Scottish Government
Riaghaltas na h-Alba

Exploring the Potential for Genetic Stock Identification of Atlantic Salmon from the large Scottish East Coast Rivers

E Cauwelier, J Gilbey, P Mollahan, J Sampayo and DJ Morris

Exploring the Potential for Genetic Stock Identification of Atlantic Salmon from the large Scottish East Coast Rivers

E Cauwelier, J Gilbey, P Mollahan, J Sampayo and DJ Morris

Published by the Marine Directorate of the Scottish Government

The Marine Directorate of the Scottish Government is responsible for the integrated management of Scotland's seas. Scottish Marine and Freshwater Science is a series of reports that publishes the results of scientific research and monitoring carried out by the Marine Directorate. It also publishes the results of marine and freshwater scientific work that has been carried out for the Marine Directorate under external commission. These reports are not subject to formal external peer-review.

This report presents the results of marine and freshwater scientific work carried out by the Marine Directorate.

This report should be quoted as:
Cauwelier, E., Gilbey, J., Mollahan, P., Sampayo, J. and Morris, D.J. (2024). Exploring the potential for genetic stock identification of Atlantic salmon from the large Scottish East Coast Rivers.

Exploring the Potential for Genetic Stock Identification of Atlantic salmon from the large Scottish East Coast Rivers

E Cauwelier, J Gilbey, P Mollahan, J Sampayo and DJ Morris*

Marine Directorate, Freshwater Laboratory
Faskally, Pitlochry, PH16 5JS
*corresponding author: E Cauwelier

Abstract

Mortality at sea is believed to be a major cause of progressive decadal declines in numbers of Atlantic salmon returning to rivers across Scotland. Different salmon populations may be distributed differently at sea. An ability to map out the distributions of river populations at sea would aid planning of new marine developments and management of existing marine activities to reduce potential impacts on particularly important salmon stocks. Recently, it has been possible to capture groups of young salmon at sea off the east coast of Scotland as they migrate to high seas feeding grounds. This project tested whether it is possible to sample the genetic constitution of such fish to work out which rivers they came from. In particular, it was investigated whether it is possible to distinguish fish from among the large east coast rivers of Scotland, which has not been feasible with genetics methods used previously. Unfortunately, even with more advanced genetics methods used in this project it was feasible only to distinguish salmon from certain rivers and groups of rivers from one another.

Executive summary

Continuing declines in adult Atlantic salmon (*Salmo salar*) returning to rivers across Scotland have been recorded. The Scottish Wild Salmon Strategy identifies multiple pressures as to the cause of this decline, with climate change a possible contributor as an overall driver, potentially resulting in increased mortality in the marine environment. These declines vary among rivers, possibly because different salmon populations may use the marine environment in different ways. Understanding these differences is important for planning and regulatory purposes, as any riverine differences in migration pathways have the potential to result in stock-specific differences in ecosystem impacts and interactions with offshore developments (e.g. marine renewable energy generation installations). The development of tools to allow insights into stock-specific migration routes is important for conservation

efforts and will inform local river managers, recreational and commercial fisheries, as well as both marine renewables and climate change mitigation planning under the Scottish Governments Blue Economy programme for sustainable marine development.

One approach to assess ocean use by salmon is the utilisation of genetic markers to determine the river of origin of salmon captured in the marine environment. This genetic stock identification (GSI) method has become an integral component of modern fisheries management, but requires knowledge of the genetic structure underlying the various populations that make up a stock. Advances in DNA profiling have allowed the development of GSI, resulting, in many cases, in increased geographic resolution and assignment of fish caught at sea to river of origin. In Scotland, although the salmon originating from some individual rivers could be identified using existing GSI protocols, these were the exception. The countrywide resolution that can be achieved using the available genetic markers is limited, in most cases, to regional assignment units that contain a number of neighbouring rivers.

Along the Scottish east coast, the main regional assignment unit stretches from the River Spey to the River Tay and, as such, contains some of the most productive salmon rivers in the country. It would therefore be of significant value to further develop GSI techniques in this region with a view to being able to robustly distinguish between salmon populations from the different rivers within it. This is particularly important to support on-going work using trawled salmon smolts to assess potential impacts of marine renewable energy developments.

Recently, a panel of genetic markers has been developed and successfully utilised in Canada to enhance GSI resolution in the western Atlantic. The aim of this project was to establish the utility of this panel in a Scottish context, focusing on developing the panel along the East coast of Scotland, with a view to determining if the same increase in GSI resolution could be achieved in this important area. To that purpose, 915 juveniles from 40 sites across nine East coast rivers were screened using this marker panel.

Genetic structuring analysis revealed the salmon to be clustered into four distinct groups: the first encompassed the Oykel/Cassley/Shin river system, the second comprised the River Tweed, a third included sites from the Rivers Don and Dee, and the fourth cluster consisted of the remaining east coast sites. Attempts to identify the rivers of origin of fish using this new marker

panel revealed poor accuracy at the river level. Increasing accuracy by combining geographically close rivers in a sequential hierarchical process resulted in accurate assignment to three regional units: the Oykel/Cassley/Shin river system, the River Tweed and the largest unit consisting of the remaining screened NE coast rivers. These were similar to the GSI resolution previously obtained with existing marker panels, despite identifying underlying population genetic structuring within the region.

Contents

Introduction

Continuing declines in Atlantic salmon numbers across Scotland have been associated with an increased mortality in the marine environment (Friedland, 1998; Potter & Crozier, 2000). Trends in mortality appear to vary among rivers, possibly because different populations of salmon use the marine environment in different ways (Davidson *et al.*, 2009; Thorstad *et al.*, 2011, Gilbey *et al.*, 2021). Understanding this variation is important for conservation efforts and will inform local managers, recreational and commercial fisheries and, at a national level, both marine renewables planning and climate change mitigation.

A number of projects are currently underway to understand the migratory routes of post-smolt salmon around Scotland, including acoustic tracking of movements and identification of distributions by trawling fish. The information gained can be used to inform plans for marine renewable energy developments and help mitigate against the impacts of other anthropogenic pressures. Further research is being carried out in the Norwegian Sea and near western Greenland in relation to impacts of climate change and ecosystem disruptions in these distant but important feeding areas for salmon. These projects include the Norwegian led 'SeaSalar' project which focuses on salmon in the Norwegian and Barents Seas, Canadian investigations of salmon around west Greenland and Marine Scotland surveys off the east coast of Scotland and Moray Firth. In each of these cases, genetic markers are being used to determine the origin of salmon in the areas to examine impacts at a regional and/or river-specific level.

