsets based on location, time, or life-history characteristics of individuals (cf. Vähä et al. 2011).

Management implications and future prospects

The present study demonstrates that a comprehensive marker panel in combination with a baseline representation of populations enables reliable genetic identification of individual populations from mixed-stock samples. The resulting resolving power is a necessity for a sustainable stock-specific management of mixed-stock harvesting (e.g., Begg et al. 1999), reducing the risk of over-exploitation of components of the population complex and biodiversity loss (Crozier et al. 2004).

Use of the salmon resource in the Teno river system is important for the local communities, their economy, and the indigenous Sámi culture. However, stock status assessments based on biological reference points (spawning targets; Falkegård et al. 2014), following the estimation procedure described by Forseth et al. (2013), reveal poor target attainment for several populations within the Teno (Anonymous 2015). The resulting trade-off, between conservation factors pointing towards a need for reduced harvesting and socio-economic factors pointing towards continued harvesting, can only be resolved through stock-specific knowledge.

The GSI procedure presented in this study is the primary solution to this. The GSI allows estimates of stock proportions in the Teno mixed-stock fishery, which can then be used to estimate stock-specific exploitation rates, which are necessary to regulate fishing activity to sustainable levels (Thorley et al. 2007). Furthermore, the GSI reveals stock-specific information on run timing (Vähä et al. 2011), necessary for implementing temporal stock-specific fisheries restrictions. Combined with population-specific biological reference points (Falkegård et al. 2014), the GSI of the Teno mixed-stock fishery catches is an imperative requirement for future adaptive management of this diverse salmon population complex. This approach follows the guidelines of the North Atlantic Salmon Conservation Organization (NASCO), where both

abundance and diversity criteria must be considered in a precautionary approach to salmon management (NASCO 2009). In addition, estimation of population-specific spawning target attainment and cumulative fishing mortality will also enable the estimation of yearly abundance of salmon at sea prior to any fisheries (pre-fishery abundance; e.g., Potter et al. 2004). Protecting individual populations from overharvesting is required to maintain the diversity that stabilizes variance in salmon returns (cf. Schindler et al. 2010). Without conserving the roles of individual populations, the resilience that population diversity provides to fisheries will deteriorate well before the Teno salmon is extirpated.

Acknowledgements

We thank Mikhail Ozerov and two anonymous referees for valuable comments on earlier versions of the manuscript, Jari Haantia and Matti Kylmäaho for scale archive mining and scale pattern analyses, and Kristiina Haapanen for laboratory assistance. Funding was provided by the Academy of Finland (project No. 133565) and the Norwegian Environment Agency.

