Three Decades of Farmed Escapees in the Wild: A Spatio-Temporal Analysis of Atlantic Salmon Population Genetic Structure throughout Norway

Kevin A. Glover1, María Quintela2, Vidar Wennevik1, François Besnier1, Anne G. E. Sørvik1, Øystein Skaala1

1 Section of Population Genetics and Ecology, Institute of Marine Research, Bergen, Norway, 2 Dept of Animal Biology, Plant Biology and Ecology, University of A Coruña, Spain

Abstract

Each year, hundreds of thousands of domesticated farmed Atlantic salmon escape into the wild. In Norway, which is the world’s largest commercial producer, many native Atlantic salmon populations have experienced large numbers of escapees on the spawning grounds for the past 15–30 years. In order to study the potential genetic impact, we conducted a spatio-temporal analysis of 3049 fish from 21 populations throughout Norway, sampled in the period 1970–2010. Based upon the analysis of 22 microsatellites, individual admixture, FST and increased allelic richness revealed temporal genetic changes in six of the populations. These changes were highly significant in four of them. For example, 76% and 100% of the fish comprising the contemporary samples for the rivers Vosso and Opo were excluded from their respective historical samples at P = 0.001. Based upon several genetic parameters, including simulations, genetic drift was excluded as the primary cause of the observed genetic changes. In the remaining 15 populations, some of which had also been exposed to high numbers of escapees, clear genetic changes were not detected. Significant population genetic structuring was observed among the 21 populations in the historical (global FST = 0.038) and contemporary data sets (global FST = 0.030), although significantly reduced with time (P = 0.008). This reduction was especially distinct when looking at the six populations displaying temporal changes (global FST dropped from 0.058 to 0.039, P = 0.006). We draw two main conclusions: 1. The majority of the historical population genetic structure throughout Norway still appears to be retained, suggesting a low to modest overall success of farmed escapees in the wild; 2. Genetic introgression of farmed escapees in native salmon populations has been strongly population-dependent, and it appears to be linked with the density of the native population.

Citation: Glover KA, Quintela M, Wennevik V, Besnier F, Sørvik AGE, et al. (2012) Three Decades of Farmed Escapees in the Wild: A Spatio-Temporal Analysis of Atlantic Salmon Population Genetic Structure throughout Norway. PLoS ONE 7(8): e43129. doi:10.1371/journal.pone.0043129

Editor: Martin Krkosek, University of Otago, New Zealand

Received February 22, 2012; Accepted July 17, 2012; Published August 15, 2012

Copyright: © 2012 Glover et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study is financed by the Norwegian Ministry of Fisheries. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Kevin.glover@imr.no

Introduction

Delineation of historical genetic structure can provide an insight into how contemporary evolutionary relationships among populations have been shaped by demographic, environmental and anthropogenic factors. Understanding these processes and their potential interactions will assist in predicting how natural populations are likely to evolve in relation to present and future challenges.

Salmonid fishes provide excellent opportunities to study evolutionary relationships among populations in both time and space. They inhabit a variety of habitats and display phenotypic and life-history variation among populations [1], some of which reflect local adaptations [2–3]. Furthermore, salmonids tend to exhibit highly distinct population genetic structuring, also in anadromous forms where high fidelity to natal stream (homing) serves to limit gene flow [4]. The Atlantic salmon (Salmo salar) is no exception to these characteristics, and the analysis of molecular genetic markers has revealed highly significant population genetic structuring throughout its entire range [5–8].

The contemporary population genetic structure of Atlantic salmon can be ascribed to a hierarchical system, whereby the largest genetic differences are observed among fish from different continents and regions [9–14]. These differences are to a large degree thought to reflect the patterns of post-glacial colonization. Within regions, highly significant genetic differentiation has been observed among salmon originating from different rivers [11,15,16], and in some cases, also between tributaries within the same river system [15,17–19]. These differences, as revealed by molecular genetic markers, primarily reflect a combination of reproductive isolation and genetic drift, whereby demographics and landscape features play a modifying role [16,17,19]. Generally, where wild populations experience low human impacts, temporal genetic stability has been reported [20,21].

Atlantic salmon populations have been heavily exploited and influenced by a wide-range of anthropogenic factors over a long period of time [22]. Adding to the list of challenges since the 1970’s, is the hundreds of thousands of domesticated salmon that escape from farms on a yearly basis, which display a wide range of interactions with wild conspecifics [23]. Although escapees display
high mortality post-escapement [24,25], they have been recorded in rivers throughout the species’ native range, such as England [26], Scotland [27,28], North America [29], and Norway [30]. Escapes have also been observed in rivers located in countries where salmon farming is not practiced [29].

Genetic changes in native Atlantic salmon populations as a result of introgression from farmed escapes have been observed in Ireland [31–34] and North America [35]. Looking beyond these studies that have been conducted in single rivers, an analysis of seven Norwegian Atlantic salmon populations revealed significant changes in several rivers that had displayed large numbers of farmed escapes on the spawning grounds [36]. However, although farmed escapes have been observed in natural populations for over three decades, and in many regions these numbers exceed wild spawner abundance, the impact this has had on population genetic structure remains elusive. It is therefore not surprising that there are global concerns regarding the genetic integrity of wild populations [23,37–41].

Norway is the world’s largest commercial producer of Atlantic salmon, and is the country where the highest numbers of farmed escapes have been recorded on the spawning grounds. Therefore, Norway represents an ideal country in which to examine how genetic structure has changed both within and among native Atlantic salmon populations in response to widespread migration of farmed escapes onto the spawning grounds. Here, we have conducted a spatio-temporal genetic analysis in order to investigate the potential genetic impacts of farmed escapes on population structure throughout an entire country.

Materials and Methods

Study Design

Atlantic salmon farming in Norway is currently based upon rearing multiple domesticated strains and sub-strains that were initially founded on fish originating from over 40 Norwegian rivers in the 1970’s [42]. Thus, while the allele frequencies of the farmed strains are generally distinct to each other due to founder effects [43], they overlap with the allele frequencies of Norwegian wild populations [43,44]. Over time, farmed escapes do not originate from a single farmed strain, but from multiple strains. The result of this is that the gene flow signal from escapes represents a dynamic mixture of allele frequencies. Thus, the detection of genetic changes in wild populations when gene flow comes from multiple farmed sources is far more complicated [45] than where a set of populations are supplemented by a single and readily defined hatchery source [32,46]. In the latter case, it is straightforward to demonstrate that the allele frequencies in the recipient wild population converges with the allele frequencies with its donor. However, for the case of multiple farmed strains, the recipient wild population will not converge with any given farmed strain over time, and genetic introgression may be partially concealed [45].

Increasing the complexity of detecting genetic introgression of farmed escapes in wild Atlantic salmon populations is that the farmed strains (and therefore their allele frequencies) have, and continue to change significantly with time, i.e., some of the populations used at an earlier stage have been terminated or combined with other strains, while new sub-strains (e.g., in response to QTL selection) have been established. Consequently, it is not possible to accurately reconstruct the allele frequencies of the farmed escapes in Norway over the 15–30 year period in which this study is conducted. Nevertheless, despite the above challenges, modeling has demonstrated that gene flow from farmed escapes will lead to a reduction in genetic structure among wild populations [45,47]. This is because over time, wild populations will be exposed to the average allele frequency from the major strains, and this will start to erode the existing allele frequency differences among wild populations. Furthermore, modeling has shown that genetic changes in wild populations as a result of farmed escapees spawning may be detected, although its likely to be underestimated [45].

As a consequence of the situation described above, the methodological approach implemented in this study is to look at both within and among-population genetic structure in the time-period where the numbers of escapes reported in Norwegian rivers has been highest (i.e., the last 15–30 years). Have native Norwegian salmon populations displayed temporal genetic changes in this period? And if so, can genetic drift be excluded as the primary driver of these temporal changes? Furthermore, where temporal genetic changes have been observed, have the populations become more similar or more differentiated to each other?

Biological Samples

Historical and contemporary samples of Atlantic salmon populations were collected from 21 rivers spanning the entire Norwegian coastline which extends over 2500 km (Fig. 1; Table 1, 2). Populations were chosen primarily due to the availability of archived scale samples which were essential to re-construct the historical baseline [pre- or early aquaculture industry], and, availability of contemporary samples (year 2000+).