Genetic stock identification (GSI) is an integral component of modern fisheries management (Begg *et al.*, 1999; Beacham *et al.*, 2021). To manage fish species successfully, it is important to understand the genetic structure underlying the various populations that make up a stock. Advances in DNA profiling and associated analytical techniques have allowed the development of GSI using different types of genetic markers (Waples *et al.*, 1990, 2008). Initially, allozymes and mitochondrial DNA were successfully used for stock identification in salmonids (Shaklee *et al.*, 1999; Moriya *et al.*, 2007; Koljonen & McKinnell, 1996). More recently, panels of highly polymorphic microsatellite markers and single nucleotide polymorphisms (SNPs) have allowed stock identification to be successfully performed with Atlantic salmon at a number of geographic scales, from inter-continental to intra-river (Gilbey *et al.*, 2005; Griffiths *et al.*, 2010, 2011; Vähä *et al.*, 2011). In Scotland, both a

microsatellite baseline (Gilbey *et al.*, 2018) and a single nucleotide polymorphism (SNP) baseline (Gilbey *et al.*, 2016a) have been developed, both of which allowed accurate assignment to regions. However, in most cases these lack the resolution needed to allow reliable assignment to the individual river level.

The development of new marker types and the advancement of analytical tools have resulted in increased geographic resolution of assignment. For example, in Scotland, the 288 SNP panel showed higher geographic resolution compared to the 14 microsatellite markers previously used (Gilbey *et al.*, 2016a, 2018). The assignment units defined by the latter, developed to encompass all eastern Atlantic Ocean countries harbouring Atlantic salmon populations, could, at the lowest level, distinguish between five regions in Scotland (Central, Cromarty Firth, NE, NW and West Central), as well as the individual River Leven and Water of Luce (Gilbey, *et al.*, 2018). The SNP baseline, on the other hand, identified thirteen assignment units, in some cases further subdividing the units identified by the microsatellites. However, a number of assignment units still encompassed large geographic areas, most notably the East coast and North and West regions (Gilbey *et al.*, 2016a).

In Gilbey *et al.* (2016a), the North and West assignment region of Scotland covered a large proportion of the country. However, river coverage was limited and the west coast is typified by a large number of relatively small rivers. The East coast assignment unit, on the other hand, included many of the largest rivers in the country and those rivers are the major contributors (~60%) to the overall Scottish wild salmon production (based on 2020 rod catch returns – Marine Scotland Science) and include nine of the seventeen Special Areas for Conservation (SACs) for the species in Scotland, under the EU Habitats Directive. A mixed stock fishery analysis of adult salmon caught in the coastal nets along the English NE coast revealed that approximately 40% of these fish were assigned to the Scottish NE region (Gilbey *et al.*, 2016b), whilst this varied between ~20% and ~40% of returning adults captured at Armadale in the far north (Cauwelier *et al.*, 2016; Armstrong *et al.*, 2018). Given both the importance of large Scottish East coast rivers towards salmon productivity and conservation (including those designated as SACs) and the ongoing development of this coastal area for marine renewable energy, it is highly desirable to be able to distinguish salmon populations from the different rivers within this region in order to obtain more accurate GSI results and, thus, allowing patterns of ocean utilisation to be examined at the river level.

Recently, a new panel of 101 microsatellite markers has been developed and successfully utilised in Canada (Bradbury *et al.*, 2018; Sylvester *et al.*, 2018) to obtain river-level resolution of Atlantic salmon, where previous marker sets identified larger regions.

The aim of this project was to establish the utility of this panel in a Scottish context, focusing on developing the panel along the East coast of Scotland. The goal was to establish a new genetic reference baseline and develop improved GSI resolution. The baseline was built with samples collected as part of the National Electrofishing Programme for Scotland (NEPS) project (Malcolm *et al.*, 2020). If higher robust assignments at an increased resolution could be obtained with this panel, then it would be used on contemporary samples of smolts caught by trawling off the Scottish east coast by Marine Scotland during the period 2017-19 for planning Marine Renewable developments. If possible, the analysis of the results of this screening would allow stock-specific patterns of migration and feeding to be examined and, hence, would provide invaluable information on ocean utilisation in the individual river stocks.

Methods

Sample collection

Under the National Electrofishing Programme for Scotland (NEPS) programme, genetic data were derived from 1,040 juvenile salmon electrofished in 2018 and 2019 from 52 point locations covering six rivers along the NE coast of Scotland (Spey - SAC, Deveron, Don, Dee - SAC, North Esk and Tweed - SAC), as well as a single site from the River Tay (SAC) and three sites from the Oykel/Cassley/Shin river system (SAC) (Fig. 1); the latter being considered the genetic outgroup. A tail fin clip was taken from anaesthetized fish, the fish released alive and the clips were stored in 99% ethanol.

Samples were combined at geographically close sites which had low numbers of samples. This resulted in a final sample size of 1040 fish, screened at 40 sites.