References

- Agashe, D. 2009. The stabilizing effect of intraspecific genetic variation on population dynamics in novel and ancestral habitats. *Am. Nat.* **174**: 255–267. doi:10.1086/600085. PMID:19519220.
- Allendorf, F.W., Luikart, G., and Aitken, S.N. 2013. Conservation and the genetics of populations. Wiley-Blackwell.
- Anderson, E.C., Waples, R.S., and Kalinowski, S.T. 2008. An improved method for predicting the accuracy of genetic stock identification. *Can. J. Fish. Aquat. Sci.* **65**(7): 1475–1486. doi:10.1139/F08-049.
- Anonymous. 2015. Status of the River Tana salmon populations 2015. Report of the Working Group on Salmon Monitoring and Research in the Tana River System [online]. Available from http://mmm.fi/documents/1410837/1801204/2015_Tenon_lohikantojen_tila_Status-of-the-river-Tana-salmon-populations.pdf/aced31a7-af81-40aa-9c82-9e6147e58134.
- Araujo, H.A., Candy, J.R., Beacham, T.D., White, B., and Wallace, C. 2014. Advantages and challenges of Genetic Stock Identification in fish stocks with low genetic resolution. *Trans. Am. Fish. Soc.* **143**: 479–488. doi:10.1080/00028487.2013.855258.
- Banks, M.A., and Jacobson, D.P. 2004. Which genetic markers and GSI methods are more appropriate for defining marine distribution and migration of salmon? North Pacific Anadromous Fish Commission Technical Note, 5: 39–42.
- Beacham, T.D., Lapointe, M., Candy, J.R., McIntosh, B., MacConnachie, C., Tabata, A., Kaukinen, K., Deng, L., Miller, K.M., and Withler, R.E. 2004. Stock identification of Fraser River sockeye salmon (*Oncorhynchus nerka*) using microsatellites and major histocompatibility complex variation. *Trans. Am. Fish. Soc.* **133**: 1117–1137. doi:10.1577/T04-001.1.
- Beacham, T.D., Candy, J.R., Jonsen, K.L., Supernault, J., Wetklo, M., Deng, L., Miller, K.M., Withler, R.E., and Varnavskaya, N. 2006. Estimation of stock composition and individual identification of Chinook salmon across the Pacific Rim by use of microsatellite variation. *Trans. Am. Fish. Soc.* **135**: 861–888. doi:10.1577/T05-241.1.
- Beacham, T.D., Winther, I., Jonsen, K.L., Wetklo, M., Deng, L., and Candy, J.R. 2008. The application of rapid microsatellite-based stock identification to management of a Chinook salmon troll fishery off the Queen Charlotte Islands, British Columbia. *N. Am. J. Fish. Manage.* **28**: 849–855. doi:10.1577/M06-167.1.
- Beacham, T.D., McIntosh, B., and Wallace, C.G. 2011. A comparison of polymorphism of genetic markers and population sample sizes required for mixed-stock analysis of sockeye salmon (*Oncorhynchus nerka*) in British Columbia. *Can. J. Fish. Aquat. Sci.* **68**(3): 550–562. doi:10.1139/F10-167.
- Begg, G.A., Friedland, K.D., and Pearce, J.B. 1999. Stock identification and its role in stock assessment and fisheries management: an overview. *Fish. Res.* **43**: 1–8. doi:10.1016/S0165-7836(99)00062-4.
- Bolnick, D.I., Amarasekare, P., Araujo, M.S., Burger, R., Levine, J.M., Novak, M., Rudolf, V.H.W., Schreiber, S.J., Urban, M.C., and Vasseur, D.A. 2011. Why intraspecific trait variation matters in community ecology. *TREE*, **26**: 183–192. PMID:21367482.
- Bradbury, I.R., Hamilton, L.C., Rafferty, S., Meerburg, D., Poole, R., Dempson, J.B., Robertson, M.J., Reddin, D.G., Bourret, V., Dionne, M., Chaput, G., Sheehan, T.F., King, T.L., Candy, J.R., and Bernatchez, L. 2015. Genetic evidence of local exploitation of Atlantic salmon in a coastal subsistence fishery in the Northwest Atlantic. *Can. J. Fish. Aquat. Sci.* **72**(1): 83–95. doi:10.1139/cjfas-2014-0058.
- Bradbury, I.R., Hamilton, L.C., Chaput, G., Robertson, M.J., Goroguer, H., Walsh, A., Morris, V., Reddin, D., Dempson, J.B., Sheehan, T.F., King, T., and Bernatchez, L. 2016. Genetic mixed stock analysis of an interceptory Atlantic salmon fishery in the Northwest Atlantic. *Fish. Res.* **174**: 234–244. doi:10.1016/j.fishres.2015.10.009.
- Carlsson, J. 2008. Effects of microsatellite null alleles on assignment testing. *J. Hered.* **99**: 616–623. doi:10.1093/jhered/esn048. PMID:18535000.
- Cavalli-Sforza, L.L., and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**: 550–570. doi:10.2307/2406616.
- Chapuis, M.P., and Estoup, A. 2007. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* **24**: 621–631. PMID:17150975.
- Crozier, W.W., Schön, P.-J., Chaput, G., Potter, E.C.E., Ó Maoiléidigh, N., and MacLean, J.C. 2004. Managing Atlantic salmon in the mixed stock environment: challenges and considerations. *ICES J. Mar. Sci.* **61**: 1344–1358. doi:10.1016/j.icesjms.2004.08.013.
