

Research article

Open Access

# Genetic characterisation of farmed rainbow trout in Norway: intra- and inter-strain variation reveals potential for identification of escapees

Kevin A Glover

Address: Institute of Marine Research, PO Box 1870, Nordnes, N-5817 Bergen, Norway

Email: Kevin A Glover - kevin.glover@imr.no

Published: 16 December 2008

Received: 12 July 2008

BMC Genetics 2008, 9:87 doi:10.1186/1471-2156-9-87

Accepted: 16 December 2008

This article is available from: <http://www.biomedcentral.com/1471-2156/9/87>

© 2008 Glover; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

**Background:** The rainbow trout (*Oncorhynchus mykiss*) is one of the most important aquaculture species in the world, and Norway is one of the largest producers. The present study was initiated in response to a request from the Norwegian police authority to identify the farm of origin for 35 escaped rainbow trout captured in a fjord. Eleven samples, each consisting of approximately 47 fish, were collected from the three farms operating in the fjord where the escapees were captured. In order to gain a better general understanding of the genetic structure of rainbow trout strains used in Norwegian aquaculture, seven samples (47 fish per sample) were collected from six farms located outside the region where the escapees were captured. All samples, including the escapees, were genotyped with 12 microsatellite loci.

**Results:** All samples displayed considerable genetic variability at all loci (mean number of alleles per locus per sample ranged from 5.4–8.6). Variable degrees of genetic differentiation were observed among the samples, with pair-wise  $F_{ST}$  values ranging from 0–0.127. Self-assignment tests conducted among the samples collected from farms outside the fjord where the escapees were observed gave an overall correct assignment of 82.5%, demonstrating potential for genetic identification of escapees. In the "real life" assignment of the 35 captured escapees, all were excluded from two of the samples included as controls in the analysis, and 26 were excluded from the third control sample. In contrast, only 1 of the escapees was excluded from the 11 pooled samples collected on the 3 farms operating in the fjord.

**Conclusion:** Considerable genetic variation exists within and among rainbow trout strains farmed in Norway. Together with modern statistical methods, this will provide commercial operators with a tool to monitor breeding and fish movements, and management authorities with the ability to identify the source of escapees. The data generated in this study were used by the Norwegian police to initiate an investigation of the company operating the three farms in the fjord where escapees were observed.

## Background

The rainbow trout (*Oncorhynchus mykiss*) is one of the most important cultured fish in the world, and is reared in over 65 countries [1]. In 2006, global production

exceeded 550 000 tonnes, and Norway, which is one of the major producing countries, accounted for 78 000 tonnes in the same year [2].

In the 1990's, only 1–2% of world aquaculture production was based upon genetically improved stocks [3]. The first successful family based genetic breeding programs for fish were initiated for the rainbow trout and Atlantic salmon (*Salmo salar* L.) in Norway in the early 1970's [4]. Since then, mass selection programs for a number of fish species have been initiated throughout the world (e.g., [5–7]), and the number of breeding programs continues to increase. However, for a number of fish species reared in culture, even major strains are often poorly or only sporadically characterised at the molecular genetic level, and, the use of genetic markers for breeding assistance is not widespread. A number of population genetic studies of wild [8–11], and hatchery [12–14] rainbow trout strains have been conducted in the USA. In contrast, only one major study of European rainbow trout strains has been published [15].

In Norway, rainbow trout production is almost entirely based upon grow-out in marine cages. Some escapement of cage reared fish into the wild is inevitable, and in the period 1998–2007, 8 000–253 000 (83 000 average) annual rainbow trout escapes were reported to the Norwegian Directorate of Fisheries (NDF) [2]. However, these numbers are thought to be an underestimation due to underreporting of escapement [16]. The rainbow trout is not native to Europe, and consequently, escapees cannot cause direct genetic effects on wild populations. Nevertheless, indirect genetic effects (e.g., [17],) and ecological effects on native salmonid populations (Atlantic salmon and brown trout *Salmo trutta* L.), may occur. Consequently, better control and monitoring of escapees is required. In a recent study, Glover et al. [18] reported, for the first time, successful identification of farmed escaped Atlantic salmon to both farm and cage of origin. This was achieved by comparing individual genetic profiles of escapees with group genetic profiles for 16 samples of Atlantic salmon taken from seven farms. That study built upon the fact that clear genetic differences have been reported among the major Atlantic salmon strains farmed in Norway [19].

The present study was initiated in response to a request by the NDF and the Norwegian police. These authorities requested genetic analyses of rainbow trout that had been captured in a fjord, in order to assist in the identification of their most likely farm(s) of origin. At the time of investigation, however, the genetic structure within and among rainbow trout strains used by the Norwegian aquaculture industry was not known. Consequently, the present study aimed to 1) investigate the extent of genetic variation within and among rainbow trout strains used in Norwegian aquaculture, 2) study the potential of genetic assignment to identify rainbow trout to strain and farm of origin, and 3) to attempt to identify the farm(s) of origin for a group of escaped rainbow trout in a Norwegian fjord based upon a "real-life" escapement episode.