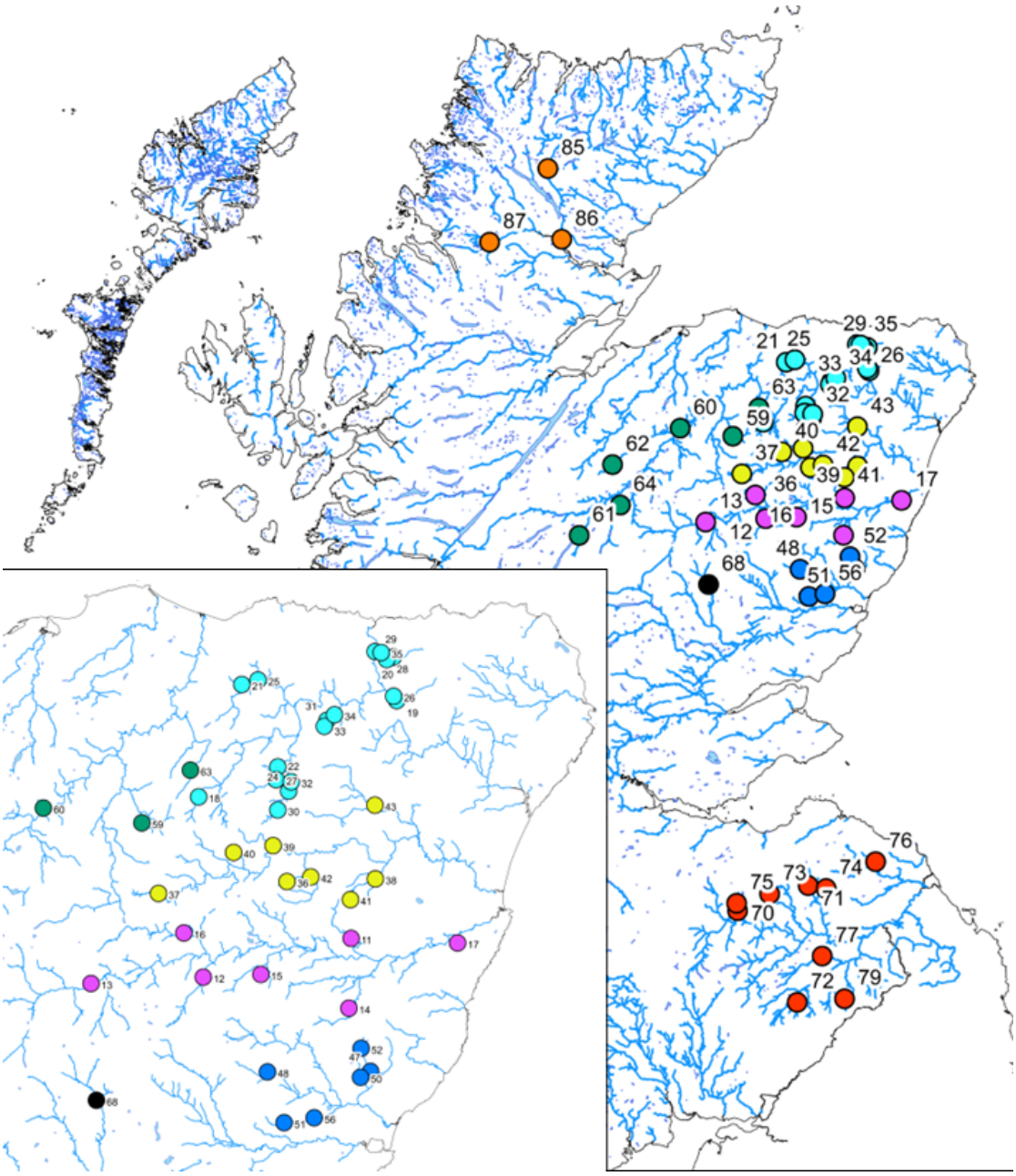

Figure 1. Geographic location of the sample sites included in this study, with each river represented by a different colour (1:2000000). Oykel (orange), Spey (green), Deveron (light blue), Don (yellow), Dee (pink), North Esk (dark blue), Tay (black) and Tweed (red). Insert focuses on the sites sampled in the Rivers Deveron, Don, Dee and North Esk. Numbers relate to site number in Table 1.

Genetic analysis

Microsatellites with a scoring percentage below 50% were removed from further analysis. Similarly, individuals with a scoring percentage below 50% post loci-removal were also removed. Lastly, full sibs in samples were

identified by COLONY2 (Jones & Wang, 2010) using the pedigree likelihood approach, assuming biparental polygamy and no inbreeding. Only a single member from each full sib family identified was retained in order to avoid inflating genetic differences among samples through family effects (Hansen *et al.*, 1997).

Genalex (Peakall and Smouse, 2006) was used to calculate the various diversity and differentiation parameters, such as site-specific observed (H_o) and expected (H_e) heterozygosity, and pairwise estimates of F_{ST} and D_A (Nei *et al.*, 1983). Hardy–Weinberg (HW) proportions and Linkage Disequilibrium (LD) were tested using the GENEPOP package (Rousset, 2008) in R (R Core Team, 2015) and a False Discovery Rate (FDR) correction (Benjamini & Hochberg, 1995) applied for multiple tests. The same package was also used to test for significance of the pairwise F_{ST} values. Allelic richness (A_r) was determined with HP-RARE (Kalinowski, 2005), standardized to a sample size of 20.

Population relatedness was assessed with STRUCTURE 2.3.4 (Pritchard *et al.*, 2000, 2004) using the admixture model assuming correlated allele frequencies (Falush *et al.*, 2003). STRUCTURE was run with a burn-in and run phase of 100,000 and 300,000 iterations, respectively, for five replicates for each number of clusters (K) and increasing K values until LnP(K) plateaued. The log-likelihood probability (Ln(K)) and ΔK (Evanno *et al.*, 2005) were calculated with STRUCTURE HARVESTER (Earl & vonHoldt, 2011) to identify the smallest K that captured the main structure in the data (Pritchard *et al.*, 2000, 2004). Individual membership coefficients were combined across replicates with CLUMPP1.1.2 (Jakobsson & Rosenberg, 2007), employing the FullSearch method with random order input (1,000 repeats) and results were visualised in DISTRUCT1.1 (Rosenberg, 2004).

Assignment analysis

All fish present in the dataset after Quality Control (QC) formed the baseline to which the assignment analysis was carried out. It was performed using the R package Rubias (Anderson, 2017) and followed the self-assignment model using a leave-one-out approach. In this approach, a single fish was taken out of the baseline at random and assigned back to it. The most likely river of origin was estimated with an associated probability. This procedure was repeated until all fish had been assigned. Assignment accuracy (i.e. how many fish assigned to a particular river/assignment unit actually originated from that river/assignment unit) was calculated.

When performing the self-assignments, a probability threshold of 0.8 was used. Thus, only fish with assignments with probabilities at or above this level were included in the analysis. Using this threshold removed fish for which it was difficult to robustly estimate river of origin, which has previously been shown to significantly increase assignment accuracy (Gilbey, *et al.*, 2016a).

Successful assignment accuracy to individual rivers or larger assignment units was defined as at or above 80% accuracy. That is, 80% or more of the fish assigned to a river or assignment unit were actually from this unit. If this condition was met for a particular river, then this river was considered to be a single assignment unit. In cases where this condition was not met, the pattern of misassignments was examined and neighbouring rivers were combined into a larger assignment unit, based on geography. The assignment analysis, as described above, was repeated, examining the results at the newly defined assignment levels. Results were assessed against the accuracy threshold and, if needed, further geographic regions were combined. This procedure was repeated until each assignment unit fulfilled the condition of at least 80% assignment accuracy to all assignment units.

Results

Quality Control

After the various QC measures outlined above were taken, combined with the full sib analysis, 125 individuals were removed, the final dataset contained genotypes for 69 microsatellites of 915 individuals from 40 sites, with an average N = 23 individuals per site. The average number of alleles per marker was 9.4 (± 4.2).

Genetic diversity

Details on within-site genetic diversity are given in Table 1, including average values across sites, together with the standard deviation (SD). After correction for multiple tests, genotypes in thirteen sites deviated significantly from the HW equilibrium.

Table 1
Site-specific genetic diversity parameters. N: number of fish, A_r: Allelic richness, H_o: Observed heterozygosity, H_e: Expected heterozygosity, F:

fixation index related to levels of inbreeding and p*HW*: probability associated with genotypes being in Hardy-Weinberg (HW) proportions. Mean and standard deviation (SD) are also presented. Significant deviations from HW equilibrium are shown in italics. Sites with multiple numbers in the first column were combined.

Site/s	Sites (NEPS or combined)	N	A_r	H_o	H_e	F	pHW
		2	3.3	0.4	0.5		
85	KSFT_Shin_5	6	3	9	0	0.04	0.070
		2	3.5	0.5	0.5	-	
86	KSFT_Shin_7	9	4	3	2	0.02	0.008
		4	3.3	0.5	0.5		<
87	Kyle_Sutherland_2523	7	5	0	2	0.03	*0.0001*
		4	3.8	0.5	0.5		<
59	Spey_001	6	8	2	5	0.07	*0.0001*
		2	4.0	0.5	0.5		<
60	Spey_019	4	7	4	6	0.05	*0.0001*
		2	3.7	0.5	0.5		
61	Spey_024	9	1	3	4	0.03	0.013
		2	3.7	0.5	0.5		
62	Spey_060	7	6	4	5	0.02	0.753
		2	3.8	0.5	0.5		
63	Spey_138	9	7	2	4	0.05	0.009
		2	3.7	0.5	0.5		
64	Spey_139	8	6	4	4	0.01	0.094
		1	3.5	0.5	0.5	-	
18	Deveron_1623	5	3	3	2	0.02	0.992
		1	4.1	0.5	0.5		
19, 26	Deveron_B	7	7	4	5	0.01	0.334
		2	3.9	0.5	0.5		<
20, 28, 29, 35	Deveron_low	2	9	2	5	0.10	*0.0001*
		1	3.7	0.5	0.5	-	
31, 33, 34	Deveron_mid	8	7	5	4	0.02	0.585
22, 24, 27, 30,		1	3.3	0.5	0.4	-	
32	Deveron_upmost	3	5	1	9	0.05	0.643
		2	3.8	0.5	0.5		
21, 25	Deveron_upper	4	7	3	4	0.02	0.417
		3	4.5	0.4	0.6		<
37	Don_1809	1	5	8	5	0.26	*0.0001*
		2	3.9	0.5	0.5		
39	Don_1827	6	5	4	5	0.03	0.010
		2	3.5	0.4	0.4		<
40	Don_1844	3	2	2	7	0.19	*0.0001*
		1	3.1	0.4	0.3	-	<
43	Don_1934	1	2	3	7	0.08	*0.0001*
		1	3.0	0.2	0.2	-	
38, 41	Don_lower	6	0	6	5	0.05	0.024