Historical samples were exclusively represented by fish scales taken from adult spawners captured in their specific rivers by rod and line (Table 1). Intermediate (neither the oldest nor newest set of samples from any given river system, and contemporary samples, were mostly represented by scale samples taken from adult fish captured by rod and line fishing or various research projects. Therefore, no specific licenses were applied for nor required to collect these samples for this study. Prior to any genetic analysis, all scale samples were analysed for growth patterns in order to exclude any salmon that had directly escaped from fish farms [48]. For some of the intermediate and contemporary samples, adult spawners were not available (for example due to closure of rod and line fishery). Instead, samples of juvenile fish were included for these populations. The historical samples were not collected from the exact same time period (Table S1), however, this was factored into some of the analyses.

Some of the relevant available information for the populations included in this study, which can be found in Norwegian reports [49–52] have been placed into Table 2. Importantly, this information includes the frequency of farmed salmon that have been observed in these populations in the period 1989–2009. Observations of farmed escaped salmon in Norwegian populations are primarily recorded by two approaches. One of the methods is based upon the percent of farmed fish in the angling catch during the summer sports fishing season, while the other is based upon the percent of farmed fish observed during dedicated autumn (spawning site) surveys. As farmed salmon tend to migrate later than wild salmon into freshwater [30], the autumn surveys tend to show higher percentages of farmed fish. However, the surveys of farmed fish frequency in the autumn usually involve sample sizes smaller than the summer angling catch surveys, are conducted less frequently, and are conducted in fewer rivers [49]. Nevertheless, the potential for genetic interaction is more tightly linked to the frequency of escapes observed on the spawning sites during the autumn than found in the summer angling catches. Therefore, we have chosen to use both estimates in the present study. First we use the un-weighted mean percent of farmed fish observed in the spawning surveys (i.e., averaging the percent farmed fish observed...
for the number of years in which they survey was conducted), in addition to using a weighted average based upon combining both summer sports fishing and autumn survey data that has been recently used to categorise over 100 Norwegian rivers in their degree of potential influence from farmed escaped salmon [52]. These estimates have then been compared with the temporal genetic changes observed for each river by regression analysis.

Samples of farmed salmon have been included for the analysis of admixture. These samples were selected from multiple data sets that have been analysed to identify the farms of origin for escapees as a DNA forensic service for the Norwegian ministry of fisheries in the period 2006 - present [53–57]. A total of nine farm samples, each of approximately 45 fish, were chosen based upon their large genetic differences to each other, and, in order to represent some of the genetic diversity found among salmon farms and farmed strains in Norway.

Figure 1. Norwegian rivers where historical and contemporary samples of Atlantic salmon populations were collected for the present study.

doi:10.1371/journal.pone.0043129.g001

Genotyping

DNA extraction was performed in 96-well plates using the Qiagen DNeasy®96 Blood & Tissue Kit. Each DNA plate contained two or more negative controls.

The following twenty two microsatellite loci were used; SSsp3016 (Genbank no. AY372820), SSsp2216, SSspG7, SSsp2201, SSsp1605, SSsp2216 [58], Ssa197, Ssa171, Ssa202 [59], SsaD157, SsaD486, SsaD144 [60], Ssa289, Ssa14 [61], SsaF43 [62], SsaOsl85 [63], MHC I [64] MHC II [65], Ssa19NVH (Genbank no. AF256670), CA060208 [66], SsalR002TKU and SsalR010TKU [67]. Amplifications were conducted in four multiplex reactions (conditions available from the authors). PCR products were analysed on an ABI 3730 Genetic Analyser and sized by a 500LIZ™ size-standard. Automatically binned alleles were manually checked by two researchers prior to exporting data for statistical analyses.
Microsatellites are known to be prone to genotyping errors [68,69], even under strict protocols [70]. Eighteen of the microsatellite markers implemented here are routinely genotyped at IMR, and have revealed low error rates [55]. Within the present data set, some samples were re-analysed in order to increase the genotyping coverage and provide an ad-hoc quantification of genotyping quality.

Statistical Analyses

For most of the statistical analyses conducted, samples were grouped into historical, intermediate and contemporary data sets. Other sub-sets of the data set were analysed for specific tests (i.e., including reduced sets of populations and markers). These variations are identified in the results. Bonferroni adjustment of the significance level for multiple testing was not presented. Instead, statistical significance was tested at $\alpha = 0.05$ and a more stringent level of $\alpha = 0.001$.

The genotype distribution of each locus in each population was compared with the expected Hardy-Weinberg distribution using the program GenePop [71] as was the linkage disequilibrium. Both were examined using the following Markov chain parameters: 10000 steps of dememorisation, 1000 batches and 10000 iterations per batch. Relative genetic variation in each population was assessed using allele frequency data from which observed heterozygosity $H_o$, expected heterozygosity $H_e$, allelic richness, $F_{IS}$ and pairwise $F_{ST}$ were calculated using MSA 4.05 [72].

In order to test whether the global $F_{ST}$ among historical populations was significantly larger than the global $F_{ST}$ among contemporary populations, a bootstrap test based on 10 000 re-sampled datasets was computed. For each resample, the global $F_{ST}$ in historical and contemporary data was calculated based on a random sample of 30% of the individuals from each population and 30% of the markers (7 out of 22). After re-sampling, the distribution of the 10 000 differences between historical and contemporary $F_{ST}$ was used to test the alternative hypothesis ($H_1$: $F_{ST}$ historical $> F_{ST}$ contemporary) against the null hypothesis ($H_0$: $F_{ST}$ historical $\leq F_{ST}$ contemporary).

The program Geneclass 2.0 [73] was used to perform genetic assignment. First, the program was used to conduct self-assignment among the 21 populations in the historical and contemporary data sets. Thereafter, the historical genetic profile for each population was used as the baseline, while individual fish representing the contemporary sample for each population was assigned to their respective baseline population. Exclusion was assessed at a significance level of $\alpha = 0.001$ using all 22 loci, and the reduced set of 14 loci, with the Rannala & Mountain simulation method [74].

| Population     | Sample size (n) | Sample type (NSR) | Population     | Sample size (n) | Sample type (NSR) |
|----------------|----------------|------------------|----------------|----------------|-------------------|
| Neiden H (1979–82) | 79             | SP (1)           | GaulaST H (1986–94) | 48             | SP (1)           |
| Neiden I (1989–93) | 43             | SP               | GaulaST C (2006–08) | 83             | SP                |
| Neiden C (2009) | 93             | SP               | Lærdalselva C (2005–08) | 53             | SP                |
| V. Jakobselva H (1989–91) | 96             | SP (1)   | V. Jakobselva C (2007–08) | 101            | SP                |
| V. Jakobselva C (2007–08) | 101            | SP          | Alta H (1988–90) | 39             | SP (1)           |
| Alta C (1988–90) | 39             | SP (1)           | Alta C (2005–2007) | 85             | SP                |
| Reisa H (1986–91) | 48             | SP (1)           | Reisa C (2006) | 61             | P                 |
| Reisa C (2008) | 30             | SP               | Målselfa H (1986–88) | 47             | SP (1)           |
| Målselfa C (2009) | 47             | SP               | Målselfa C (2008) | 30             | P                 |
| Roksdalsvassdraget H (1987–93) | 37             | SP (1)   | Roksdalsvassdraget C (2008) | 94             | SP                |
| Roksdalsvassdraget C (2008) | 94             | SP          | Namsen H (1977) | 92             | SP (1)           |
| Namsen C (2000) | 58             | SP               | Namsen C (2008) | 102            | SP                |
| GaulaST H (1986–94) | 48             | SP (1)           | GaulaST C (2006–08) | 106            | SP                |
| GaulaST C (2006–08) | 106            | SP            | Suna H (1986–89) | 30             | SP (1)           |
| Suna C (2005–08) | 52             | SP               | Suna C (2005–08) | 52             | SP                |
| Eira H (1986–94) | 34             | SP (0)           | Eira C (2005–2008) | 50             | SP                |
| Eira C (2005–2008) | 50             | SP            | Bondalselva H (1986–88) | 44             | SP (0)           |
| Bondalselva C (2007) | 16             | P              | Bondalselva C (2007–08) | 94             | SP                |
| Ørsta H (1986–89) | 40             | SP (1)           | Ørsta C (2006–08) | 34             | SP                |

Population = name of river with postscript letter H = historical sample, I = intermediate sample, C = contemporary sample. Life stage sampled = SP = spawners, E = eggs, A = alevins, F = fry, P = parr, SM = smolt, NSR = National Salmon River (river protected by extra legislation from government): 1 = yes, 0 = no.

doi:10.1371/journal.pone.0043129.t001
In order to investigate the potential relationship between geographic and genetic distance (FG) in the historic and contemporary data sets, Mantel tests were conducted with the software PASSaGE [73] and significance was found after 10,000 permutations. Genetic differentiation among populations was estimated by the Analysis of Molecular Variance, AMOVA [76] implemented in the program Arlequin [77], and significance was tested after 10,000 permutations.

