

Evidence of neutral and adaptive genetic divergence between European trout populations sampled along altitudinal gradients

I. KELLER,*† A. TAVERNA* and O. SEEHAUSEN*‡

*Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Center of Ecology, Evolution and Biochemistry, Seestrasse 79, CH-6047 Kastanienbaum, Switzerland, †Department of Aquatic Ecology, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland, ‡Department of Aquatic Ecology and Macroevolution, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland

Abstract

Species with a wide geographical distribution are often composed of distinct subgroups which may be adapted to their local environment. European trout (*Salmo trutta* species complex) provide an example of such a complex consisting of several genetically and ecologically distinct forms. However, trout populations are strongly influenced by human activities, and it is unclear to what extent neutral and adaptive genetic differences have persisted. We sampled 30 Swiss trout populations from heterogeneous environments along replicated altitudinal gradients in three major European drainages. More than 850 individuals were genotyped at 18 microsatellite loci which included loci diagnostic for evolutionary lineages and candidate markers associated with temperature tolerance, reproductive timing and immune defence. We find that the phylogeographic structure of Swiss trout populations has not been completely erased by stocking. Distinct genetic clusters corresponding to the different drainages could be identified, although nonindigenous alleles were clearly present, especially in the two Mediterranean drainages. We also still detected neutral genetic differentiation within rivers which was often associated with the geographical distance between populations. Five loci showed evidence of divergent selection between populations with several drainage-specific patterns. Lineage-diagnostic markers, a marker linked to a quantitative trait locus for upper temperature tolerance in other salmonids and a marker linked to the major histocompatibility class I gene were implicated in local adaptation and some patterns were associated with altitude. In contrast, tentative evidence suggests a signal of balancing selection at a second immune relevant gene (TAP2). Our results confirm the persistence of both neutral and potentially adaptive genetic differences between trout populations in the face of massive human-mediated dispersal.

Keywords: candidate loci, environmental gradient, genome scan, local adaptation, population structure, trout

Received 24 June 2010; revision received 31 January 2011; accepted 9 February 2011

Introduction

Many species are geographically widespread and their genetic structure will be heavily influenced by land-

scape features that act as barriers to dispersal. Even within continuous habitats, populations may face a range of biotic and abiotic conditions, and evolutionary ecology provides plenty of examples of local adaptation to such heterogeneous environments (Endler 1977; Herford 2009). While local adaptation is perhaps most often demonstrated between allopatric populations, it

Correspondence: Irene Keller, Fax: +41 44 823 5028; E-mail: irene.keller@eawag.ch

can, under more restrictive conditions, arise in the absence of geographical isolation, i.e. in parapatry or sympatry (e.g. Schluter 2000). The genes underlying adaptation are unknown in most systems (for exceptions see e.g. Colosimo *et al.* 2005; Seehausen *et al.* 2008; Steiner *et al.* 2009; Bernatchez *et al.* 2010) but would be of great interest especially in cases where adaptive traits are not immediately phenotypically apparent. The spatial and ecological distribution of particular alleles could then be used to characterize adaptive diversity across natural populations at a resolution not amenable through experiments. To predict the response of natural populations to changing environments, it will be crucial to know which components of adaptive diversity are maintained between and which within populations and sites (Barrett & Schluter 2008).

Although evidence of local adaptation has been reported from other salmonid fishes (e.g. Vasemägi *et al.* 2005; Dionne *et al.* 2008), the extent of divergent adaptation between populations of European trout (*Salmo trutta* species complex) is difficult to predict. On the one hand, trout are very widely distributed and harbour extensive genetic diversity which is expected to promote local adaptation. On the other hand, adaptive divergence between populations may be impeded by potentially high levels of gene flow owing to natural dispersal and, perhaps more importantly, human management practices.

Based on mitochondrial sequence diversity, Bernatchez (2001) described five evolutionarily distinct European trout lineages which probably split in the Pleistocene some 0.5–2 Ma and are mostly considered separate species in a recent taxonomic reassessment (Kottelat & Freyhof 2007; *S. trutta* = Atlantic lineage of Bernatchez (2001); *Salmo rhodanensis* = Mediterranean; *Salmo cenerinus* = Adriatic; *Salmo marmoratus* = marmorata; *Salmo labrax* = Danubian). All five evolutionary lineages are expected in Switzerland and, historically, their distribution would have been largely allopatric with each major drainage system containing one native trout form. An exception were the rivers from the Po system where Adriatic and marble trout co-occurred (Giuffra *et al.* 1994, 1996; Meraner *et al.* 2007).

Each year, Swiss rivers are stocked with several million captive-reared trout (Swiss Federal Office for the Environment). In the past, these measures often relied on hatchery fish from abroad (e.g. Denmark) and also included translocations across major drainage boundaries. In particular, Atlantic trout were introduced in large numbers into the Doubs (Rhône) and Po drainages, and previous allozyme-based studies suggest that they hybridized extensively with the native lineages (Largiadèr & Scholl 1995, 1996).