		1	3.5	0.3	0.4		<
36, 42	Don_middle	6	1	9	5	0.20	*0.0001*
		1	4.1	0.4	0.5		<
11	Dee_186	9	7	6	7	0.20	*0.0001*
		4	3.9	0.5	0.5		<
12	Dee_190	7	3	2	4	0.07	*0.0001*
		1	3.6	0.5	0.5	-	
13	Dee_195	5	5	4	2	0.02	0.996
		3		0.5	0.5		<
14	Dee_313	9	3.7	0	3	0.07	*0.0001*
		2	4.6	0.4	0.6		<
15	Dee_314	9	5	8	3	0.25	*0.0001*
		2	3.7	0.5	0.5		
16	Dee_316	9	1	3	3	0.04	0.014
		1	5.2	0.3	0.2	-	
17	Dee_345	4	7	3	7	0.19	0.073
		2	3.7	0.5	0.5		
48	Esk_2000	8	9	4	3	0.00	0.334
		1	3.6	0.5	0.5	-	
51	Esk_2027	0	8	3	2	0.04	0.801
		1	4.3	0.5	0.6		<
56	Esk_2111	5	2	2	0	0.14	*0.0001*
		1	4.0	0.5	0.5		
47, 50, 52	Esk_Luther	6	7	1	5	0.07	0.003
		2	3.6	0.5	0.5	-	
68	Tay_4106	4	3	5	3	0.04	0.003
		2	3.7	0.5	0.5		
72	Tweed_4148	1	1	2	2	0.01	0.295
		2	3.4	0.5	0.5		
73	Tweed_4153	1	3	0	1	0.04	0.036
		1	3.3	0.5	0.4	-	
76	Tweed_4159	2	4	1	8	0.06	0.997
		1	3.6	0.5	0.5		
77	Tweed_4160	4	9	3	3	0.00	0.118
		1		0.5	0.5	-	
79	Tweed_4276	6	3.7	3	2	0.01	0.797
		1	3.6	0.5	0.5		
71, 74	Tweed_Leader	6	5	0	1	0.05	0.148
		1	3.6	0.5	0.5	-	
70, 75	Tweed_upper	3	5	2	1	0.01	0.826
		2	**3.7**	**0.5**	**0.5**		
	Mean	**3**	**8**	**0**	**2**	**0.04**	
		1	**0.4**	**0.0**	**0.0**		
	SD	**0**	**2**	**6**	**7**	**0.09**	

Genetic structure

Supplementary table S1 details the differentiation parameters F_{ST} and D_A. Average and SD values for those were 0.096 ± 0.122 and 0.317 ± 0.465, respectively. 507 out of the 780 (65%) pairwise comparisons were significant after FDR correction.

The optimal number of clusters in the dataset was K = 4. At that level, the sites from the rivers Oykel/Cassley/Shin and those from the River Tweed each form their own cluster (1: purple and 2: dark green in Fig. 2A, respectively). The majority of the sites from the River Don and a few from the River Dee (Dee_186, Dee_314 and Dee_345) are grouped in a third cluster (light green in Fig. 2A). The remainder of the sites encompassed the fourth group (lila in Fig. 2A). A geographical presentation of the clustering results is also given in Figure 2B.