## Methods

### Samples

For legal reasons, the names and locations of all farms and samples used in this study, including exact dates of collection, are anonymous. The present study was initiated after the NDF received several reports detailing observations of escaped rainbow trout in a Norwegian fjord in 2008. No farm in the area reported any losses of fish. A total of three operational rainbow trout farms, all owned by the same company, were located in the inner part of the fjord. All were sited within a 5 km radius from where the escapees were captured, and the closest alternative source for the rainbow trout escapees was >130 km distant by sea. The escapees were uniform in size (1.5–3.8 kg), and overlapped with the size range of rainbow trout held on the three farms in the vicinity. Losses of cage-reared salmonids are often detected in the local area by a sudden increase in the catch per unit effort [20]. Whilst the circumstances surrounding the escapement indicated that the most likely source of escapement was from one or more of these three farms, the Norwegian police required additional evidence to initiate a formal investigation. A simple question was posed. Does the genetic profile of the escapees fit in with one or more of the suspected farms, or can they be excluded as potential sources of the escapement? In order to address this question, a total of 11 samples, representing all cages on these farms, were taken (Table 1). Multiple cage sampling was conducted as the fish reared on the three farms were declared as a non-random and unidentified mixture of two genetic strains, and it was essential to sample all possible sources of the escapees. These farms (farms 7–9, samples 8–18) are here on referred to as the suspected farms/samples collectively. In addition, the NDF collected 35 escaped rainbow trout by contacting local fishermen who had captured escapees in the area in the period January–March 2008.

To investigate genetic variation observed within and among rainbow trout strains used in Norwegian aquaculture in general, seven samples from other sites were also analysed (Table 1). The locations of these samples were such that they were not regarded as potential sources for the 35 escapees. Approximately 47 fin clips from individual fish were taken for each sample, and stored in ethanol until DNA extraction.

### Microsatellite screening

DNA was extracted from fin clips using a Qiagen DNAeasy 96 tissue kit. Twelve rainbow trout derived microsatellite loci (Table 2) were amplified in two separate multiplexes using a mixture of the reaction and amplification conditions presented in two earlier articles [12,21] as the start point for optimisation. PCR amplification was performed as a multiplex in 96 well plates with a total reaction volume of 12  $\mu$ l. For both multiplexes, each reaction con-

**Table 1: Samples used in the present study.**

Sample	Farm or Location	Sample size (n)	Specification of sample	Genetic strain
1	1	47	Slaughterhouse A. (mixed cages)	Not specified
2	2	47	Slaughterhouse A. (mixed cages)	SSF
3	3	47	Slaughterhouse B. (mixed cages)	ISV
4	4	47	Single production cage sampled (sea)	AG
5	5	47	Single tank sampled (freshwater)	AL
6	5	47	Single tank sampled (freshwater)	SB
7	6	47	Single production cage sampled (sea)	AG
8	7	47	Single production cage sampled (sea)	AG + ISV mixed
9	7	47	Single production cage sampled (sea)	AG + ISV mixed
10	7	47	Single production cage sampled (sea)	AG + ISV mixed
11	7	47	Single production cage sampled (sea)	AG + ISV mixed
12	8	47	Single production cage sampled (sea)	AG + ISV mixed
13	8	47	Single production cage sampled (sea)	AG + ISV mixed
14	8	47	Single production cage sampled (sea)	AG + ISV mixed
15	9	47	Single production cage sampled (sea)	AG + ISV mixed
16	9	47	Single production cage sampled (sea)	AG + ISV mixed
17	9	47	Single production cage sampled (sea)	AG + ISV mixed
18	9	47	Single production cage sampled (sea)	AG + ISV mixed
19	10	35	Escapes captured in sea close to farms 7–9	Unknown

Samples 8–18 were collected on the three rainbow trout farms located in the fjord where the escapees were captured. Samples 1–7 were collected from different locations throughout Norway and were not regarded as potential sources of origin for the escapees (sample 19). SSF = Stolt Sea Farms, ISV = IIsvåg, AG = Aqua Gen, AL = Alf Lone, SB = Salmobreed.

sisted of 2 µl DNA, 2.0 mM MgCl<sub>2</sub>, 0.2 mM dNTP's and 0.5 U Go *Taq* polymerase and the manufacturers' 1× reaction buffer. In all cases, forward primers were labelled with fluorescent dyes for fragment detection. Primer concentrations for both multiplexes are presented (Table 2). Amplification was performed in an Eppendorf Master Cycler using the following programme for multiplex 1: 95°C for 10 minutes, then 2 cycles of 94°C for 1 minute, 62°C for 45 seconds and 72°C for 2 minutes, then 29 cycles of 94°C for 1 minute, 58°C for 45 seconds and 72°C for 2 minutes, and finishing with a 72°C extension for 45 minutes and hold at 4°C, and for multiplex 2: 95°C for 10 minutes, 29 cycles of 94°C for 1 minute, 58°C for 45 seconds and 72°C for 2 minutes, and finishing with a 72°C extension for 45 minutes and hold at

4°C. PCR fragments were separated using an ABI 3730 sequencing machine and sized relative to the Applied Biosystem GeneScan™-500 LIZ™ size standard. Alleles were scored using automatic binning implemented in the GenMapper software (V4.0) with manual checking before data were exported for further analyses.

Following the main screening, DNA from the 35 escapee fish was isolated for a second time. These samples, in addition to 60 DNA samples randomly selected from the first screening were then re-analysed (PCR amplification followed by allele detection in ABI) in order to check for accuracy of genotyping. Routinely carrying out such checks in DNA data sets has recently been recommended as standard practice [22,23].