- Dillane, E., McGinnity, P., Coughlan, J.P., Cross, M.C., De Eyto, E., Kenchington, E., Prodöhl, P., and Cross, T.F. 2008. Demographics and landscape features determine intrariver population structure in Atlantic salmon (*Salmo salar* L.): the case of the River Moy in Ireland. *Mol. Ecol.* **17**: 4786–4800. doi:10.1111/j.1365-294X.2008.03939.x. PMID:19140972.
- Dionne, M., Caron, F., Dodson, J.J., and Bernatchez, L. 2009. Comparative survey of within-river genetic structure in Atlantic salmon; relevance for management and conservation. *Conserv. Genet.* **10**: 869–879. doi:10.1007/s10592-008-9647-5.
- Ensing, D., Crozier, W.W., Boylan, P., Ó Maoiléidigh, N., and McGinnity, P. 2013. An analysis of genetic stock identification on a small geographical scale using microsatellite markers, and its application in the management of a mixed-stock fishery for Atlantic salmon (*Salmo salar*) in Ireland. *J. Fish Biol.* **82**: 2080–2094. doi:10.1111/jfb.12139. PMID:23731153.
- Erkinaro, J., Niemelä, E., Vähä, J.-P., Primmer, C.R., Brørs, S., and Hassinen, E. 2010. Distribution and biological characteristics of escaped farmed salmon in a major subarctic wild salmon river: implications for monitoring. *Can. J. Fish. Aquat. Sci.* **67**(1): 130–142. doi:10.1139/F09-173.
- Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online*, **1**: 47–50.
- Falkegård, M., Foldvik, A., Fiske, P., Erkinaro, J., Orell, P., Niemelä, E., Kuusela, J., Finstad, A.G., and Hindar, K. 2014. Revised first-generation spawning targets for the Tana/Teno river system. NINA Report 1087 [online]. Available from <http://www.nina.no/archive/nina/PppBasePdf/rapport/2014/1087.pdf>.
- Falush, D., Stephens, M., and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, **164**: 1567–1587. PMID:12930761.
- Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
- Foll, M., and Gaggiotti, O.E. 2006. Identifying the environmental factors that determine the genetic structure of Populations. *Genetics*, **174**: 875–891. doi:10.1534/genetics.106.059451. PMID:16951078.
- Forseth, T., Fiske, P., Barlaup, B., Gjosæter, H., Hindar, K., and Diserud, O.H. 2013. Reference point based management of Norwegian Atlantic salmon populations. *Environ. Conserv.* **40**: 356–366. doi:10.1017/S03768929130000416.
- Gamfeldt, L., Hillebrand, H., and Jonsson, P.R. 2008. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology*, **89**: 1223–1231. doi:10.1890/06-2091.1. PMID:18543617.
- Gauthier-Ouellet, M., Dionne, M., Caron, F., King, T.L., and Bernatchez, L. 2009. Spatiotemporal dynamics of the Atlantic salmon (*Salmo salar*) Greenland fishery inferred from mixed-stock analysis. *Can. J. Fish. Aquat. Sci.* **66**(12): 2040–2051. doi:10.1139/F09-147.
- Gilbey, J., Stradmeyer, L., Cauwelier, E., and Middlemas, S. 2012. Genetic investigation of the North East English net fisheries. Marine Scotland Science Report 04/12.
- Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3) [online]. Available from <http://www2.unil.ch/popgen/softwares/fstat.htm> [updated from Goudet (1995)].
- Griffiths, A.M., Machado-Schiaffino, G., Dillane, E., Coughlan, J., Horreo, J.L., Bowkett, A.E., Minton, P., Toms, S., Roche, W., Gargan, P., McGinnity, P., Cross, T., Bright, D., Garcia-Vazquez, E., and Stevens, J.R. 2010. Genetic stock identification of Atlantic salmon (*Salmo salar*) populations in the southern part of the European range. *BMC Genet.* **11**: 31. doi:10.1186/1471-2156-11-31. PMID:20429926.
- Grimholt, U., Drablos, F., Jorgensen, S.M., Høyheim, B., and Stet, R.J.M. 2002. The major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics*, **54**: 570–581.
- Habicht, C., Seeb, L.W., Myers, K.W., Farley, E.V., and Seeb, J.E. 2010. Summer-fall distribution of stocks of immature sockeye salmon in the Bering Sea as revealed by single-nucleotide polymorphisms (SNPs). *Trans. Am. Fish. Soc.* **139**: 1171–1191. doi:10.1577/T09-149.1.
- Harris, L.N., and Taylor, E.B. 2010. Genetic population structure of broad whitefish, *Coregonus nasus*, from the Mackenzie River, Northwest Territories: implications for subsistence fishery management. *Can. J. Fish. Aquat. Sci.* **67**(6): 905–918. doi:10.1139/F10-027.
- Hess, J.E., Whiteaker, J.M., Fryer, J.K., and Narum, S.R. 2014. Monitoring stock-specific abundance, run timing, and straying of Chinook salmon in the Columbia River using genetic stock identification (GSI). *N. Am. J. Fish. Manage.* **34**: 184–201. doi:10.1080/02755947.2013.862192.
- Hilborn, R., Quinn, T.P., Schindler, D.E., and Rogers, D.E. 2003. Biocomplexity