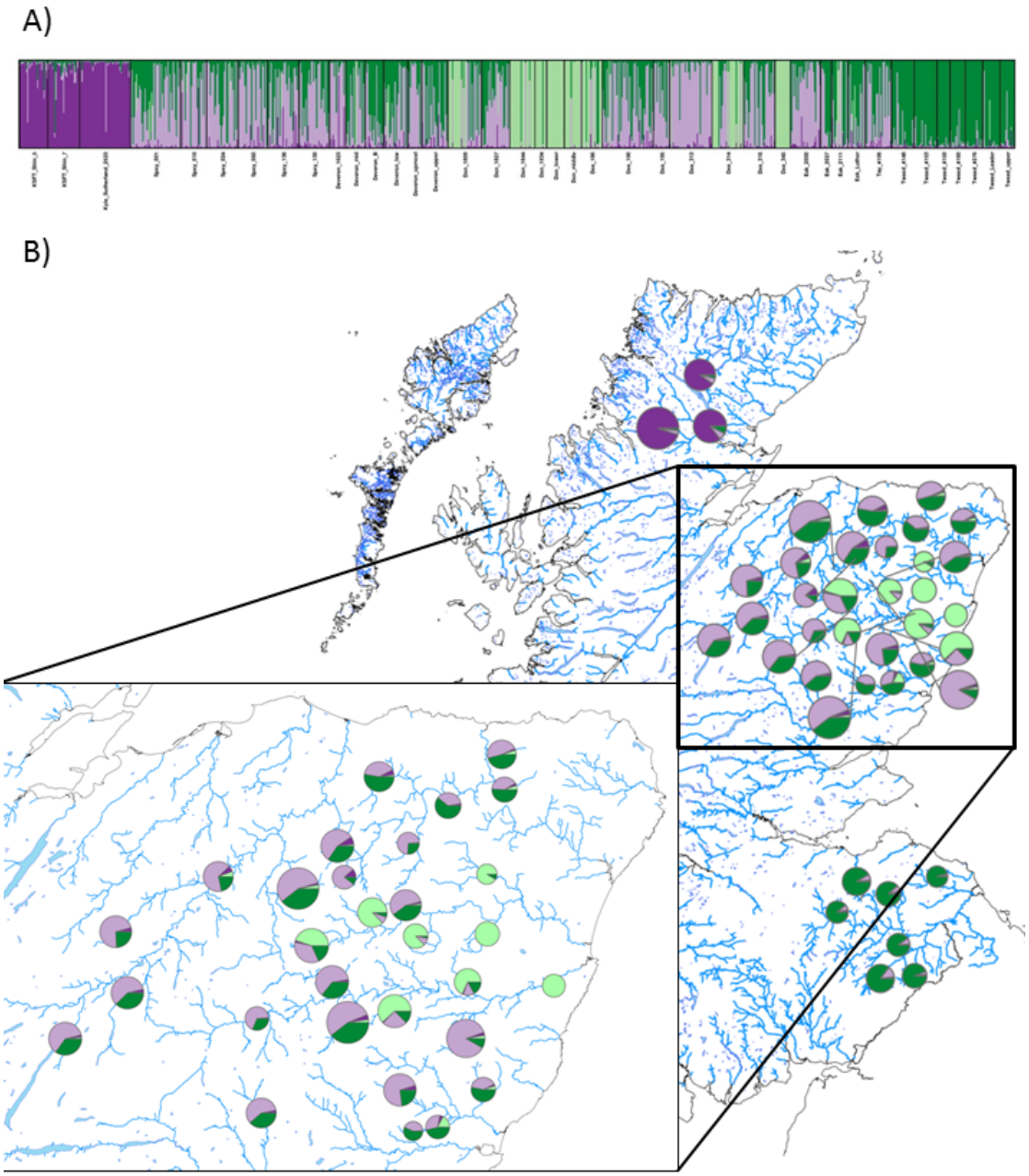

Figure 2. A) Genetic clustering of individuals/sites into four groups as revealed by STRUCTURE. Each vertical line represents an individual and sites are delineated by a black line. Different clusters are indicated by different colours and individual membership to each of these clusters is given by those colours. B) Pie charts depicting the results of the clustering analysis to provide geographic representation of the various clusters. Bottom right: Enlargement of area highlighted by square on the bottom left.

Self-assignment

At the first river-level and based on all fish (i.e. not using a probability cut-off), assignment accuracy varied from 33.3% for the River North Esk to 94.1% for the River Oykel (Table 2). After applying the 0.8 probability cut-off level, assignment accuracies increased. However, only the rivers Oykel/Cassley/Shin and Tweed met the required assignment accuracy condition (80% accuracy). These two rivers were thus retained as separate assignment units. The misassignments, together with geographic proximity, were used to combine low accuracy rivers into larger assignment units. This was carried out by combining the adjacent rivers Spey and Deveron into a new unit (SpeyDev); and the Don and Dee into another new unit (DonDee). Self-assignment was then performed to the new Level 2 assignment units, using the 0.8 probability cut-off.

Results of the first, river-level, assignment analysis. Table 2 presents the results of all fish assignments, whilst Table 3 includes the results when a probability cut-off of 0.8 is applied. Both accuracy and number of fish assigned for each river is shown in bold.

Table 2

All assignments	**Assigned to**							
River of origin	**Oykel/Cassley/Shin**	**Spey**	**Deveron**	**Don**	**Dee**	**North Esk**	**Tay**	**Tweed**
Oykel/Cassley/Shin	111	2	1	1	0	2	0	1
Spey	1	132	12	14	21	16	2	4
Deveron	1	41	29	8	19	10	1	4
Don	0	20	3	98	9	1	3	6
Dee	3	52	8	63	75	19	2	7
North Esk	2	25	3	12	12	26	2	5
Tay	0	6	0	1	1	1	16	0
Tweed	0	14	6	11	7	3	1	84
Accuracy %	**94.1**	**45.2**	**46.8**	**47.1**	**52.1**	**33.3**	**59.3**	**75.7**

Table 3

Prob cut-off >0.8	**Assigned to**							
River of origin	**Oykel/Cassley/Shin**	**Spey**	**Deveron**	**Don**	**Dee**	**North Esk**	**Tay**	**Tweed**
Oykel/Cassley/Shin	107	2	1	0	0	1	0	1
Spey	1	115	6	6	13	11	2	2
Deveron	1	29	28	4	10	6	1	3
Don	0	12	1	96	3	0	1	5
Dee	2	36	6	56	60	14	1	4
North Esk	0	20	3	10	8	24	1	4

Tay	0	6	0	1	0	1	16	0
Tweed	0	11	3	5	5	3	1	79
Assigned %	**94.1**	**79.1**	**77.4**	**85.6**	**68.8**	**76.9**	**85.2**	**88.3**
Accuracy %	**96.4**	**49.8**	**58.3**	**53.9**	**60.6**	**40.0**	**69.6**	**80.6**

At Level 2, both accuracy and percentage of fish assigned for SpeyDev and DonDee had increased, but were still not meeting the conditions required (Table 4). The majority of misassignments were between the regions SpeyDev, DonDee and the River North Esk. As such, these were combined into a larger unit, termed “East” and a Level 3 self-assignment carried out.

Table 4

Level 2 assignment results, with a probability cut-off of 0.8 applied.

Prob cut-off >0.8	**Assigned to**					
Assignment unit	**Oykel/Cassley/Shin**	**SpeyDev**	**DonDee**	**North Esk**	**Tay**	**Tweed**
Oykel/Cassley/Shin	107	3	0	1	0	1
SpeyDev	2	192	39	17	3	5
DonDee	2	59	222	14	2	9
North Esk	0	24	18	24	1	4
Tay	0	6	1	1	16	0
Tweed	0	14	12	3	1	79
Assigned %	**94.1**	**84.7**	**82.5**	**76.9**	**85.2**	**88.3**
Accuracy %	**96.4**	**64.4**	**76.0**	**40.0**	**69.6**	**80.6**

At Level 3, assignment accuracy of fish assigned to the “East” met the accuracy threshold (Table 5). However, the assignment accuracy of the River Tay was still below the set threshold and misassignments were to the “East” group. As such, the Tay site was included within the larger “East” region and a final self-assignment carried out.

Table 5

Level 3 assignment results, with a probability cut-off of 0.8 applied.

Prob cut-off >0.8	**Assigned to**			
Assignment unit	**Oykel/Cassley/Shin**	**East1**	**Tay**	**Tweed**
Oykel/Cassley/Shin	107	6	0	1
East1	4	707	6	18
Tay	0	9	16	0

Tweed	0	32	1	79
Assigned %	**94.1**	**95.7**	**88.5**	**90.7**
Accuracy %	**96.4**	**93.8**	**69.6**	**80.6**

At the final level, all rivers/regions met the assignment accuracy threshold, with 93.8% of fish assigned with 95.9% accuracy to the three defined assignment units. Assignment details are given in Table 6.

Table 6

Final assignment results, with a probability cut-off of 0.80 applied.

Prob cut-off >0.8	**Assigned to**		
Assignment unit	**Oykel/Cassley/Shin**	**East2**	**Tweed**
Oykel/Cassley/Shin	107	6	1
East2	4	749	18
Tweed	0	33	79
Assigned %	**94.1**	**96.8**	**90.7**
Accuracy %	**96.4**	**95.1**	**80.6**

Discussion

Analysis of this dataset has indicated the presence of four genetic clusters, consisting of both individual rivers and larger regions spanning multiple rivers. However, despite significant phylogenetic structuring being present, the levels of differentiation between the units identified were not large enough to translate into robust river-specific assignments using GSI. Indeed, the same three assignment units as had previously been observed were identified, two comprising single river systems (Oykel/Cassley/Shin and Tweed) and the other consisting of all the other sampled North East coast rivers (Gilbey *et al.*, 2016a).