**Table 2: Summary data for 12 microsatellite markers.**

Hexaplex 1				Hexaplex 2			
Marker	Primer µM	ABI dye	Allele size	Marker	Primer µM	ABI dye	Allele size
OMM5132 <sup>1</sup>	0.25	6-Fam	94–126*	OMM5177 <sup>1</sup>	0.1	6-Fam	115–142
OMM5047 <sup>2</sup>	0.1	6-Fam	259–278*	OMM1051 <sup>4</sup>	0.075	6-Fam	210–290
OMM1303 <sup>3</sup>	0.2	Pet	286–362	OMM5264 <sup>1</sup>	0.12	Pet	104–117
OMM5007 <sup>2</sup>	0.08	Ned	155–199	OMM1097 <sup>5</sup>	0.1	Ned	213–319*
OMM5233 <sup>1</sup>	0.1	Vik	112–138	OMM1088 <sup>5</sup>	0.075	Vik	113–159
OMM1008 <sup>4</sup>	0.1	Vik	258–279	OMM1325 <sup>3</sup>	0.075	Vik	278–296*

Concentration is for each primer (i.e., forward and reverse). <sup>1</sup> = [36], <sup>2</sup> = [37], <sup>3</sup> = [38], <sup>4</sup> = [39], <sup>5</sup> = [40]. \* denotes marker displaying one or more alleles separated by 1 bp.

### Statistical analysis

Genepop V3.4 [24] was used to test for deviation from Hardy Weinberg equilibrium, calculate pair-wise  $F_{ST}$  values among the samples, and test for sample differentiation using the Fishers method (Markov chain parameters: dememorization number 1000, number of batches 100, iterations per batch 1000). FSTAT V2.9.3.2 [25] was used to compute allele frequencies for all loci in all samples, and estimate allele richness. TFGA V1.3 [26] was used to calculate pair-wise genetic distance according to Nei's 1978 genetic distance [27], calculate heterozygosity for all samples across all loci, and produce the UPMGA diagram. Assignment tests were conducted using GeneClass V1.0.02 [28]. Self-assignment was conducted using the direct assignment option, Bayesian method of computation, and the leave one out sub-option. The 35 escapees were assigned to farm samples using a mixture of direct assignment, which places each individual into the sample that it is most similar to irrespective of absolute value of genetic similarity, in addition to probability based exclusion, which rejects individuals from samples at a set probability, both using the Bayesian method of computation.

In order to investigate different escapement and identification scenarios, self-assignment simulations were conducted on several data sub-sets. First, self-assignment was conducted using only the samples collected from outside of the fjord where the escapees were captured (samples 1–7). Second, self-assignment was conducted for samples 1–8, 10, 12 and 15. This sub-data set included all samples collected from outside the fjord where the escapees were captured, in addition to a sample from each of the three farms within the fjord. For farm 7 however, two samples were represented. This is because two genetically distinct groups of fish were observed on this farm (see results). Third, self-assignment was conducted among all farm samples (1–18).

In order to identify the source of origin for the 35 escapees, a combination of direct assignment, in addition to exclusion (at 0.05 probability), were carried out using a genetic baseline consisting of all samples collected from the three farms located in the vicinity of the escapees (samples 8–18), and three samples collected from other locations (samples 5–7). Due to the high degree of miss-assignment observed for the self-assignment simulations within the entire data set (see results), and, that none of the samples collected from outside of the fjord were considered to be real alternative sources of origin, only three of these samples, randomly selected to serve as controls, were included for these computations.

## Results

### Genotyping quality

The control re-screening of 95 DNA samples produced identical genotype scores for ten of the markers. For *OMM5233* however, 7 of the 95 individuals produced genotyping discrepancies between first and second runs. This occurred as the shorter allele in many individuals was significantly lower in amplitude than the longer allele, and was as a consequence inconsistently scored. For *OMM5132*, a single genotyping error was observed. This individual was incorrectly genotyped as a homozygote in the first run. Consequently, with the exception of locus *OMM5233*, genotyping was highly consistent. This is despite the fact that four of the loci (*OMM1097*, *OMM1325*, *OMM5047*, *OMM5132*) displayed one or more alleles just a single base pair distant from a neighbouring allele.

Potential departure from Hardy Weinberg equilibrium was examined in all loci and samples. A total of 40 significant departures out of 228 tests were observed ( $P < 0.05$ ). Looking at loci first, with the exception of *OMM5233* (6 departures), and *OMM1051* (14 departures), departures were evenly distributed among loci (1–3 departures per locus). A similar pattern of distribution of departure from Hardy Weinberg equilibrium was observed for samples across loci. Most samples displayed 0–3 significant departures with the exception of sample 2 (5 departures), sample 14 (5 departures), and sample 15 (4 departures). Due to genotyping inconsistency and/or extensive departure from Hardy Weinberg equilibrium, both *OMM5233* and *OMM1051* were excluded from all further statistical analyses. Following this, and implementation of Bonferroni correction for multiple independent tests (new  $P = 0.005$ ), only two departures, both in sample 14, remained significant.

### Genetic variation within samples

The amount of genetic variation observed in each sample for each locus is presented (Table 3). A total of 136 alleles were observed in the entire data set (mean per locus = 11.3), ranging from 5–24 for loci across samples, and 65–103 for samples across loci. For the samples taken outside of the fjord where the escapees were captured (samples 1–7), total number of alleles per sample ranged from 65–93, whereas the total number of alleles for the samples collected on farms located in the fjord where escapees were captured (samples 8–18) ranged from 83–103. Individual samples contained 48–76% of the allelic variation observed in the pooled material. Observed heterozygosity averaged across all 12 loci ranged from 0.65–0.79 among all of the samples (Table 3).