A growing number of statistical approaches are available to identify putative non-neutral loci [78]. First, we used a hierarchical Bayesian method [79] as implemented in BayeScan software [80]. Secondly, we used the Fdist approach [81], implemented in the program Arlequin [77], and significance was tested after 10,000 permutations.

In order to investigate population structure we identified genetic clusters in the total and neutral dataset with the Bayesian model-based clustering algorithms implemented in STRUCTURE v. 2.3.3 [83–85] under a model assuming admixture and correlated allele frequencies without using population information. Five to ten runs with a burn-in period of 50000–100000 replications and a run length of 500000–1000000 Markov chain Monte Carlo (MCMC) iterations were performed for a variable number of clusters (see footnotes of corresponding barplots for more detailed information). We then applied an ad hoc summary statistic ΔK which is based on the rate of change of the ‘estimated likelihood’ between successive K values [86]. When needed, runs of the selected K were averaged with CLUMPP version 1.1.1 [87] using the LargeKGreedy algorithm and the G’ pairwise matrix similarity statistics and results were visualized as a barplot. Admixture analyses were conducted both with wild salmon and with a combination of wild and farmed salmon (see results).

Genetic drift may be considered as a random evolutionary process whereby a population’s allele frequency at one or more loci can change through time. This process is especially influential in small populations [88,89]. Thus, in order to evaluate whether any of the populations included in the present study were very...
small and likely to be strongly influenced by genetic drift, the effective population size ($N_e$) was computed in each river. This was conducted separately for both the historical and contemporary samples, using the one sample linkage disequilibrium method implemented in the program LDNE [90]. Furthermore, in order to investigate the plausibility that genetic drift could have been the primary driver of the temporal genetic changes observed in some of the populations studied (see results), we simulated genetic drift on these historical populations. For these computations, a methodological approach inspired by an available software for simulating genetic drift [91] was implemented in R (R Development Core Team). Starting from the observed historical sample, additional generations were simulated by gene dropping, so that every additional generation were obtained from the previous one assuming random mating, equal sex proportions, no migration, selection nor migration. Drift was thus assumed to be the only evolutionary force acting upon the populations and markers were unlinked. In order to investigate how $N_e$ influences genetic drift over multiple loci simultaneously, these simulations were conducted 1000 times for each population assuming $N_e$ of 25, 50, 75, 100, 200, 300, 400 and 500, and setting a non-overlapping generation interval to 5 years. The number of generations in which drift was simulated was thereafter a function of the number of years between the historical sample and the corresponding contemporary one, divided by 5, and then rounded up to the nearest whole number. The genetic distance ($F_{ST}$) between the observed historical genetic profile for that population, and the 1000 simulated contemporary populations at each level of $N_e$ were then compared to the genetic distance that was actually observed between the historical and contemporary sample. The probability that the observed pair-wise $F_{ST}$ was greater than the genetic drift simulated $F_{ST}$ was thereafter computed. As in [91], this was achieved by comparing the proportion of the observed $F_{ST}$ values exceeding the genetic-drift simulated $F_{ST}$ values for that population. These simulations were also used to look at global $F_{ST}$ values, and evaluate allelic richness in the presence of genetic drift.

**Results**

**Genotyping Quality**

The final data set consisted of 3049 salmon displaying a mean genotyping coverage of 96.1%. Coverage ranged from 87.1% for the marker SsaD157, to 99.4% for the marker SsaF43. When genotyping success was broken down into the historical and contemporary data sets, coverage was 94.8% and 97.9% respectively.

From 9314 alleles scored independently on two occasions, a mean genotyping error rate (defined here as inconsistent scoring between two independent runs of the same sample) of 0.1% was computed. The absolute number of alleles scored twice/errors observed $= 7506/7$, 806/1, and 1002/2 for the historical, intermediate and contemporary samples respectively. This is consistent with previous estimates for these [55] and other genetic markers [70,92] in this laboratory. Allelic distribution in the historical and contemporary data sets (pooled populations) did not reveal a disproportionate loss of the large alleles in the historical samples (Table S2).

**HWE, LD and Potential Neutrality of Markers**

Analysis of HWE and LD can identify technical issues (marker robustness and genetic linkage between loci) and biological processes (mixing of populations and population disturbance through introgression). At the significance level of $\alpha = 0.05$, a total of 32 (7.1%), 5 (2.9%) and 32 (7.2%) loci by sample combinations displayed significant deviations from HWE in the historical, intermediate, and contemporary samples respectively (Table 3; Table S3—supporting information). At $\alpha = 0.001$, the number of deviations dropped to 2, 1, and 1 in the three data sets respectively. No more than 4 of the 21 populations deviated for any given locus in any of the three data sets demonstrating once again that the markers were of high technical quality. Excluding the historical sample for Vestre Jacobselv, where 9 loci departed from equilibrium at $\alpha = 0.05$ (one of which remained significant at $\alpha = 0.001$), deviations from HWE were distributed among the rivers, with most displaying deviations in 0–3 loci at $\alpha = 0.05$ (Table 3; Table S3).

When computed for all combinations of pairs of loci, within each population separately, LD was detected 309 (6.4%) and 35 (0.7%) times among the historical samples, 122 (6.6%) and 12 (0.6%) times in the intermediate samples, and 422 (8.7%) and 25 (0.5%) times in the contemporary samples at $\alpha = 0.05$, and $\alpha = 0.001$, respectively. Deviations were distributed evenly among the different combinations of pairs of loci, but unevenly distributed among the samples (Table 3). For example, in the historical samples, Vestre Jacobselv displayed 85 pair-wise LD combinations among loci (28% of all LD observed in the historical samples). Together, HWE and LD suggest some form of disturbance in the Vestre Jacobselv in the historical sample. Within the contemporary samples, three populations (Rokdalvassdraget, Reisa and Opo) accounted for 44% of the pair-wise LD combinations observed.

All loci displayed statistically significant global $F_{ST}$ estimates in the historical and contemporary data sets (Table S3). Samples corresponding to the historical data set identified three loci under possible directional selection (MHCC, SsaF43, Ssa289) and five under possible stabilizing selection (SSp2216, Ssa197, Ssd157, Ssd144, Ssd2201), whereas the contemporary data set showed the same loci under possible directional selection but only two of the former ones under possible stabilizing selection (Ssd157, Ssd2201) (Text S1). Subsequently, analyses have been conducted on data sets comprised of the full (all 22 loci) and the neutral (14 loci only) markers.

**Temporal Genetic Variation within Populations**

The number of alleles observed among populations, and between temporal samples within populations varied greatly (Table S3). Differences in sample size were accounted for by computing allelic richness $A_R$. Looking specifically at temporal variation of $A_R$ within populations, most showed a very slight increase with time, however, the populations Vosso, Opo and Loneelva increased by 18–27 (Table 3).

When considering data from the set of 22 loci, and the 14 neutral ones separately, statistically significant temporal genetic change, as measured by $F_{ST}$, was detected in 6 of the 21 populations (Table 3). Populations displaying LD, or distinctly increased $A_R$ in the contemporary samples, were all among those displaying temporal genetic changes. In three of the populations the $F_{ST}$ estimates between historical and contemporary samples exceed 0.01 (i.e., Opo, Vosso and Loneelva). The change in $A_R$ from the historical to the contemporary samples was significantly higher ($P = 0.003$; non-parametric Mann-Whitney test) in the six populations showing temporal genetic changes (mean increase per population $= 13.8$), than in the six ones displaying the strongest temporal stability (mean increase per population $= 2.6$).