The samples screened were collected as part of NEPS programme, whereby sites were selected using a Generalised Random Tessellation Stratified (GRTS) design rather than targeted sampling based on knowledge of areas of high juvenile density. Such a design allows unbiased sampling to be undertaken across the area of interest. In some cases, however, this resulted in low numbers of fish caught in a number of sites. Notwithstanding this, the aim of the analysis was to characterise populations at the river level, for which the GRTS design delivered an unbiased sampling of sites across the different

rivers. It was important that this was achieved, as it is well known that there is population genetic variance within rivers. The design allowed for this variation to be captured utilising the fixed number of samples to be screened within the project.

Microsatellites can be highly variable, with large numbers or variants (alleles). As a result, sufficient sample numbers are required to ensure that all variation has been captured and to be able to accurately estimate allele frequencies. To address this issue, we combined geographically close sites to represent the wider river area. Combining sites in this way may have resulted in disruption of the unique genetic signature of a site (as reflected in the number of sites out of Hardy-Weinberg equilibrium). This may, in turn, have compromised, to some degree, the power to differentiate between sites. However, this study focused on the assignment potential at the river-level, not at the site-level.

For technical reasons during the panel development, the 101 microsatellites were initially selected to be limited in numbers of alleles in comparison to most microsatellites, with an average of eight alleles per marker as reported by Bradbury *et al.* (2018). Having a lower number of alleles would reduce the impact of small sample sizes, though may not have completely alleviated them. However, considering the close match with previous analysis of the regions with differing numbers of sites/fish and different marker types, it is unlikely that such influences have had a significant impact on the results of the analysis.

Out of the 101 microsatellites, 69 produced reliable and high quality genotypes. Though the omission of nearly a third of the original markers could possibly have influenced the results, we consider that it is unlikely that the addition of the remaining ones would have substantially altered the findings. The markers are randomly distributed throughout the salmon genome (Bradbury *et al.*, 2018) and are generally considered adaptively neutral. A larger number of markers would have resulted in an increased number of alleles, which would, in turn, have increased the power of the assignment analysis. However, power of assignment would not necessarily have related to higher accuracy at the levels of differentiation observed here. That said, the effect of microsatellite number on assignment success could be investigated in the future.

The most distinct sites, with the highest differentiation values from the other sites were Dee_345, Don_1934 and Don_lower, whilst the lowest differentiation was found between and among sites within the rivers Dee, Spey and Deveron. However, despite the low levels of differentiation observed within and between those latter three rivers, many pairwise comparisons were still significantly different, suggesting that these sites harboured genetically distinct populations. Furthermore, based on this panel of neutral markers, an absence of statistical significance in differentiation between sites does not equate to a lack of natural and adaptive differences between these populations. Indeed, previous work has shown both differences between upper and lower parts within each river (Cauwelier *et al.*, 2018a), as well as differences in a set of adaptive markers related to adult return timing (Cauwelier *et al.*, 2018b).

The structuring analysis revealed four genetic groups, two of which included sites from a single river. Previous work has shown the Oykel/Cassley/Shin river system to be genetically distinct from the other East coast rivers further south (Gilbey *et al.*, 2016a). Similar, though less pronounced, differentiation has previously been reported between the River Tweed and the more northern East coast rivers (Gilbey *et al.*, 2016a,b). The third cluster, consisting of sites from both the rivers Don and Dee, including those three most distinct sites, has not been observed before. It is possible that the new panel of microsatellites has resulted in higher resolution amongst this Northeast coast region at these two rivers. However, successful genotyping proved difficult for samples from these Dee and Don sites compared to the remainder of the dataset. All samples were processed using our laboratory standard operating procedures and samples from the various rivers were spread over multiple plates/runs, so as to minimise processing effects related to reagents, staff member etc. Internal QC checks indicated that the quantity of the DNA was insufficient and/or the quality poor from samples originating at these sites, with average scoring percentage for those sites (69.5% ± 6.9%) significantly lower (t-test: $t = 19.9$, $df = 38$, $p < 0.0001$) than for the others (96.1% ± 1.8%). The reasons for the lower DNA quality/quantity from these samples remain unclear. Further analysis would be required to disentangle any issues with sample quality from the potential for the microsatellite panel to have identified a true distinct genetic cluster not seen in previous work.

The fourth cluster encompassed all the other sites in the dataset, from the River Spey down to the River Tay. Despite the absence of clear structuring between these rivers and an associated inability to robustly assign fish to

them, as reported here and in previous work, within-river differences have been observed (Cauwelier *et al.*, 2018a, b). Though this picture is not obvious here, there is an indication that the proportion of “Cluster 2” increases downstream, whilst the proportion of “Cluster 4” tends to decrease (Figure 2B). Again, further, more detailed work would be required to assess this in more detail.

Despite the clustering analysis showing a number of distinct groups, assignment success at the river level did not meet the accuracy threshold, with the exception of the Oykel/Cassley/Shin river system and the River Tweed. Rivers, therefore, had to be combined to represent larger geographic areas and the final set of assignment units was similar to those found previously (Gilbey *et al.*, 2016b). This would suggest that, despite there being genetic differences among sites and rivers, the differences were not sufficient to translate into high assignment success.

The salmon populations from the river Tweed form an interesting complex. Indeed, the Tweed sites have previously both been combined with either the River Tay populations (Gilbey *et al.*, 2016b) or within the larger NE assignment unit (Gilbey *et al.* 2016a), which also included the River Tay. Assignment success for fish originating from the River Tweed has, in all these cases, been around 80%, this also being our threshold value for combining (or not) rivers into larger units. This would suggest that the genetic differences between salmon populations from the River Tweed and the other NE coast rivers are borderline to being large or not large enough to result in robust assignment.

A number of genetic marker panels have now been employed in rivers covering the East coast of Scotland and all have failed to provide high resolution genetic stock identification, despite there being indications of population structuring within and between these rivers. These panels have consisted of neutral markers that are not directly associated with specific traits and their variants are thus not considered to confer any fitness differences. An alternative approach to the use of such markers is to examine variation in DNA sequences that are part of genome regions coding for proteins and are linked to trait variations and, as such, are subject to natural selection (i.e. adaptive markers). These adaptive markers could be identified using a whole-genome sequencing approach. In this method, an organism’s whole genome is sequenced and, when applied to multiple individuals and populations, this approach could result in both the detection of neutral and adaptive population

divergence (Larson *et al.*, 2014; Narum *et al.*, 2018). This technique has been shown to have the potential to enhance GSI resolution (Fuentes-Pardo *et al.*, 2020).

Recently, other non-genetic based methods have been used for stock identification. These methods examine environmental differences between locations that get incorporated into a number of structures within growing fish, such as their scales, otoliths and eye lenses in the form of trace elements (Veinott & Porter, 2005; Marklevitz *et al.*, 2011; Perrier *et al.*, 2011). Similar principles are applied here as in genetic-based approaches, whereby a trace elements baseline is created to which unknown individuals are assigned. This approach has been employed to successfully distinguish and assign Chinook (*Oncorhynchus tshawytscha*) (Marklevitz *et al.*, 2011), brown/sea trout (*Salmo trutta*) (Ramsay *et al.*, 2011; Veinott *et al.*, 2012) and Atlantic salmon (Veinott & Porter, 2005; Perrier *et al.*, 2011).