**Table 3: Allelic variation observed at 12 loci in 19 samples of rainbow trout.**

Sample	N	Locus (all have OMM as pre-fix)												Allele summary			
		1008	1051	1088	1097	1303	1325	5007	5047	5132	5177	5233	5264	A <sub>T</sub>	A <sub>M</sub>	A <sub>A</sub>	Ho
1	47	4	6	7	8	6	4	6	5	6	4	8	4	68	5.7	66.2	0.67
2	46	6	8	7	10	9	7	8	7	9	5	8	5	89	7.4	84.9	0.74
3	39	4	6	7	9	5	4	7	3	7	4	6	3	65	5.4	63.7	0.68
4	47	5	11	8	13	8	6	7	6	12	4	8	5	93	7.8	90.2	0.77
5	44	6	7	6	9	7	4	7	4	9	4	10	5	78	6.5	72.9	0.71
6	46	5	9	6	9	8	5	10	4	9	4	4	5	78	6.5	75.3	0.73
7	46	4	5	7	10	6	5	7	3	7	4	7	3	68	5.7	65.3	0.65
8	46	5	8	7	12	9	3	8	7	7	4	9	4	83	6.9	77.3	0.70
9	42	5	9	7	13	8	5	7	6	8	4	9	4	85	7.1	79.3	0.70
10	43	6	13	8	11	11	6	10	6	12	5	7	5	100	8.3	93.8	0.77
11	46	6	12	9	10	13	5	8	7	12	5	7	5	99	8.3	93.1	0.78
12	41	6	10	8	14	11	5	11	7	11	5	7	5	100	8.3	94.8	0.74
13	30	6	11	8	11	11	6	8	7	9	4	8	4	93	7.8	92.4	0.77
14	46	6	10	8	14	11	5	9	7	10	5	10	4	99	8.3	92.7	0.77
15	45	6	12	8	14	11	6	11	7	9	5	8	5	102	8.5	94.0	0.77
16	45	6	11	8	15	11	6	11	7	11	5	7	5	103	8.6	95.9	0.77
17	43	6	11	8	11	10	6	10	7	10	5	9	5	98	8.2	92.8	0.77
18	43	6	10	8	15	11	6	11	8	11	5	8	4	103	8.6	97.2	0.76
19	35	6	11	8	13	12	5	9	6	11	4	9	5	99	8.3	96.8	0.79
Total	820	6	16	9	24	14	7	14	8	17	5	11	5	136	11.3	100.0	---

N = mean number of observations per locus, A<sub>T</sub> = total number of alleles, A<sub>M</sub> = mean number of alleles, A<sub>A</sub> = allelic richness based upon a minimum sample size of 28 diploid individuals, Ho = observed heterozygosity.

**Genetic variation among samples**

A matrix of genetic distances among all samples as estimated by  $F_{ST}$  and Nei's 1978 genetic distance [27] is presented (additional file 1). Pair-wise  $F_{ST}$  values ranged from 0–0.127, and in total, 21 of the 171 pair-wise comparisons were greater than 0.1. The greatest differentiation was observed between samples 5 and 7, and, the majority of the highest values (for example those over 0.1) involved samples 1, 5, 7. Genetic relationships among the samples are also presented using Nei's 1978 [27] genetic distance (Fig. 1). The genetic material reared in the three farms under suspicion (samples 8–18), grouped into two major clusters. One of the clusters was genetically distinct from all of the other samples in the total data set, and involved samples 10–18. Within this cluster, all pair-wise  $F_{ST}$  values were under 0.022, and the majority were smaller than 0.01. In addition, the vast majority of the few pair-wise sample comparisons that were not genetically differentiated according to the Fishers method involved samples from this cluster (additional file 1). The second major cluster involving the samples originating from the three farms under suspicion involved samples 8 and 9. This cluster also included control samples 1, 3 and 7.

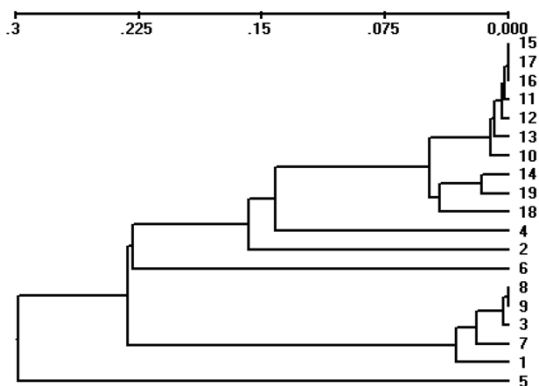
Pair-wise  $F_{ST}$  values between samples 8 and 9 were very small, and did not display significant differentiation (additional file 1). Likewise, pair-wise  $F_{ST}$  values between

samples 10 and 11 were very small, and lacked significance (additional file 1). These pairs of samples were both reported to have been the result of splitting fish from two to four cages prior to the NDF taking samples of the fish.

**Self-assignment simulations**

For the computation including samples 1–7, overall correct self-assignment was calculated at 83%, and ranged from 63–96% for individual samples (Table 4). When self-assignment was computed for samples 1–7, 8, 10, 12 and 15, overall correct self-assignment dropped to 61%, and, when computed for the entire data set (escapees excluded), correct overall self-assignment dropped further to 38%.