No statistically significant correlation was observed between the frequency of farmed escapes observed in a given population in the period 1989–2009 based upon the un-weighted mean from the autumn spawning surveys (see Table 2), and pair-wise $F_{ST}$
Table 3. Effective population size, within-sample genetic diversity estimates, and temporal genetic stability between historical and contemporary samples within 21 Atlantic salmon rivers located throughout Norway. For full data, including locus specific statistics see Table S2.

| Rivers              | Within-sample diversity | Temporal stability | FST, historical vs. contemporary | Exclusion from hist. <0.001 | Temporal change? |
|---------------------|-------------------------|--------------------|---------------------------------|----------------------------|-----------------|
|                     | Historical              | Contemporary       | 22 loci | 14 loci | 22 loci | 14 loci |
| LD                  | HW A_r Ne (95% CI)      | LD                 | HW A_r Ne (95% CI)              |                           |                 |
| Neiden              | 22 0 201 430 (296–760) | 7 1 203 Inf (3179-inf) | 0.0009 | 0.0011 | 6% 3% | No |               |
| V. Jakobselv        | 85 9 190 79 (71–91)     | 32 0 200 169 (148–196) | 0.0064** | 0.0076** | 16% 7% | Yes |               |
| Alta                | 5 2 187 Inf (990-inf)   | 13 1 190 4860 (856-inf) | −0.0002 | 0.0010 | 2% 1% | No |               |
| Reisa               | 11 2 185 272 (180–533) | 61 1 179 80 (69–94)   | 0.0041** | 0.0020 | 15% 10% | No |               |
| Målselv             | 10 2 199 Inf (–1361-inf) | 3 0 207 411332 # (322-inf) | −0.0026 | −0.0011 | 13% 7% | No |               |
| Roksdalsvass.       | 9 0 205 516 (241-inf)   | 66 2 206 384 (291–554) | 0.0014 | 0.0023 | 20% 12% | No |               |
| Namsen              | 10 0 208 3526 (835-inf) | 14 1 209 914 (549–2550) | 0.0013* | −0.0012 | 9% 3% | No |               |
| GaulaST             | 4 0 206 Inf (2162-inf)  | 10 1 208 24753 (1358-inf) | 0.0012 | 0.0018 | 12% 14% | No |               |
| Surna               | 9 0 203 1530 # (252-inf) | 11 1 216 Inf (965-inf)   | 0.0025 | 0.0035 | 34% 17% | No |               |
| Eira                | 11 2 209 378 (196–3201) | 11 0 211 498 (293–1519) | 0.0005 | 0.0000 | 14% 10% | No |               |
| Bondelselva         | 9 0 209 1288 (418-inf) | 12 3 NC. 34 # (26–47)     | 0.0043 | 0.0017 | 6% 0% | No |               |
| Ørstaella           | 6 1 214 3678 (450-inf)  | 17 0 210 400 (202–6501) | 0.0003 | −0.0013 | 0% 0% | No |               |
| GaulaSF             | 7 3 211 1193 (371-inf)  | 19 2 205 439 (311–727)  | 0.0001 | 0.0008 | 17% 1% | No |               |
| Lærdalselva         | 8 1 193 Inf (–506-inf)  | 13 2 200 333 (216–468)  | 0.0015 | 0.0010 | 15% 6% | No |               |
| Vosso               | 14 1 175 Inf (–304-inf) | 8 4 202 189 (138–294)   | 0.0179** | 0.0213** | 76% 67% | Yes |               |
| Loneelva            | 17 5 176 984 (348-inf)  | 8 2 200 241 (172–390)  | 0.0120** | 0.0116** | 52% 29% | Yes |               |
| Opo                 | 10 1 166 Inf # (–14-inf) | 58 1 184 68 (60–76)     | 0.0258** | 0.0279** | 100% 90% | Yes |               |
| Etne                | 25 1 209 752 (439–2405) | 12 3 209 917 (507–4135) | 0.0006 | 0.0000 | 5% 5% | No |               |
| Figgjo              | 9 1 204 Inf (–1638-inf) | 14 2 210 Inf (1070-inf)  | 0.0048** | 0.0058** | 38% 4% | Yes |               |
| Numedalslagen       | 9 1 194 Inf (1194-inf)  | 14 1 210 653 (383–2050) | 0.0032* | 0.0051* | 29% 18% | No |               |
| Berbeylva           | 19 0 156 81 (67–101)    | 19 4 166 245 (194–327)  | 0.0053** | 0.0071** | 16% 7% | Yes |               |

Within samples: LD = observed number of deviations from linkage disequilibrium (231 pair-wise tests per population, 211 tests for Opo) at α = 0.05, HW = observed deviations from Hardy Weinberg Equilibrium (22 tests per population, 21 tests for Opo) at α = 0.05, A_r = allelic richness computed using re-sample size of 25 (note Opo samples only computed with 21 loci therefore not directly comparable to other populations), N_r = effective population size as computed from LD method in LDNE [90] inf = Infinity suggesting that the population is "relatively large" (i.e., >200) [93], # = harmonic mean sample size less than 30 and therefore estimated N_r not to be trusted. Between temporal samples: *= FST significant at α = 0.05, ** = FST significant at α = 0.001 (and following Bonferroni), NC = not computed, Exclusion from hist. = percentage of fish from the contemporary population that are excluded from the historical population profile in the program Geneclass at a cut off of α = 0.001, temporal change? = whether significant temporal genetic change is reported within rivers at α = 0.001 based upon pair-wise FST for both sets of microsatellites.

doi:10.1371/journal.pone.0043129.t003

between the historical and contemporary samples for the same population (R^2 = 0.16, P = 0.052) (Fig. 2a). When using the weighted mean number of escapes reported in a combination of the summer-sports-fishing catch and the autumn spawning counts for each population [52], the correlation with pair-wise FST was statistically strong (R^2 = 0.56, P < 0.0001) (Fig. 2b). However, when the river Opo was excluded (this river displayed by both the highest percentages of escapes and greatest temporal genetic change) the correlation was not significant (R^2 = 0.09, P = 0.20) (Fig. 2c). The lack of a clear relationship between percentage of farmed fish (by either of the two estimations) and observed genetic changes is readily illustrated by the fact that two of the populations (e.g., Opo and Vosso) displayed high numbers of escapes on the spawning grounds and large temporal genetic changes, while other populations (e.g., Ørsta and Etne) also displayed high numbers of escapes but did not reveal genetic change with time. Furthermore, several other rivers had been exposed to >10% escapes in the period 1909–2009 without displaying statistically significant temporal genetic changes (Table 2, 3, Fig. 2).

Individual admixture analysis was also applied to evaluate within-population temporal stability, using historical, intermediate (when available) and the contemporary samples both for the total and neutral sets of microsatellites. The assessment of ΔK in single-population assignment analyses revealed that the most likely number of clusters ranged between two and three (Fig. 3; Fig. S1), although in one population, Berbeylva, this was ≥4 [86]. Admixture analysis supported the results of temporal change from the F-statistics. Thus, populations such as Opo, Vosso, Loneelva and Vestre Jakobselv, which showed temporal genetic changes in FST, also showed evident signs of admixture (Fig. 3).

The percentage of fish from each contemporary sample that was excluded from its historical population sample when conducting genetic assignment ranged from 0–100% when using all 22 loci, and 0–90% when using the reduced set of neutral loci (Table 3). There was a strong correlation between percentage of fish that were excluded from their respective historical populations, and the pair-wise FST values (R^2 = 0.86 P < 0.0001). For example, the populations Opo, Vosso and Loneelva displayed the highest pair-wise FST values between historical and contemporary samples.
Genetic Impact of Farmed Escapees

![Graphs showing the relationship between Pair-wise $F_{ST}$ and Mean percent farmed escapees (1989-2009)]
(0.028, 0.021, 0.012 respectively) in addition to the highest exclusion rates (100%, 76%, and 52% respectively). While other assignment methods implemented in the program Geneclass gave different absolute exclusion percentages, the above trend remained.

Spatio-temporal Genetic Variation

Global FST among the 21 historical samples was significantly larger than among the 21 contemporary ones (Table 4). Significantly, the reduction in global FST with time was observed in 21 of the 22 loci (Fig. 4, Table S3). This trend was also reflected in the self-assignment analyses conducted in Geneclass which showed a drop from 61.6% of fish correctly assigned to their source populations in the historical data set, to 57.6% in the contemporary. Finally, the AMOVA analysis revealed that the amount of genetic variation observed among populations dropped from 4.1% in the historical data set to 2.9% in the contemporary one.

The historical data set was drawn from a wider time-interval than the contemporary one (Table S1). Therefore, in order to test whether this was spuriously responsible for the drop in global FST between the two data sets, a reduced historical data set was established from 12 populations where samples were available from the interval 1986–1994. Likewise, a temporal reduction in global FST was still observed for the 12 populations (Table 4).