It may be that enhanced stock identification could be achieved by the combination of both genetic and trace element analysis. Perrier *et al.* (2011) combined a panel of six microsatellites markers with trace element analysis gathered from otoliths and found high levels of accurate assignment (83% - 100%) to river of origin using this combined approach, even though assignment success based solely on the genetic markers varied between 25% and 47%. This approach of combining both genetic markers and environmental differences reflected and traced in scales/otoliths could be explored in future to assess if enhanced stock identification can be achieved within this large and biologically and commercially important region of Scotland.

Acknowledgements

This work was partly funded by Crown Estate Scotland as part of joint Crown Estate Scotland and Marine Scotland Priority Projects 2020/2021 Programme. Our thanks extend to the members of the steering group for this project: Ross Gardiner (MSS), Dr. Janelle Braithwaite (MSS), Fiona Simpson (Crown Estate Scotland) and Professor Colin Bean (NatureScot) for their valued contribution throughout the project.

References

Anderson, E. C. 2017. *Rubias-genetic stock identification (GSI) in the tidyverse*. https://github.com/eriqande/rubias

Armstrong, J. D., Gauld, N. R., Gilbey, J., & Morris, D. J. 2018. Application of acoustic tagging, satellite tracking and genetics to assess the mixed stock nature of coastal net fisheries. Scottish Marine and Freshwater Science, 9,27pp.

Beacham, T.D., Wallace, C.G., Jonsen, K., McIntosh, B., Candy, J.R., Horst, K., Lynch, C., Willis, D., Luedke, W., Kearey, L. & Rondeau, E.B. 2021. Parentage-based tagging combined with genetic stock identification is a cost-effective and viable replacement for coded-wire tagging in large-scale assessments of marine Chinook salmon fisheries in British Columbia, Canada. Evolutionary Applications*, 14*, 1365-1389.

Begg, G.A., Friedland, K.D. & Pearce, J.B. 1999. Stock identification and its role in stock assessment and fisheries management: an overview*. Fisheries Research, 43*, 1-8.

Benjamini, Y. & Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57, 289-300.

Bradbury, I.R., Wringe, B.F., Watson, B., Paterson, I., Horne, J., Beiko, R., Lehnert, S.J., Clément, M., Anderson, E.C., Jeffery, N.W., Duffy, S., Sylvester, E., Roberston, M. & Bentzen, P. 2018. Genotyping-by-sequencing of genome-wide microsatellite loci reveals fine-scale harvest composition in a coastal Atlantic salmon fishery. Evolutionary Applications, 11, 918-930.

Cauwelier, E., Gilbey, J. & Middlemas, S. 2016. Genetic assignment of marine-caught adult salmon at Armadale to region of origin. Aberdeen, Scotland, U.K. Scottish Marine and Freshwater Science Report 6/16. 19 pp.

Cauwelier, E., Gilbey, J., Sampayo, J., Stradmeyer, L. & Middlemas, S.J. 2018b. Identification of a single genomic region associated with seasonal river return timing in adult Scottish Atlantic salmon (*Salmo salar*), using a genome-wide association study. Canadian Journal of Fisheries and Aquatic Sciences, 75, 1427-1435.

Cauwelier, E., Stewart, D.C., Millar, C.P., Gilbey, J., & Middlemas, S.J. 2018a. Across rather than between river genetic structure in Atlantic salmon *Salmo salar* in north-east Scotland, UK: Potential causes and management implications. Journal of Fish Biology, 92, 607-620.

Davidsen, J.G., Rikardsen, A. H., Halttunen, E., Thorstad, E.B., Økland, F., Letcher, B.H., Skarðhamar, T. & Næsje, T.F. 2009. Migratory behaviour and survival rates of wild northern Atlantic salmon *Salmo salar* post-smolts: effects of environmental factors. Journal of Fish Biology, 75, 1700-1718.

Earl, D.A. & vonHoldt, B.M. 2011. Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 3, 429-431.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14, 2611-2620.

Falush, D., Stephens, M. & Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164, 1567-1587.

Friedland, K.D. 1998. Ocean climate influences on critical Atlantic salmon (*Salmo salar*) life history events. Canadian Journal of Fisheries and Aquatic Sciences, 55, 119-130.

Fuentes-Pardo, A.P., Pettersson, M., Sprehn, C. G., Andersson, L. & Farrell, E. 2020. Population structure of the Atlantic horse mackerel (*Trachurus trachurus*) revealed by whole-genome sequencing. A report prepared for the members of the Northern Pelagic Working Group and the Pelagic Advisory Council. 38pp.

Gilbey, J., Knox, D., O'Sullivan, M. & Verspoor, E. 2005. Novel DNA markers for rapid, accurate, and cost-effective discrimination of the continental origin of Atlantic salmon (*Salmo salar* L.). ICES Journal of Marine Science, 62, 1606-1616.

Gilbey, J., Cauwelier, E., Coulson, M.W., Stradmeyer, L., Sampayo, J.N., Armstrong, A., Verspoor, E., Corrigan, L., Shelley, J. & Middlemas, S. 2016a. Accuracy of assignment of Atlantic salmon (*Salmo salar* L.) to rivers and

regions in Scotland and northeast England based on single nucleotide polymorphism (SNP) markers. PLoS One, 11, e0164327.

Gilbey, J., Cauwelier, E., Stradmeyer, L., Sampayo, J., Middlemas, S., Corrigan, L., and Shelly, J. 2016b. Assignment of fish from the north east English net fishery to origin using SNP genetic markers. Marine Scotland Science and the Environment Agency Report. Faskally, Pitlochry, Scotland. U.K. 45pp.

Gilbey, J., Coughlan, J., Wennevik, V., Prodöhl, P., Stevens, J. R., Garcia de Leaniz, C., Ensing, D., Cauwelier, E., Cherbonnel, C., Consuegra, S., Coulson, M.W., Cross, T.F., Crozier, W., Dillane, E., Ellis, J.S., García-Vázquez, E., Griffiths, A., Gudjonsson, S., Hindar, K., Karlsson, S., Knox, D., Machado-Schiaffino, G., Meldrup, D., Nielsen, E.G., Ólafsson, K., Primmer, C.R., Prusov, S., Stradmeyer, L., Vähä, J.-P., Veselov, A.J., Webster, L,M.I., McGinnity, P. & Verspoor, E. 2018. A microsatellite baseline for genetic stock identification of European Atlantic salmon (*Salmo salar* L.). ICES Journal of Marine Science, 75, 662-674.

Gilbey, J., Utne, K. R., Wennevik, V., Beck, A. C., Kausrud, K., Hindar, K., Garcia de Leaniz, C., Cherbonnel, C., Coughlan, J., Cross, T.F., Dillane, E., Ensing, D., García-Vázquez, E., Hole, L.R., Holm, M., Holst, J.C., Jacobsen, J.A., Jensen, A.J., Karlsson, S., Ó Maoiléidigh, N., Mork, K.A., Nielsen, E.E., Nøttestad, L., Primmer, C.R., Prodöhl, P., Prusov, S., Stevens, J.R., Thomas, K., Whelan, K., McGinnity, P. & Verspoor, E. 2021. The early marine distribution of Atlantic salmon in the North-east Atlantic: A genetically informed stock-specific synthesis. Fish and Fisheries, 22, 1274-1306.