Miss-assignment was not random among the samples. Taking sub-set 1 first (Table 4), samples 1, 3 and 7 displayed a moderate degree of miss-assignment among themselves. This is consistent with the fact that samples 1, 3 and 7 represent part of a cluster which also includes samples 8 and 9 (Fig. 1). The relationship between pair-wise genetic distance and miss-assignment is further apparent in the partial matrix for sub-set 2 (Table 4), where sample 8 displayed a high degree of miss-assignment to samples 1, 3 and 7. However, in contrast to samples 10, 12 and 15, sample 8 did not display any significant degree of miss-assignment to the other cage



**Figure 1**  
**UPMGA diagram based upon Nei's (1978) unbiased genetic distance among 19 samples of rainbow trout.**

sampled on the same farm (sample 10) nor cages sampled in neighbouring farms (samples 12 and 15) originating from the same material that was reported to be mixed.

**Assignment of the escapees**

Assignment of the 35 rainbow trout escapees is presented (Table 5). Direct assignment to sample resulted in the 35 escapees being scattered evenly among the samples used to generate the baseline. This is consistent with the patterns of miss-assignment described above. Nevertheless, when data from the farms were combined (Table 5), 32 of the 35 escapees were directly placed into the samples collected from the farms located in the fjord where the escapees were captured. Direct assignment places each individual into the sample that it is genetically most similar to, irrespective of the degree of similarity. In the context of forensics, the most robust form of genetic identification is through exclusion rather than inclusion. Taking the samples collected from outside the fjord where the escapees were captured first, at a significance level of  $P = 0.05$ , all of the 35 escapees were excluded from samples 5 and 6, and, 26 of the 35 escapees were excluded from sample 7. For nine of the 11 samples collected on the three farms under suspicion, 19–23 of the 35 rainbow trout could not be excluded on an individual sample basis. However, for samples 8 and 9, only 10 and 11 of the escapees could originate from those samples at the calculated level of exclusion. Interestingly, samples 8 and 9 are those that display the greatest degree of genetic similarity to sample 7 (additional file 1, Fig. 1), where a similar level of exclusion was estimated. When samples from the three farms under suspicion were combined, only one out of 35 escapees could be excluded.

**Discussion**

Results of the "real-life" assignment for 35 escaped rainbow trout did not identify a single farm, or cage of origin.

This was expected, since the three farms operating within the fjord where the escapees were captured, shared genetic material. This was confirmed by the low pair-wise  $F_{ST}$  values among many of these samples, and the resultant miss-assignment among them when simulating with the data (Table 4). Nevertheless, all 35 escapees were excluded from two of the samples included as controls in the baseline for this analysis, and, the majority were excluded from the third control sample. As only one of the 35 escapees could be excluded from having originated from one or more of the three farms located in the fjord where the escapees were captured, and the nearest other rainbow trout farm was >130 km distance by sea, the genetic analyses presented here provide clear evidence that the rainbow trout escapees probably originated from one or several of the three farms under suspicion. As a single company owned all three farms, this permitted the Norwegian police to initiate an investigation of the company.

Significant among-sample variation is a pre-requisite for genetic assignment, with greater differentiation increasing accuracy [28-30]. The  $F_{ST}$  values observed among samples in the present study are similar to those observed among some of the major Atlantic salmon breeding strains used in aquaculture [19], and, similar to those observed among cage reared salmon [18]. Furthermore, whilst the overall correct self-assignment reported here (83%) among the 7 samples collected outside the location where escapees were captured, is slightly less than reported for 12 European rainbow trout strains (92%) [15], it is greater than observed by Glover et al. [18] in a study of cage-reared Atlantic salmon. Comparison to the latter study is important, because it described, for the first time, use of genetic assignment to identify the cage and farm of origin for a group of escaped farmed Atlantic salmon.

Throughout the world, over 75 strains of rainbow trout have been produced in culture [31], and over 65 countries have reported aquaculture activity of this species [1]. Norway is one of the largest producers, accounting for approximately 15% of global production in 2006 [2]. In 2008, when samples for this study were collected, the majority of the Norwegian production was based upon five strains (Aqua Gen AS, Salmobreed AS, Ilsvåg, Alf Lone and Stolt sea farms). Data from this study demonstrated variable but nevertheless significant genetic differentiation among samples of fish reported to originate from these strains, with clear opportunities for genetic assignment.

The degree of genetic differentiation observed among the samples in this study (highest pair-wise  $F_{ST} = 0.127$ ) is within the range of values observed in other studies of rainbow trout throughout the world. Among 12 European strains, Gross et al. [15] reported an average  $F_{ST}$  of 0.14. When looking specifically among Danish and Finnish

**Table 4: Results of self-assignment tests.**

Sub-set 1													
Sample	1	2	3	4	5	6	7	Correctly assigned					
1	<b>32</b>	-	12	-	-	-	3	68%					
2	-	<b>43</b>	-	2	-	-	1	93%					
3	3	1	<b>28</b>	-	-	-	7	72%					
4	-	2	-	<b>45</b>	-	-	-	96%					
5	1	-	-	-	<b>40</b>	3	-	91%					
6	-	1	-	2	-	<b>43</b>	-	94%					
7	11	-	6	-	-	-	<b>29</b>	63%					
Overall								83%					