Looking specifically at the six populations displaying temporal genetic changes, global FST decreased from 0.058 among the historical samples, to 0.039 among the contemporary ones. In contrast, global FST estimated among the six populations that showed the highest level of within-river temporal stability did not display any change between the historical (0.026) and contemporary (0.027) data sets. Inspection of the pair-wise FST values among the six populations displaying within-population changes showed that all of them contributed to the distinct temporal decrease in global FST (Table 5, 6).

Using data from all 22 markers, a significant relationship between geographic and genetic distance was observed for the total set of populations both in the historical (R² = 0.365, P < 0.0001) and in the contemporary samples (R² = 0.377, P < 0.0001). When looking specifically at the six populations not displaying temporal genetic change, a strong relationship was found in the historical (R² = 0.758, P = 0.0011), and contemporary data sets (R² = 0.668, P = 0.0013). When examining the six populations displaying temporal genetic change, the relationship between genetic and geographic distance was not statistically significant in either the historical (R² = 0.279, P = 0.1013) nor the contemporary data sets (R² = 0.221, P = 0.1411).

Admixture analyses conducted on the 21 populations provided the strongest support for K = 2, both when considering the probability of the data [P(D)] and the ad hoc statistic ΔK, for historical and contemporary samples when using the 22 loci (Fig. 5) and the 14 neutral loci (Fig. S2). In both cases, the five northernmost populations formed a very distinct separate cluster. Following a hierarchical approach, we split the data set into the corresponding five and sixteen populations respectively and conducted the assignment analyses separately. Looking at the full set of markers, the five northernmost populations yielded K3 in the historic dataset and K4 in the contemporary one. Visual inspection of either K3 or K4 for the northern populations revealed increased admixture in several of the rivers over time. This was most apparent for the rivers Vestre Jakobselv, and interestingly, Måselselva, the latter of which did not display temporal genetic change as computed by FST, nor by single-river admixture analysis (Fig. S1). Turning to the remaining sixteen populations, both the historical and contemporary data sets revealed K = 3 as the most likely number of clusters. The southernmost population, Berbyelva was the most distinct (especially in the contemporary data set), and therefore, admixture analyses were also computed with this population excluded. Changes in genetic structure between the historical and contemporary data sets across these sixteen populations were subtle, and not as distinct as for changes within populations (Fig. 3, Fig. S1).

In order to investigate whether the inclusion of farmed salmon would improve the power to detect temporal genetic changes in population genetic structure (either within or among populations), samples from nine genetically distinct farm sources were included in the admixture analyses. Runs were conducted for K = 12 and K = 13 as the analyses included salmon from 9 distinct farm samples, and, that K for the northern and southern clusters had already been estimated at 3 or 4. Both sets of analyses were conducted with and without a prior for the farm samples (which made no difference to the result). As expected, samples from the farms were confirmed to be highly distinct to each other, whereas wild populations were strongly admixed in both the historical and contemporary samples (Fig. S3). Thus, inclusion of farmed fish did not reveal additional temporal genetic changes not already detected.

Effective Population Size and Simulations of Genetic Drift

In most of the historical and contemporary samples representing each population, the computed effective population size (Ne) was larger than 200 (Table 3). Confidence intervals associated with these estimates were large, often reaching infinity in the upper bound (Table 3, Table S4). Several of the samples also showed negative values, both in the upper and lower bound. Negative values occur when the variance observed can be ascribed entirely to sampling error alone, and suggests that these samples displayed relatively high Ne (i.e., >200) [93].

Simulations of genetic drift were conducted for the six populations identified as displaying statistically significant temporal genetic changes. These simulations were conducted in order to evaluate the possibility that genetic drift could have caused the observed changes given the number of generations that have occurred between the historical and contemporary samples.

Unsurprisingly, the mean pair-wise FST between the historical sample and the simulated contemporary population was strongly influenced by Ne (Fig. 6); small Ne leading to large FST. For five of the six populations, a value of Ne of 100 was sufficient to exclude genetic drift as the primary driver of the observed temporal genetic changes (P < 0.001). In these cases the pair-wise FST that was observed between the historical and contemporary sample was greater than the pair-wise FST between the historical sample and the simulated population in all the replicates (i.e., P < 0.001 for 1000 replicates). In the river Figgjo, a value of Ne of 300 or more would be required to achieve the same level of significance.
Figure 3. Bayesian clustering of historical (H), intermediate (I) and contemporary (C) samples representing the four rivers displaying the largest temporal genetic changes at 22 microsatellite loci. For the river Vosso, a total of four samples were available. Thus, the two intermediate samples for this river include a suffix I1 and I2 (linking to these specific samples to Table 1). These analyses were conducted on each river separately. Inferred ancestry was computed using STRUCTURE v. 2.3.3 [83,84], under a model assuming admixture and correlated allele frequencies without using population information. Ten runs with a burn-in period consisting of 100000 replications and a run length of 1000000 Markov chain Monte Carlo (MCMC) iterations were performed for a number of clusters ranging from K 1 to 5. Then an ad hoc summary statistic ΔK [86] was used to calculate the number of clusters (K) that best fitted the data for each river separately. For full computation details and results for all populations using both 22 and 14 markers see Fig. S1 (supporting information).
doi:10.1371/journal.pone.0043129.g003
Comparing these genetic drift simulations with the computed Ne values (Table 3) revealed that genetic drift can be confidently excluded as the driver of the observed temporal genetic changes in the rivers Vosso, Loneelva and Figgjo. This is due to the fact that their Ne values ranged between several hundred and infinity in both the historical and contemporary samples (Table 3). For the rivers V. Jakobselv, Opo and Berbyelva, either the historical or contemporary sample displayed a Ne lower than 100 (79, 68 and 81 respectively). This is at the level of Ne where the potential for genetic drift to contribute to temporal genetic changes on the time-scale studied can be excluded at modest levels of statistical significance ($P = 0.04$, 0.01, and 0.01 for V. Jacobseivl, Opo and Berbyelva respectively for Ne = 75) (Fig. 6, Table 7). Nevertheless, all of these three populations displayed Ne values >150 in one of the samples.

Strong genetic drift in small populations is not only expected to lead to within-population temporal instability, it is expected to simultaneously lead to increased inter-population differentiation (on average) when it is stronger than the influence of gene-flow [88,89]. The genetic drift based simulations reported above were also used to re-compute the global FST value between the six populations displaying statistically significant temporal genetic changes after having simulated genetic drift independently within each population (Fig. 6). The “global” plot illustrates that as Ne decreases, and genetic drift becomes more pronounced within each population, the level of inter-population genetic differentiation increases rapidly. This is in stark contrast to the large and statistically显著的全球FST估计，以及P值表明全球FST估计在历史和当代样本之间是否显著不同。

Table 4. Summary of global FST estimates, and, P values indicating whether the global FST estimates are significantly different between the historical and contemporary samples.

| COMPARISON FST BETWEEN GROUPS (Historical vs. contemporary) | TOTAL LOCI | Neutral LOCI |
|-------------------------------------------------|------------|--------------|
| FST histor. | FST contemp. | $P$ value | FST histor. | FST contemp. | $P$ value |
| All 21 populations | 0.038 | 0.030 | 0.008 | 0.038 | 0.028 | 0.001 |
| 20 populations (excluding Opo) | 0.038 | 0.030 | 0.010 | 0.034 | 0.026 | 0.006 |
| 12 populations in restricted data set* | 0.039 | 0.032 | 0.078 | 0.032 | 0.025 | 0.042 |
| 6 populations displaying temporal changes | 0.058 | 0.039 | 0.006 | 0.057 | 0.032 | 0.001 |
| 6 populations displaying the strongest temporal stability | 0.027 | 0.028 | 0.550 | 0.027 | 0.026 | 0.470 |

All global FST estimates were significant at $\alpha = 0.001$.

*These 12 populations were selected due to narrow the historical temporal data-set to the period 1986–1994, Opo was excluded due to the fact that it was only genotyped for 21 of the 22 loci.

doi:10.1371/journal.pone.0043129.t004

Figure 4. Ratio between global FST computed among the 21 contemporary samples divided by the global FST computed among the 21 historical samples for 22 microsatellite markers. Locus number is consequent with locus names and other locus-specific details available in Table S3.

doi:10.1371/journal.pone.0043129.g004
significant drop actually observed in the global $F_{ST}$ among these six populations with time (Table 5).