- and fisheries sustainability. *Proc. Nat. Acad. Sci.* **100**: 6564–6568. doi:10.1073/pnas.1037274100. PMID:12743372.
- Huson, D.H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**: 254–267. PMID:16221896.
- Hutchinson, W.F. 2008. The dangers of ignoring stock complexity in fishery management: the case of the North Sea cod. *Biol. Lett.* **4**: 693–695. doi:10.1098/rsbl.2008.0443. PMID:18782730.
- ICES. 2015. Report of the Working Group on North Atlantic Salmon (WGNAS). ICES CM 2015, ACOM:09.
- Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S., Reich, P.B., Scherer-Lorenzen, M., Schmid, B., Tilman, D., van Ruijven, J., Weigelt, A., Wilsey, B.J., Zavaleta, E.S., and Loreau, M. 2011. High plant diversity is needed to maintain ecosystem services. *Nature*, **477**: 199–202. doi:10.1038/nature10282.
- Kalinowski, S.T. 2004. Genetic polymorphism and mixed-stock fisheries analysis. *Can. J. Fish. Aquat. Sci.* **61**(7): 1075–1082. doi:10.1139/f04-060.
- Kalinowski, S.T., Manlove, K.R., and Taper, M.L. 2007. ONCOR: a computer program for Genetic Stock Identification [online]. Department of Ecology, Montana State University, Bozeman, MT 59717, USA. Available from <http://www.montana.edu/kalinowski>.
- Koljonen, M.-L. 2006. Annual changes in the proportions of wild and hatchery Atlantic salmon (*Salmo salar*) caught in the Baltic Sea. *ICES J. Mar. Sci.* **63**: 1274–1285. doi:10.1016/j.icesjms.2006.04.010.
- Liu, K., and Muse, S.V. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, **21**: 2128–2129. doi:10.1093/bioinformatics/bti282. PMID:15705655.
- Mills, L.S. 2012. Conservation of wildlife populations. Demography, genetics and management. Wiley-Blackwell.
- Milner, G.B., Teel, D.J., Utter, F.M., and Winans, G.A. 1985. A genetic method of stock identification in mixed populations of Pacific salmon, *Oncorhynchus* spp. *Mar. Fish. Rev.* **47**: 1–8.
- Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., Chaput, G., and Bernatchez, L. 2014. Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. *Mol. Ecol.* **23**: 5680–5697. doi:10.1111/mec.12972. PMID:25327895.
- NASCO. 2009. Guidelines for the management of salmon fisheries. North Atlantic Salmon Conservation Organization (NASCO), Edinburgh, Scotland, UK. NASCO Council Document CNL 43.
- Neaves, P.I., Wallace, J.R., Candy, J.R., and Beacham, T.D. 2005. CBayes: computer program for mixed stock analysis of allelic data, version v4.02 [online]. Available from pac.dfo-mpo.gc.ca/sci/mlg/CBayes_e.htm.
- Nehls, W., Williams, J.E., and Lichatowich, J.A. 1991. Pacific Salmon at the Crossroads: Stocks at Risk from California, Oregon, Idaho, and Washington. *Fisheries*, **16**: 4–21. doi:10.1577/1548-8446(1991)016<0004:PSATCS>2.0.CO;2.
- Niemelä, E., Erkinaro, J., Julkunen, M., Hassinen, E., Lämsä, M., and Brørs, S. 2006. Temporal variation in abundance, return rate and life histories of previously spawned Atlantic salmon in a large subarctic river. *J. Fish Biol.* **68**: 1222–1240. doi:10.1111/j.0022-1112.2006.001012.x.
- Northrup, S., Connor, M., and Taylor, E.B. 2010. Population structure of lake trout (*Salvelinus namaycush*) in a large glacial-fed lake inferred from microsatellite DNA and morphological analysis. *Can. J. Fish. Aquat. Sci.* **67**(7): 1171–1186. doi:10.1139/F10-054.
- Ozerov, M.Y., Veselov, A.E., Lumme, J., and Primmer, C.R. 2012. “Riverscape” genetics: river characteristics influence the genetic structure and diversity of anadromous and freshwater Atlantic salmon (*Salmo salar*) populations in northwest Russia. *Can. J. Fish. Aquat. Sci.* **69**(12): 1947–1958. doi:10.1139/f2012-114.
- Ozerov, M., Vasemägi, A., Wennevik, V., Diaz-Fernandez, R., Niemelä, E., Prusov, S., Kent, M.P., and Vähä, J.-P. 2013. Finding markers that make a difference: DNA pooling and SNP-arrays identify population informative markers for genetic stock identification. *PLOS ONE*, **8**: e82434. doi:10.1371/journal.pone.0082434. PMID:24358184.
- Paterson, S., Piertney, S.B., Knox, D., Gilbey, J., and Verspoor, E. 2004. Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Mol. Ecol. Notes*, **4**: 160–162. doi:10.1111/j.1471-8286.2004.00598.x.