Sub-set 2													
Sample	1	2	3	4	5	6	7	8	10	12	15	Correctly assigned	
8	3	-	15	-	-	-	<b>8</b>	<b>18</b>	-	-	1	40%	
10	-	2	-	1	-	1	-	-	<b>14</b>	17	<b>8</b>	36%	
12	-	1	-	3	-	2	-	-	10	<b>18</b>	<b>8</b>	43%	
15	-	3	-	1	1	-	-	-	12	14	<b>14</b>	31%	
Overall												37%	

Sub-set 1 represents a full matrix of self-assignment among samples 1–7. Sub-set 2 represents a partial matrix of self-assignment for samples 8, 10, 12 and 15 combined in sub-set 1. Numbers in bold represent exact numbers of fish correctly assigned to their sample of origin.

strains however, these authors reported a mean  $F_{ST}$  of 0.1 and 0.05 respectively. Among 24 anadromous samples in California, Aguilar & Garza [11] reported the largest pair-wise  $F_{ST}$  value of 0.171, whilst Deiner et al. [32] reported pair-wise  $F_{ST}$  values as high as 0.385 among 20 land-locked populations in California. Johnson et al. [12] reported an  $F_{ST}$  of 0.0481 between two hatchery strains of

rainbow trout in West Virginia, whilst Silverstein et al. [14] reported an average  $F_{ST}$  of 0.089 among three distinct hatchery strains in USA.

It is difficult to make direct comparisons of genetic variation to other studies where different marker sets have been used. However, a tentative comparison with the 12 Euro-

**Table 5: Assignment of 35 captured rainbow trout escapees.**

Single samples														
8	9	10	11	12	13	14	15	16	17	18	5	6	7	
<u>Direct assignment</u>														
6	4	2	1	3	2	1	3	5	2	3	0	1	2	
<u>Not excluded at 0.05</u>														
10	11	19	20	20	19	21	20	22	20	23	0	0	9	
Pooled samples														
Farm 7 (samples 8–11)				Farm 8 (samples 12–14)			Farm 9 (samples 15–18)			Suspected farms (samples 8–18)			Control farms (samples 5–7)	
<u>Direct assignment</u>														
13				6			13			32			3	
<u>Not excluded at 0.05</u>														
31				31			28			34			9	

The genetic baseline comprises of 11 samples taken from the three farms in the fjord where the escapees were captured, and 3 control samples taken from two farms outside the region.

pean strains analysed by Gross et al. [15], analysed by 10 microsatellite loci (no marker overlap with present study) indicate a similar range of mean number of alleles per sample between studies (3.2–7.2 for 12 European strains, 5.4–7.8 for samples 1–7 in this study). In a study of three domesticated rainbow trout strains in USA, Silverstein et al. [14] reported mean number of alleles per locus per sample of 8.1–9.6, with 9 microsatellite loci. A direct comparison with the study of two hatchery strains maintained at the National Centre for Cool and Cold Water Aquaculture in USA (NCCCWA) [12] is possible as nine of the loci analysed by these authors overlapped with loci used here. These authors reported an average number of alleles per strain from 11.1–13.6, which is much higher than the values reported here. This was also clear on an individual locus basis. For example, 15–22 alleles per strain for locus *OMM1097* was reported in the strains reared in the NCCCWA, whilst only 8–13 alleles were observed per sample for samples 1–7 in the present study. Whilst 15 alleles were observed for *OMM1097* in samples 16 and 18, these were reported to represent a mixture of strains. Consequently, although allelic variation is positively linked with sample size [33,34], and the sample sizes screened by Johnson et al. [12] ranged from 60–67, in contrast to 39–47 in the present study, it cannot be excluded that rainbow trout strains used in Norwegian aquaculture display reduced allelic diversity compared to the NCCCWA strains. This is consistent with the manner by which rainbow trout have been distributed beyond their native range [31,35], and the potential for multiple founder effects and/or bottlenecks. However, it is also possible that the observed differences may be a consequence of sampling. Samples taken at the NCCCWA were obtained directly from the breeding stock, whilst in the present study, samples were taken from production units, which may not necessarily contain all of the genetic variation present in the breeding nucleus. This requires further elucidation.

In the present study, two samples (4 and 7) were reported to originate from the Aqua Gen strain. Both pair-wise  $F_{ST}$  values and total number of alleles varied greatly between the two samples. Whilst it cannot be excluded that this is a result of incorrect information regarding the origin of these samples, or mixing of genetic material from several sources, it is possible that this is a consequence of two different smolt-producers delivered rainbow trout to the two marine farms where samples 4 and 7 were collected. Genetic differences between samples originating from the same strain can be generated by non-random sorting of gametes, fertilised eggs, fry, and smolts in the process of distributing genetic material from the central breeding stations, to marine producers, as has been previously suggested [18].

The detailed examination of genetic variation in and among all cages of rainbow trout collected from the three farms under suspicion, highlight the possibility to use genetic tools to monitor production. For example, this could assist in the reconstruction and verification of fish movements among cages and sites, check for illicit (re-)production of a specific strain by a competing company or customer, or confirm genetic background to a new delivery of fish. This potential is best illustrated by looking at the four samples taken on farm 7 (samples 8–11). Samples 8 and 9 were collected from two adjacent cages, and were reported to have been recently split from a single cage prior to sampling. This was also reported for the pair of samples 10 and 11. The very low  $F_{ST}$  values observed between samples 8 and 9, and, 10 and 11, combined with the almost identical total number of alleles observed for each pair of samples (83 and 85, 100 and 99, for samples 8, 9, 10 and 11 respectively), are consistent with the information supplied by the farmer. However, the very clear genetic differences between pairs of samples 8 and 9 vs. 10 and 11, as revealed by  $F_{ST}$  and allelic variation, provided extra information. Clearly, these two groups of fish did not represent a homogenous mixture of fish from a single source.