**Discussion**

This study represents one of the largest temporal analyses of population genetic structure conducted thus far. Samples covering an entire country, and spanning up to four decades, have permitted the identification of genetic changes occurring both within and among 21 populations, through time. Two main conclusions can be drawn from these analyses. First, despite the fact that farmed escapes have been recorded on the spawning grounds for all of the populations studied, outnumbering wild conspecifics in some years in some of the populations, only weak to moderate changes in among-population genetic structure have been observed in the time-period studied, and in most rivers, statistically significant temporal genetic changes were not observed. This demonstrates that generally, farmed escaped salmon have had poor to moderate success in the wild. Second, not all populations were equally resilient. Genetic changes were observed in six of the populations (29% of those studied), and in four of them, the changes were highly significant. For example, 100%, 76% and 52% of the fish comprising the contemporary samples for Opø, Vosso and Loneelva were distinct and highly significant. For example, 100%, 75% and 52% of the contemporary samples from Opo, Vosso and Loneelva were excluded from their respective historical profiles. When focusing on the six populations displaying temporal changes, global $F_{ST}$ nearly halved between the historical and contemporary data sets. From population genetics theory [88], classical experimental studies [89], and the simulations conducted within this study, genetic drift is expected to lead to greater differentiation among populations. This has been documented for example in the Spanish imperial eagle (*Aquila adalberti*) [101] and forest jaguars (*Panther onca*) [102] in response to habitat fragmentation, and among Atlantic salmon populations that have experienced significant population declines at the southernmost part of their natural distribution [103]. In addition, none of the six populations displaying temporal genetic changes had very low $N_e$ estimates, and based upon simulations, genetic drift was conclusively excluded as the primary driver of the observed temporal genetic changes within most of these rivers. Furthermore, genetic drift was demonstrated to be incompatible with the observed drop in differentiation among these populations with time, and not least, cannot explain the increase in the number of alleles observed in all of these populations. Therefore, in consideration of the genetic data and simulations presented, characteristics of these populations, the high numbers of escapes observed on the spawning grounds (Table 2), and the fact that successful spawning of farmed escaped salmon has been documented in several Norwegian rivers observed with time, significant population genetic structure was still observed in the contemporary data set. Both the historical and contemporary datasets displayed a clear pattern of isolation by distance which is characteristic for Atlantic salmon [15,16]. In 15 of the 21 populations, temporal genetic changes were not detected despite the fact that all of them had experienced farmed escapes on the spawning grounds, and in some years, escapes had outnumbered wild spawners (Table 2). While it is possible that the set of markers implemented here may have failed to detect low-levels of introgression in some populations (see discussion below), it is concluded that the gene flow from farmed escapes into native populations throughout Norway, has been less than the numbers of escapes observed on the spawning grounds. We suggest that this is primarily due to the fact that farmed escapes display reduced spawning success [96–98], in addition to the fact that their offspring display lower survival in the wild when compared with native conspecifics [96,99,100].

Not all of the populations studied were equally resilient. Statistically significant temporal genetic changes were observed in six populations, and for some of these, the changes were very distinct and highly significant. For example, 100%, 75% and 52% of the contemporary samples from Opo, Vosso and Loneelva were excluded from their respective historical profiles. When focusing on the six populations displaying temporal changes, global $F_{ST}$ nearly halved between the historical and contemporary data sets. From population genetics theory [88], classical experimental studies [89], and the simulations conducted within this study, genetic drift is expected to lead to greater differentiation among populations. This has been documented for example in the Spanish imperial eagle (*Aquila adalberti*) [101] and forest jaguars (*Panther onca*) [102] in response to habitat fragmentation, and among Atlantic salmon populations that have experienced significant population declines at the southernmost part of their natural distribution [103]. In addition, none of the six populations displaying temporal genetic changes had very low $N_e$ estimates, and based upon simulations, genetic drift was conclusively excluded as the primary driver of the observed temporal genetic changes within most of these rivers. Furthermore, genetic drift was demonstrated to be incompatible with the observed drop in differentiation among these populations with time, and not least, cannot explain the increase in the number of alleles observed in all of these populations. Therefore, in consideration of the genetic data and simulations presented, characteristics of these populations, the high numbers of escapes observed on the spawning grounds (Table 2), and the fact that successful spawning of farmed escaped salmon has been documented in several Norwegian rivers

| Table 5. Pair-wise genetic distance as computed by $F_{ST}$ among the 6 populations displaying within-river temporal genetic changes. Computed for historical (bottom left) and contemporary samples (top right), and based upon the analysis of 22 loci. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **V. Jakobselva** | **Loneelva** | **Vosso** | **Opo** | **Figgjo** | **Berbyelva** |
| 0.035 | 0.013 | 0.031 | 0.035 | 0.074 |
| 0.067 | 0.013 | 0.017 | 0.020 | 0.063 |
| 0.061 | 0.038 | 0.033 | 0.015 | 0.051 |
| 0.055 | 0.047 | 0.037 | 0.039 | 0.042 |
| 0.086 | 0.086 | 0.078 | 0.069 | 0.053 |

Computed for historical (bottom left) and contemporary samples (top right), and based upon the analysis of 22 loci. All $F_{ST}$ values significant at $P<0.001$ with the exception of those in bold.

| Table 6. Pair-wise genetic distance as computed by $F_{ST}$ among the 6 populations displaying the greatest within-river temporal stability. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Alta** | **Måselselva** | **Eira** | **Ørstaælva** | **GaulaSF** | **Etna** |
| 0.020 | 0.056 | 0.054 | 0.046 | 0.051 |
| 0.021 | 0.029 | 0.024 | 0.023 | 0.026 |
| 0.053 | 0.031 | 0.012 | 0.015 | 0.012 |
| 0.051 | 0.029 | 0.009 | 0.003 | 0.002 |
| 0.049 | 0.027 | 0.009 | 0.007 | 0.004 |
| 0.053 | 0.035 | 0.012 | 0.006 | 0.005 |

Computed for historical (bottom left) and contemporary samples (top right), and based upon the analysis of 22 loci. All $F_{ST}$ values significant at $P<0.001$ with the exception of those in bold.
doi:10.1371/journal.pone.0043129.t006
in the time period studied [104,105], it is concluded that genetic introgression of farmed escaped salmon represents the primary cause of the observed temporal genetic changes. Specifically in the case of the river Vosso, extensive spawning of farmed females has been documented by size and pigment measurements conducted on eggs deposited in the river, leading to the conclusion that the population in this river had been replaced by farmed escapees in the 1990’s [105]. The results of that field experiment are highly consistent with both the timing and magnitude of genetic changes observed in the river Vosso in the present study. Nevertheless, it is worthy of note that the populations in Berbyelva and Figgjo both displayed relatively small temporal genetic changes. For these two populations, the influence of non-biological factors, for example sampling bias in the historical or contemporary samples, or unidentified natural or anthropogenic disturbances, may have had a proportionately high contribution to the observed changes.

No clear relationship between the reported frequency of farmed fish in each population, and the degree of within river genetic changes were revealed in this study. This was true when using both the unweighted mean percent of farmed fish observed in the autumn survey, and the weighted mean combining data from summer sports-fishing catches and autumn surveys [52] (in combination with removing the single river sample Opo which

Figure 5. Hierarchical Bayesian clustering for the historical and contemporary data sets for 21 populations genotyped at 22 microsatellite loci. Inferred ancestry was computed using STRUCTURE v. 2.3.3 [83,84], under a model assuming admixture and correlated allele frequencies without using population information. Ten runs with a burn-in period consisting of 100000 replications and a run length of 1000000 Markov chain Monte Carlo (MCMC) iterations were performed for a number of clusters ranging from K 1 to 5. Then, the ad hoc summary statistic AK [86] was used to calculate the number of clusters (K) that best fitted the data. Populations are ordered North to South, thus corresponding with Tables 1 and 2. Barplots for K3 and K4 are presented for comparison between the historical and contemporary data sets (see results section). For full computation details and results using both 22 and 14 markers see Fig. S2 (supporting information).

doi:10.1371/journal.pone.0043129.g005

Table 7. P-values testing whether the observed pair-wise F<sub>ST</sub> between each population’s historical and contemporary sample was significantly larger than the F<sub>ST</sub> between each population’s observed historical sample and 1000 computer simulated contemporary samples.