- Pella, J., and Masuda, M. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. *Fish. Bull.* **99**: 151–167.
- Potter, E.C.E., Crozier, W.W., Jan-Schön, P.-J., Nicholson, M.D., Maxwell, D.L., Prevost, E., Erkinaro, J., Gudbergsson, G., Karlsson, L., Hansen, L.P., Maclean, J.C., O'Maoileidigh, N., and Prusov, S. 2004. Estimating and forecasting pre-fishery abundance of Atlantic salmon in the Northeast Atlantic for the management of mixed stock fisheries. *ICES J. Mar. Sci.* **61**: 1359–1369. doi:10.1016/j.icesjms.2004.08.012.
- Primmer, C.R., Veselov, A.E., Zubchenko, A., Poututkin, A., Bakhmet, I., and 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. *Mol. Ecol.* **15**: 653–666. doi:10.1111/j.1365-294X.2005.02844.x. PMID:16499692.
- Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945–959. PMID:10835412.
- Quinn, T.P. 1993. A review of homing and straying of wild and hatchery-produced salmon. *Fish. Res.* **18**: 29–44. doi:10.1016/0165-7836(93)90038-9.
- Rannala, B., and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. U.S.A.* **94**: 9197–9201. doi:10.1073/pnas.94.17.9197. PMID:9256459.
- Rosenberg, M.S., and Anderson, C.D. 2011. PASSAGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods Ecol. Evol.* **2**: 229–232. doi:10.1111/j.2041-210X.2010.00081.x.
- Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**: 1219–1228. PMID:9093870.
- Saitou, N., and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425. PMID:3447015.
- Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A., and Webster, M.S. 2010. Population diversity and the portfolio effect in an exploited species. *Nature*, **465**: 609–612. doi:10.1038/nature09060. PMID:20520713.
- Swattdipong, A., Vasemägi, A., Niva, T., Koljonen, M.-L., and Primmer, C.R. 2013. Genetic mixed-stock analysis of lake-run brown trout *Salmo trutta* fishery catches in the Inari Basin, northern Finland: implications for conservation and management. *J. Fish Biol.* **83**: 598–617. doi:10.1111/jfb.12199. PMID:23991877.
- Thorley, J.L., Youngson, A.F., and Laughton, R. 2007. Seasonal variation in rod recapture rates indicates differential exploitation of Atlantic salmon, *Salmo salar*, stock components. *Fish. Manage. Ecol.* **14**: 191–198. doi:10.1111/j.1365-2400.2007.00540.x.
- Tilman, D., Wedin, D., and Knops, J. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, **379**: 718–720. doi:10.1038/379718a0.
- Vähä, J.-P., and Primmer, C.R. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**: 63–72. PMID:16367830.
- Vähä, J.-P., Erkinaro, J., Niemelä, E., and Primmer, C. 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Mol. Ecol.* **16**: 2638–2654. doi:10.1111/j.1365-294X.2007.03329.x. PMID:17594436.
- Vähä, J.-P., Erkinaro, J., Niemelä, E., and Primmer, C. 2008. Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. *Evol. Appl.* **1**: 137–154. doi:10.1111/j.1752-4571.2007.00007.x. PMID:25567497.
- Vähä, J.-P., Erkinaro, J., Niemelä, E., Saloniemi, I., Primmer, C.R., Johansen, M., Svenning, M., and Brørs, S. 2011. Temporally stable population-specific differences in run timing of one-sea-winter Atlantic salmon returning to a large river system. *Evol. Appl.* **4**: 39–53. doi:10.1111/j.1752-4571.2010.00131.x. PMID:25567952.
- Vähä, J.-P., Wennevik, V., Ozerov, M., Diaz-Fernandez, R., Unneland, L., Haapanen, K., Niemelä, E., Svenning, M., Falkegård, M., Prusov, S., Lyzhov, I., Rysakova, K., Kalske, T., Christiansen, B., and Ustyuzhinsky, G. 2014. Genetic structure of Atlantic salmon in the Barents region and genetic stock identification of coastal fishery catches from Norway and Russia [online]. Kolarctic ENPI CBC — Kolarctic salmon project (KO197) report. Available from www.fylkesmenn.no/kolarcticsalmon.
- Verspoor, E., Stradmeyer, L., and Nielsen, J.L. (Editors). 2007. The Atlantic salmon: genetics, conservation and management. Blackwell.
- Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**: 1358–1370. doi:10.2307/2408641.
- Youngson, A.F., Jordan, W.C., Verspoor, E., McGinnity, P., Cross, T., and Ferguson, A. 2003. Management of salmonid fisheries in the British Isles: towards a practical approach based on population genetics. *Fish. Res.* **62**: 193–209. doi:10.1016/S0165-7836(02)00162-5.