## Conclusion

This study represents the first molecular genetic characterisation of farmed rainbow trout in Norway, one of the worlds largest producers. A significant amount of genetic variation was observed both within, and, among the samples examined. Among sample variation reflected genetic differences between groups of fish reared on different farms, in different cages on the same farm, and among strains. In the "real-life" assignment of 35 rainbow trout escapees, with the exception of one individual, none could be excluded from originating in one or several of the three farms operating in the fjord where they were captured. In contrast, two of the samples used as a control were excluded as potential donors for all of the escapees, and the third sample used as a control was excluded for the majority of the escapees. These data were used by the Norwegian police to initiate an investigation of the company operating the three farms in the fjord. This study, has clearly demonstrated potential for genetic assignment of rainbow trout in aquaculture, providing farmers with the ability to monitor various aspects of production, and management authorities with a tool with which to trace escapees.

## Authors' contributions

KAG designed the study, supervised laboratory work, carried out data analysis, and wrote the article.



## Additional material

### Additional file 1

Pair-wise genetic distances among 19 samples of rainbow trout. This matrix of values includes estimations by  $F_{ST}$  in the lower diagonal and Nei's (1978) unbiased genetic distance in the upper diagonal.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2156-9-87-S1.doc>]

## Acknowledgements

Employees at the NDF are acknowledged for collecting samples used in this study, and the farmers for their cooperation. Harald Breilid, Anne-Grete Søvik, and Bjørghild Seliussen are gratefully acknowledged for conducting all laboratory work under supervision of KAG. John B. Taggart and Øystein Skaala are acknowledged for comments on this work. This study was financed by the Institute of Marine Research in Bergen as part of their basic funding from the Norwegian government.