| Ne  | Population      | V. Jakobselv | Vosso | Loneelva | Opo  | Figgjo | Berbyelva |
|-----|-----------------|--------------|-------|----------|------|--------|-----------|
| 25  | 0.99            | 0.99         | 0.85  | 1.0      | 1.0  | 0.97   |
| 50  | 0.3             | 0.4          | 0.03  | 0.57     | 1.0  | 0.2    |
| 75  | 0.04            | 0.01         | 0.02  | 0.01     | 1.0  | 0.01   |
| 100 | <0.001          | <0.001       | <0.001| <0.001   | 0.95 | <0.001 |
| 200 | <0.001          | <0.001       | <0.001| <0.001   | 0.03 | <0.001 |
| 300 | <0.001          | <0.001       | <0.001| <0.001   | <0.001| <0.001 |
| 500 | <0.001          | <0.001       | <0.001| <0.001   | <0.001| <0.001 |

Simulations were based upon genetic drift at different Ne. Plots of observed and simulated F<sub>ST</sub> values are presented in Fig. 6.
doi:10.1371/journal.pone.0043129.t007
was solely responsible for the statistically significant relationship) (Fig. 2a, b, c). There are many potential explanations for this result. Firstly, it is important to consider the fact that the numbers of rivers investigated is only 21, limiting the ability to test for such a relationship in a statistically robust manner. Furthermore, and importantly, the data relating to the frequency of farmed fish in these populations (either the summer sports-fishing data or the autumn surveys) has limitations, such as missing counts in some years (Table 2), and the fact that the maturity status of these escapees is not often recorded. Nevertheless, the question still remains; why did some populations (e.g., Opo and Vosso) experiencing large numbers of domesticated escapees display very large temporal genetic changes, while other populations (e.g., Ørsta and Etne), also displaying high percentages of escapees, not reveal detectable temporal genetic changes? From both ecological and conservation viewpoints, these are vital questions in order to understand the evolutionary processes underlying the potential for natural populations to persist in the face of migration and potential gene flow from non-native sources. We suggest that there are both ecological and technical reasons for this. First we address the ecological reasons.

Farmed salmon are competitively inferior to wild salmon in spawning [96–98], and their relative spawning success is density-dependent [106]. Density-dependant spawning success has also been observed for hatchery reared salmon [107]. Together, these studies suggest that farmed escaped salmon will have a higher probability of introgression in native populations with low adult densities, than in populations with high adult densities. Once introgression has occurred, it is likely that the relative survival of the domesticated offspring and admixed individuals will be higher in rivers displaying low juvenile density and accordingly low intra-specific competition. This is because the offspring of domesticated and non-native conspecifics tend to display lower survival in the wild when compared to native fish [96,99,100]. This is consistent with the fact that successful introgression of hatchery reared brown trout in native Danish populations has been partially explained by low wild fish population density [46], and with a recent study that concluded that wild population density is the most important factor affecting the competitive balance between hatchery-reared and wild fish [108]. Furthermore, the two populations (Opo and Vosso) displaying the greatest genetic changes in the present study, have both experienced low numbers of adult spawners in the period where high numbers of escapees were reported. In contrast, two other populations (e.g., Ørsta and Etne) displaying relatively high numbers of wild adult spawners in the population, did not display temporal genetic changes, despite high numbers of escapees.

For several technical reasons, it is possible that the estimated level of within-population temporal genetic changes, as estimated by the 22 microsatellites implemented here, is lower than the true level of genetic introgression by farmed escapes. As detailed in the Materials and Methods, gene flow from farmed fish into wild populations may be concealed and thus underestimated [45]. Several of the populations studied here displayed close to significant temporal genetic changes in \( F_{ST} \), relatively high exclusion rates from the historical population, and, some evidence of linkage disequilibrium (Table 3). Furthermore, the ability to detect statistically significant temporal genetic changes is influenced by the ratio between sample and effective population size (\( S/\bar{N} \)) [109]. Given that both factors varied among the samples and populations in this study (i.e., the contemporary sample for Bondalselva, which represented the smallest sample, was only \( N = 16 \)), this may have limited the ability to detect temporal changes in some of the populations. It is possible however, that analysis of genetic markers putatively under domestication selection [44] may provide the ability to quantify introgression of escapces in rivers where this has occurred at a low level.

The effective population size (\( \bar{N} \)) represents an important parameter in conservation genetics as it provides information about the potential for genetic drift, inbreeding and natural selection to act upon populations. A range of methods for computing \( \bar{N} \) are available, and may be broadly split into temporal [109–112] and one-sample [90,113–115] based approaches. Here, we applied a one-sample based method [90] that utilizes a bias correction [116]. This provided us with the ability to compute \( \bar{N} \) for both the historical and contemporary samples separately, in order to estimate whether these were small populations likely to be under the influence of genetic drift. All methods of computing \( \bar{N} \) include underlying assumptions that are rarely fulfilled in the populations in which they are implemented. For example, linkage disequilibrium, which is the primary parameter used to estimate \( \bar{N} \) in single-sample methods, can be caused by several factors not related to \( \bar{N} \), such as immigration and overlapping generations. Both of these two underlying assumptions were violated by the populations in the present study, although the LD method implemented by [90] has been demonstrated to be robust to equilibrium migration [117]. Thus, while the \( \bar{N} \) estimations presented here should be treated with some caution, they nevertheless provide indications regarding each population’s effective size, and thus potential for genetic drift.

The genetic changes observed here occurred over a period of 15–30 years, which is equivalent to approximately 3–6 generations for these wild populations. This time-scale is consistent with predictions from models of gene flow based upon experimental data in which it has been suggested that under high intrusion scenarios, it will be difficult to obtain broodstock from the original population after just a few generations [118]. This correlates strongly with the results of our genetic assignment tests, where over half of the contemporary populations for Opo, Vosso and Lonelva could be excluded from their historical population profiles at \( P = 0.001 \) (Table 3). Given that farmed salmon continue to escape into the natural environment, it is likely that the number of populations where introgression is observed, and the magnitude of introgression within each population, will increase with time. Several of the salmonid species in the Pacific are monitored, and in some circumstances, actively managed using genetic based methods [119]. Furthermore, there are a range of advantages in using genetic methods to monitoring populations for conservation and management [120]. Here, it is suggested that if farmed salmon continue to escape into the wild, a monitoring program to assess genetic stability in native salmon populations will be necessary.
order to produce science-based management strategies in the future.

Salmonid fish populations are often regarded as locally adapted to their native environments [1–3], and supplementation with hatchery produced or non-native conspecifics is potentially negative to wild populations [121]. Farmed salmon have been selected for a range of economically important traits for approximately ten generations, and as a result, they display genetic differences to wild salmon. For example, farmed salmon grow significantly faster [122], transcribe genes differently [123], exhibit reduced anti-predator responses [124], and display lower fitness in the natural environment [96,99,100]. Nevertheless, analysis of neutral, or nearly neutral genetic markers as has been conducted here, can only describe changes in population genetic structure due to gene flow. While this represents a necessary step towards understanding the level of genetic-impact that farmed escaped fish may cause in native populations, such data cannot directly infer biological consequences in recipient wild populations. Ultimately, a major question will be how allele frequencies in genes causatively linked to adaptive traits have changed in these populations.

Supporting Information

Figure S1 Bayesian clustering of historical (H), intermediate (I) and contemporary (C) samples for 21 Atlantic salmon rivers separately.

Figure S2 Hierarchical Bayesian clustering of the 21 rivers in the historical and contemporary data sets.

Figure S3 Bayesian clustering of the 21 rivers in the historical and contemporary data sets when combined together with data from 9 distinct farm sources.

Table S1 Years in which samples were taken for the historical, intermediate, and contemporary data sets.