Griffiths, A.M., Machado-Schiaffino, G., Dillane, E., Coughlan, J., Horreo, J.L., Bowkett, A. E., Minting, P., Toms, S., Roche, W., Gargan, P., McGinnity, P., Cross, T., Bright, D., García-Vázquez, E. & Stevens, J.R. 2010. Genetic stock identification of Atlantic salmon (*Salmo salar*) populations in the southern part of the European range. BMC Genetics, 11, 1-27.

Griffiths, A.M., Ellis, J.S., Clifton-Dey, D., Machado-Schiaffino, G., Bright, D., Garcia-Vazquez, E. & Stevens, J.R. 2011. Restoration versus recolonisation: The origin of Atlantic salmon (*Salmo salar* L.) currently in the River Thames. Biological Conservation, 144, 2733-2738.

Hansen, M.M., Nielsen, E.E. & Mensberg, K.L.(1997. The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. Molecular Ecology, 6, 469-474.

Jakobsson, M. & Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, 23, 1801-1806.

Jones, O.R. & Wang, J. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources, 10, 551-555.

Kalinowski, S.T. 2005. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes, 5, 187-189.

Koljonen, M.L. & McKinnell, S. 1996. Assessing seasonal changes in stock composition of Atlantic salmon catches in the Baltic Sea with genetic stock identification. Journal of Fish Biology, 49, 998-1018.

Larson, W.A., Seeb, L.W., Everett, M.V., Waples, R.K., Templin, W.D. & Seeb, J. E. 2014. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). Evolutionary Applications, 7, 355-369.

Malcolm, I. A., Millidine, K. J., Jackson, F.L., Glover, R.S. & Fryer, R.J. (2020). The National Electrofishing Programme for Scotland (NEPS) 2019. Scottish Marine and Freshwater Science, 11/9. 57pp.

Marine Scotland Science 2020 rod catch returns data: https://data.marine.gov.scot/dataset/salmon-and-sea-trout-fishery-statistics-2020-season-reported-catch

Marklevitz, S.A., Fryer, B.J., Gonder, D., Yang, Z., Johnson, J., Moerke, A. & Morbey, Y.E., 2011. Use of otolith chemistry to discriminate juvenile Chinook salmon (*Oncorhynchus tshawytscha*) from different wild populations and hatcheries in Lake Huron. Journal of Great Lakes Research, 37, pp.698-706.

Moriya, S., Sato, S., Azumaya, T., Suzuki, O., Urawa, S., Urano, A. & Abe, S. 2007. Genetic stock identification of chum salmon in the Bering Sea and North

Pacific Ocean using mitochondrial DNA microarray. Marine Biotechnology, 9, 179-191.

Narum, S.R., Di Genova, A., Micheletti, S.J. & Maass, A. 2018. Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon. Proceedings of the Royal Society B: Biological Sciences, 285, 20180935.

Nei, M., Tajima, F. & Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. Journal of Molecular Evolution, 19, 153-170.

Peakall, R.O.D. & Smouse, P.E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288-295.

Perrier, C., Daverat, F., Evanno, G., Pécheyran, C., Bagliniere, J.L. & Roussel, J.M. 2011. Coupling genetic and otolith trace element analyses to identify river-born fish with hatchery pedigrees in stocked Atlantic salmon (*Salmo salar*) populations. Canadian Journal of Fisheries and Aquatic Sciences, 68, pp.977-987.

Potter E.C.E., & Crozier, W.W. 2000. A perspective on the marine survival of Atlantic salmon. In: Mills, D., ed. The Ocean Life of Atlantic salmon: Environmental and Biological Factors Influencing Survival. Oxford: Fishing News Books, Blackwell Science, 19–36.

Pritchard, J.K., Wen, W. & Falush, D. 2004. Documentation for STRUCTURE software: ver. 2.3.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Ramsay, A.L., Milner, N. J., Hughes, R.N. & McCarthy, I.D. 2011. Comparison of the performance of scale and otolith microchemistry as fisheries research tools in a small upland catchment. Canadian Journal of Fisheries and Aquatic Sciences, 68, 823-833.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes, 4, 137-138.

Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8 103-106.

Shaklee, J.B., Beacham, T.D., Seeb, L. & White, B.A. 1999. Managing fisheries using genetic data: case studies from four species of Pacific salmon. Fisheries Research, 43, 45-78.

Sylvester, E.V.A., Beiko, R.G., Bentzen, P., Paterson, I., Horne, J.B., Watson, B., Lehnert, S., Duffy, S., Clément, M., Robertson, M.J. & Bradbury, I.R. 2018. Environmental extremes drive population structure at the northern range limit of Atlantic salmon in North America. Molecular Ecology, 27, 4026-4040.

Thorstad, E.B., Whoriskey, F., Rikardsen, A.H. & Aarestrup, K. 2011. Aquatic Nomads: The Life and Migrations of the Atlantic Salmon. In: Aas, O., Einum, S., Klemetsen, A. & Skurdal, J. eds. Atlantic Salmon Ecology. Wiley-Blackwell. 1–32.

Vähä, J. P., Erkinaro, J., Niemelä, E., Primmer, C. R., Saloniemi, I., Johansen, M., Svenning, M. & Brørs, S. 2011. Temporally stable population-specific differences in run timing of one-sea-winter Atlantic salmon returning to a large river system. Evolutionary Applications, 4, 39-53.

Veinott, G. & Porter, R. 2005. Using otolith microchemistry to distinguish Atlantic salmon (*Salmo salar*) parr from different natal streams. Fisheries Research, 71, pp.349-355.

Veinott, G., Westley, P.A., Warner, L. & Purchase, C.F. 2012. Assigning origins in a potentially mixed-stock recreational sea trout (*Salmo trutta*) fishery. Ecology of Freshwater Fish, 21, 541-551.

Walsh, P.S., Metzger, D.A. & Higuchi, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques, 10, 506-513.

Waples, R.S., Winans, G.A., Utter, F.M. & Mahnken, C. 1990. Genetic approaches to the management of Pacific salmon. Fisheries, 15, 19-25.

Waples, R.S., Dickhoff, W.W., Hauser, L. & Ryman, N. 2008. Six decades of fishery genetics: taking stock. Fisheries, 33, 76-79.

Zhan, L., Paterson, I.G., Fraser, B.A., Watson, B., Bradbury, I.R., Nadukkalam Ravindran, P., Reznick, D., Beiko, R.G. & Bentzen, P. 2017. MEGASAT: automated inference of microsatellite genotypes from sequence data. Molecular Ecology Resources, 17, 247-256.

Marine Directorate of the Scottish Government
Freshwater Laboratory
Faskally
Pitlochry
PH16 5LB

Copies of this report are available from the Scottish Government Publications website at https://www.gov.scot/publications/

This publication is available at **www.gov.scot**

Any enquiries regarding this publication should be sent to us at

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-83601-329-7 (web only)

Published by The Scottish Government, October 2024

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS1457078 (10/24)

w w w . g o v . s c o t