## References

1. FAO: **Cultured Aquatic Species Information Programme × *Oncorhynchus mykiss***. In **FAO Fisheries and Aquaculture Department** [online]. Rome. Updated 15 Jun 2005. [Cited 11 Jul 2008]. [[http://www.fao.org/fishery/culturedspecies/Oncorhynchus\\_mykiss](http://www.fao.org/fishery/culturedspecies/Oncorhynchus_mykiss)].
2. Anon: **The ministry of Fisheries and Coastal Affairs. Facts about Fisheries and Aquaculture**. [<http://www.fisheries.no>].
3. Gjedrem T: **Flesh quality improvement in fish through breeding**. *Aquaculture International* 1997, **5(3)**:197-206.
4. Gjedrem T: **Genetic improvement of cold-water fish species**. *Aquaculture Research* 2000, **31(1)**:25-33.
5. Garber AF, Sullivan CV: **Selective breeding for the hybrid striped bass (*Morone chrysops*, *Rafinesque* × *M-saxatilis*, *Walbaum*) industry: status and perspectives**. *Aquaculture Research* 2006, **37(4)**:319-338.
6. Kettunen A, Fjalestad KT: **Genetic parameters for important traits in the breeding program for Atlantic cod (*Gadus morhua* L.)**. *Aquaculture* 2007, **272**:276.
7. de Leon FJG, Canonne M, Quillet E, Bonhomme F, Chatain B: **The application of microsatellite markers to breeding programmes in the sea bass, *Dicentrarchus labrax***. *Aquaculture* 1998, **159(3-4)**:303-316.
8. Nielsen JL, Fountain MC: **Microsatellite diversity in sympatric reproductive ecotypes of Pacific steelhead (*Oncorhynchus mykiss*) from the Middle Fork Eel River, California**. *Ecology of Freshwater Fish* 1999, **8(3)**:159-168.
9. Gregg RE, Rhodes OE, Armstrong G: **A genetic analysis of the London strain of rainbow trout**. *Animal Genetics* 2001, **32(4)**:210-214.
10. Narum SR, Contor C, Talbot A, Powell MS: **Genetic divergence of sympatric resident and anadromous forms of *Oncorhynchus mykiss* in the Walla Walla River, USA**. *Journal of Fish Biology* 2004, **65(2)**:471-488.
11. Aguilar A, Garza JC: **A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum)**. *Molecular Ecology* 2006, **15(4)**:923-937.
12. Johnson NA, Rexroad CE, Hallerman EM, Vallejo RL, Palti Y: **Development and evaluation of a new microsatellite multiplex system for parental allocation and management of rainbow trout (*Oncorhynchus mykiss*) broodstocks**. *Aquaculture* 2007, **266(1-4)**:53-62.
13. Hauser L, Seamons TR, Dauer M, Naish KA, Quinn TP: **An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchus mykiss*) in Forks Creek, Washington, USA**. *Molecular Ecology* 2006, **15(11)**:3157-3173.
14. Silverstein JT, Rexroad CE, King TL: **Genetic variation measured by microsatellites among three strains of domesticated rainbow trout (*Oncorhynchus mykiss*, Walbaum)**. *Aquaculture Research* 2004, **35(1)**:40-48.
15. Gross R, Lulla P, Naver T: **Genetic variability and differentiation of rainbow trout (*Oncorhynchus mykiss*) strains in northern and Eastern Europe**. *Aquaculture* 2007, **272**:139-146.
16. Baarøy V, Gjerde B, Heggberget TG, Jensen PE, Maroni K, Sandvik S, Skaala Ø: **Identifisering av rømt oppdrettlaks. Utredning av utvalg nedsett av Fiskeridirektøren (Identification of escaped farmed salmon. Report from the committee to the Director of Fisheries)**. In **Norwegian**. 2004.
17. Coughlan J, McGinnity P, O'Farrell B, Dillane E, Diserud O, de Eyto E, Farrell K, Whelan K, Stet RJM, Cross TF: **Temporal variation in an immune response gene (MHC I) in anadromous *Salmo trutta* in an Irish river before and during aquaculture activities**. *ICES Journal of Marine Science* 2006, **63(7)**:1248-1255.
18. Glover KA, Skilbrei OT, Skaala Ø: **Genetic assignment identifies farm of origin for Atlantic salmon (*Salmo salar* L.) escapees in a Norwegian fjord**. *ICES Journal of Marine Science* 2008, **65**:1-9. advanced access
19. Skaala O, Hoyheim B, Glover K, Dahle G: **Microsatellite analysis in domesticated and wild Atlantic salmon (*Salmo salar* L.): allelic diversity and identification of individuals**. *Aquaculture* 2004, **240(1-4)**:131-143.
20. Skilbrei OT, Wennevik V: **The use of catch statistics to monitor the abundance of escaped farmed Atlantic salmon and rainbow trout in the sea**. *ICES Journal of Marine Science* 2006, **63(7)**:1190-1200.
21. Pierce LR, Palti Y, Silverstein JT, Barrows FT, Hallerman EM, Parsons JE: **Family growth response to fishmeal and plant-based diets shows genotype × diet interaction in rainbow trout (*Oncorhynchus mykiss*)**. *Aquaculture* 2008, **278(1-4)**:37-42.
22. Hoffman JL, Amos W: **Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion**. *Molecular Ecology* 2005, **14(2)**:599-612.
23. Pompanon F, Bonin A, Bellemain E, Taberlet P: **Genotyping errors: Causes, consequences and solutions**. *Nature Reviews Genetics* 2005, **6(11)**:847-859.
24. Raymond M, Rousset F: **Genepop (Version-1.2) × Population-Genetics Software for Exact Tests and Ecumenicism**. *Journal of Heredity* 1995, **86(3)**:248-249.
25. Goudet J: **FSTAT (Version 1.2): A computer program to calculate F-statistics**. *Journal of Heredity* 1995, **86(6)**:485-486.
26. Miller MP: **Tools for population genetic analysis v1.3**. 1999 [<http://www.marksgeneticssoftware.net/tfpga.htm>].
27. Nei M: **Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals**. *Genetics* 1978, **89(3)**:583-590.
28. Cornuet JM, Piry S, Luikart G, Estoup A, Solognac M: **New methods employing multilocus genotypes to select or exclude populations as origins of individuals**. *Genetics* 1999, **153(4)**:1989-2000.
29. Mank JE, Avise JC: **Individual organisms as units of analysis: Bayesian-clustering alternatives in population genetics**. *Genetical Research* 2004, **84(3)**:135-143.
30. Paetkau D: **The optimal number of markers in genetic capture-mark-recapture studies**. *Journal of Wildlife Management* 2004, **68(3)**:449-452.
31. MacCrimmon HR: **World Distribution of Rainbow Trout (*Salmo-Gairdneri*)**. *Journal of the Fisheries Research Board of Canada* 1971, **28(5)**:663-704.
32. Deiner K, Garza JC, Coey R, Girman DJ: **Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California**. *Conservation Genetics* 2007, **8(2)**:437-454.
33. Sjogren P, Wyoni PI: **Conservation Genetics and Detection of Rare Alleles in Finite Populations**. *Conservation Biology* 1994, **8(1)**:267-270.
34. Leberg PL: **Estimating allelic richness: Effects of sample size and bottlenecks**. *Molecular Ecology* 2002, **11(11)**:2445-2449.
35. Gall GAE, Crandell PA: **The Rainbow-Trout**. *Aquaculture* 1992, **100(1-3)**:1-10.
36. Coulibaly I, Gharbi K, Danzmann RG, Yao J, Rexroad CE: **Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags**. *Animal Genetics* 2005, **36(4)**:309-315.

37. Rexroad CE, Rodriguez MF, Coulibaly I, Gharbi K, Danzmann RG, DeKoning J, Phillips R, Palti Y: **Comparative mapping of expressed sequence tags containing microsatellites in rainbow trout (*Oncorhynchus mykiss*)**. *BMC Genomics* 2005, **6(1)**:54.
38. Palti Y, Fincham MR, Rexroad CE: **Characterization of 38 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*)**. *Molecular Ecology Notes* 2002, **2(4)**:449-452.
39. Rexroad CE, Coleman RL, Gustafson AL, Hershberger WK, Killefer J: **Development of rainbow trout microsatellite markers from repeat enriched libraries**. *Marine Biotechnology* 2002, **4(1)**:12-16.
40. Rexroad CE, Coleman RL, Hershberger WK, Killefer J: **Rapid communication: Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout**. *Journal of Animal Science* 2002, **80(2)**:541-542.

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