References

1. Garcia de Leaniz C, Fleming IA, Einum S, Verspoor E, Jordan WC, et al. (2007) A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. Biological Reviews 82:173–211.
2. Taylor EB (1991) A review of local adaptation in salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture 98: 183–207.
3. Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity 106:404–420.
4. Stabell OB (1984) Homing and allocthon in salmonids - a critical review with special reference to the Atlantic salmon. Biological Reviews of the Cambridge Philosophical Society 59: 333–388.
5. Stahl G (1987) Genetic population structure of Atlantic salmon. In: Ryman N, Utter F, editors. Population genetics and fishery management: University of Washington press, Seattle. 121–146.
6. Verspoor E, Beardsmore JA, Consuegra S, De Leaniz CG, Hindar K, et al. (2005) Population structure in the Atlantic salmon: insights from 40 years of research into genetic protein variation. Journal of Fish Biology 67: 3–54.
7. Dionne M, Caron F, Dodson JJ, Bernatchez L (2008) Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Molecular Ecology 17: 2382–2396.
8. Griffiths AM, Machado-Schiaffino G, Dillane E, Coughlan J, Horree JL, et al. (2003) Genetic stock identification of Atlantic salmon (Salmo salar) populations in the southern part of the European range. Bmc Genetics 11.
9. Taggart JB, Verspoor E, Galvin PT, Moran P, Ferguson A (1995) A minisatellite DNA marker for discriminating between European and North American Atlantic salmon (Salmo salar). Canadian Journal of Fisheries and Aquatic Sciences 52: 2905–2911.
10. Gilby 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: 1609–1616.
11. Trouten A, Veselov AJ, Zuschensken AV, Lumme J, Primemer CR (2009) Microsatellites reveal clear genetic boundaries among Atlantic salmon (Salmo salar) populations from the Barents and White seas, northwest Russia. Canadian Journal of Fisheries and Aquatic Sciences 66: 717–735.
12. King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski BA (2001) Population structure of Atlantic salmon (Salmo salar L.): a range-wide perspective from microsatellite DNA variation. Molecular Ecology 10: 807–821.
13. Lubienczi K, Jones SL, Davidson EA, Park J, Koope BF, et al. (2010) Comparative genomic analysis of Atlantic salmon, Salmo salar, from Europe and North America. Bmc Genetics 11.
14. Wennenv K, Skaala O, Tito SV, Studnov J, Naeverd G (2004) Microsatellite variation in populations of Atlantic salmon from North Europe. Environmental Biology of Fishes 69: 143–152.
15. Dillane E, Cross MC, McGinnity P, Coughlan JP, Galvin PT, et al. (2007) Spatial and temporal patterns in microsatellite DNA variation of wild Atlantic salmon, Salmo salar, in Irish rivers. Fisheries Management and Ecology 14: 209–219.
16. Perrier C, Guymard R, Bagliniere JL, Evanno G (2011) Determinants of hierarchical genetic structure in Atlantic salmon populations: environmental factors vs. anthropogenic influences. Molecular Ecology 20: 4231–4245.
17. Dillane E, McGinnity P, Coughlan JP, Cross MC, de Eyto E, et al. (2008) Demographics and landscape features determine intrariver population structure in Atlantic salmon (Salmo salar L.): the case of the River Moy in Ireland. Molecular Ecology 17: 4786–4800.
18. Dionne M, Caron F, Dodson JJ, Bernatchez L (2009) Comparative survey of within-river genetic structure in Atlantic salmon; relevance for management and conservation. Conservation Genetics 10: 409–429.
19. Vaha JP, Erkkilä J, Niemela E, Primmer CR (2007) Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Molecular Ecology 16: 2638–2654.
34. Clifford SL, McGinnity P, Ferguson A (1998) Genetic changes in an Atlantic salmon population following introgression by farmed escapees. Aquaculture Environment Interactions 10: 123–146.

35. Hindar K, Ryman N, Utter F (1991) Genetic effects of cultured fish on natural populations. Canadian Journal of Fisheries and Aquatic Sciences 54: 1224–1233.

36. Hansen MM, Fraser DJ, Meier K, Mensberg KLD (2009) Sixty years of salmon farming in Norwegian sea-cages reveals a non-random distribution of genetic material from a breeding programme to commercial production. Aquaculture Research 40: 1182–1190.

37. Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (Salmo salar L.) microsatellites. Molecular Ecology Notes 4: 160–162.

38. O'Reilly PT, Hamilton LC, McConnell SK, Wright JM (1996) Rapid analysis and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (Salmo salar L.) microsatellites. Molecular Ecology Notes 5: 130–132.

39. Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14: 599–612.

40. Ferguson A, Fleming IA, Hindar K, Skaala Ø, McGinnity P, et al. (2007) Farm escapees. Ices Journal of Marine Science 63: 1162–1181.

41. Hansen LP (2006) Migration and survival of farmed Atlantic salmon (Salmo salar L.) released from two Norwegian fish farms. Ices Journal of Marine Science 63: 1211–1217.

42. Gjedrem T, Gjoen HM, Gjerde B (1991) Genetic-Origin of Norwegian Farmed salmon (Salmo salar) L. Genetics, Conservation and Management. Oxford, UK: Blackwell.

43. Hansen MM, Fraser DJ, Meier K, Mensberg KLD (2009) Sixty years of salmon farming in Norwegian sea-cages reveals a non-random distribution of genetic material from a breeding programme to commercial production. Aquaculture Research 40: 1182–1190.

44. O'Reilly PT, Hamilton LC, McConnell SK, Wright JM (1996) Rapid analysis and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (Salmo salar L.) microsatellites. Molecular Ecology Notes 4: 160–162.

45. Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14: 599–612.

46. Hansen MM, Fraser DJ, Meier K, Mensberg KLD (2009) Sixty years of anthropogenic pressure: a spatio-temporal genetic analysis of brown trout populations subject to stock and population declines. Molecular Ecology 18: 2549–2562.

The text continues with more scientific references and methodologies, focusing on genetic interactions and escapees in aquatic systems.
98. Weir LK, Hutchings JA, Fleming IA (2004) Dominance relationships involving maternal allele frequencies: a QTL-mapping approach to identify QTLs underlying visual and mechanical performance traits in wild, farmed and hybrid progeny in a natural river environment. Ices Journal of Marine Science 54: 986–1000.

99. Martinez-Cruz B, Godoy JA, Negro JJ (2007) Population fragmentation leads to spatial and temporal genetic structure in the endangered Spanish imperial eagle. Molecular Ecology 16: 477–486.

100. Haag T, Santos AS, Sana DA, Morato RG, Cullen L, et al. (2010) The effect of habitat fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (Panthera onca). Molecular Ecology 19: 4906–4921.

101. Tatarca CP, Berejikian BA (2012) Mechanisms influencing competition between hatchery and wild juvenile anadromous Pacific salmonids in fresh water and their relative competitive abilities. Environmental Biology of Fishes 94: 7–19.

102. Saegrov H, Hindar K, Kalas S, Laura H (1997) Escaped farmed Atlantic salmon replace the original salmon stock in the River Vosso, western Norway. Ices Journal of Marine Science 54: 1166–1172.

103. Borrell YJ, Bernardo D, Blanco G, Vazquez E, Sanchez JA (2008) Spatial and temporal variation of genetic diversity and estimation of effective population sizes in Atlantic salmon (Salmo salar L.) populations from Asturias (Northern Spain) using microsatellites. Conservation Genetics 9: 807–819.

104. Lura H, Saegrov H (1991) Documentation of successful spawning of escaped farmed female Atlantic salmon, Salmo salar, in Norwegian rivers. Aquaculture 98: 151–159.

105. Saegrov H, Hindar K, Kalas S, Laura H (1997) Escaped farmed Atlantic salmon replace the original salmon stock in the River Vosso, western Norway. Ices Journal of Marine Science 54: 1166–1172.

106. Kara PE (2012) Allele frequency covariance among cohorts and its use in estimating effective size of age-structured populations. Molecular Ecology Resources 12: 476–480.

107. Jorde PE, Ryman N (2007) Unbiased estimator for genetic drift and effective population size. Genetics 172: 927–935.

108. Pudovkin AI, Zaykin DV, Hedgecock D (1996) On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. Genetics 143: 327–336.

109. Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONESAMP: a program to estimate effective population size using approximate Bayesian computation. Molecular Ecology Resources 8: 299–301.

110. Wang JL, Whitlock MC (2003) A Genome-Scan Method to Identify Selected Loci appropriate for Both Dominant and Codominant Markers: A Bayesian Approach. Genetics 168: 977–993.

111. Rosenberg MS, Anderson CD (2011) PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. Methods in Ecology and Evolution 2: 29–32.

112. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from matric distances among DNA haplotypes - Application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

113. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1: 47–50.

114. Sandoval-Castellanos E (2010) Testing temporal changes in allele frequencies: a genome-scan approach. Molecular Ecology 19: 969–980.

115. Folli M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Approach. Genetics 168: 977–993.