In Switzerland, as elsewhere in Europe, trout are common along a very broad altitudinal gradient from the low-elevation midlands to alpine rivers. No other fish taxon in Europe occupies a similarly broad altitudinal niche. Water temperature is one obvious (but clearly not the only) abiotic factor that will differ markedly along altitudinal gradients and will have very direct effects on ectotherms (Angilletta 2009), influencing not only the rate of biochemical reactions but also processes at the level of the whole organism such as rates of locomotion, growth and development (Ohlberger *et al.* 2008; Kingsolver 2009). The biological interactions experienced by a population are also expected to vary strongly with altitude. The number of potentially competing fish species decreases rapidly with increasing elevation. Moreover, numerous studies have shown that both the distribution range and the virulence of parasites and pathogens may increase at higher temperatures (e.g. Marcogliese 2008). A specific example is that of proliferative kidney disease (PKD), a salmonid disease mediated by a myxozoan parasite, for which serious outbreaks are typically observed only above a certain temperature threshold (Hari *et al.* 2006; Wahli *et al.* 2007).

The diverse nature of the environmental changes expected along altitudinal gradients implies that adaptive genetic differentiation probably involves a range of traits. Of particular importance may be the tolerance to high temperatures and the timing of reproduction, traits known to exhibit genetically based intraspecific variation in salmonids (Jackson *et al.* 1998; Hodgson & Quinn 2002; O'Malley *et al.* 2003; Somorjai *et al.* 2003; Carlson & Seamons 2008). An important role is also expected for traits related to immune defence.

Central genes in the vertebrate immune system are those of the major histocompatibility complex (MHC), which code for molecules responsible for antigen presentation. While the often exceptional polymorphism of these genes is probably maintained by some form of balancing selection, the MHC has also been implicated in local adaptation in a range of species (reviewed e.g. in Klein 1986; Muirhead 2001; Bernatchez & Landry 2003; Garrigan & Hedrick 2003; Piertney & Oliver 2006; Wegner 2008; Spurgin & Richardson 2010), including several examples from salmonids (Landry & Bernatchez 2001; Vasemägi *et al.* 2005; Aguilar & Garza 2006; Dionne *et al.* 2007; Hansen *et al.* 2007). A role for divergent adaptation can be envisaged if environments differ in parasite/pathogen communities and MHC alleles exhibit at least some specificity against particular challenges (Palti *et al.* 2001; Miller *et al.* 2004). Other non-MHC genes may contribute to immune defence (Acevedo-Whitehouse & Cunningham 2006; Tonteri *et al.* 2010), including the Transporter associated with Antigen

Processing (TAP2; Grimholt 1997; Grimholt *et al.* 2002) involved in the transport of foreign peptides into the endoplasmic reticulum where they are loaded onto MHC I molecules (e.g. Abele & Tampé 2006). Generally, little is known about TAP diversity in natural populations, but Jensen *et al.* (2008) report evidence of spatially divergent and temporally fluctuating selection in brown trout. This finding is in line with observations from genomic model organisms that several viruses are able to suppress TAP function (Abele & Tampé 2006) which suggests the potential for strong pathogen-mediated selection.

In this study, we investigate the patterns of within- and between-population diversity in trout populations sampled along replicated altitudinal gradients on both sides of the Alps using microsatellite markers. Our marker panel includes two loci with alleles diagnostic for some of the major trout lineages (Estoup *et al.* 2000), microsatellites linked to salmonid quantitative trait loci (QTL) for reproductive timing and temperature tolerance and markers close to the MHC class I and II genes and both copies of the duplicated TAP2 gene. First, we describe the neutral genetic substructure between and within drainages, and then use outlier scan approaches (e.g. Beaumont & Balding 2004; Storz 2005) to identify loci showing unusually elevated or reduced population divergence consistent with the action of selection. Finally, we investigate whether the allele frequencies at MHC-linked loci are more even than expected at mutation-drift equilibrium, which could suggest some form of balancing selection.

Materials and methods

Sampling

Trout (*S. trutta* species complex) samples were collected in 2007 and 2008 from 30 sites in the Rhine, Rhone and Po drainage systems in Switzerland and northern Italy (Fig. 1). A single stream, the Allaine, was studied in the Doubs sub-drainage of the Rhone system. We sampled several smaller rivers from the Rhine and two distinct sub-drainages from the Po, the Ticino and the Poschiavino, which were separated by a large waterway distance. Sampling sites were arranged along altitudinal transects replicated in different rivers and spanned a range of *c.* 300–1800 m above sea level (Fig. 1, Table S1, Supporting information). The maximum water temperatures at these sites differ by at least 10 °C (our own unpublished data). The animals were caught by electro-fishing within a short river section of typically *c.* 100–200 m. Tissue samples were taken from a fin and stored in absolute ethanol. We specifically targeted large fish to avoid sampling recently introduced hatchery fish but

included smaller individuals if necessary. Sample size was at least 25 individuals in all but two sites (Table S1, Supporting information), resulting in a total of 853 individuals.

Choice of molecular markers

We genotyped all individuals at 18 microsatellite loci selected from the literature (Tables 1 and S2, Supporting information). Eight markers with no known functional association were selected from different linkage groups based on the trout linkage map (Gharbi *et al.* 2006). For two of them, one (Str85-reverse) or both (Str543) primers were newly designed based on the published sequence (accession nos AB001059 and AB001062) to change the length of the amplified fragment. Two diagnostic markers (Table 1) had previously been shown to have nonoverlapping allele size ranges in three of the Pleistocene trout lineages (Atlantic, Mediterranean and marble trout; Estoup *et al.* 2000) and can hence provide both an estimate of some of the historical genetic diversity present in a particular system and the phylogenetic affiliation of a population. Finally, nine candidate loci were genotyped (one only in a subset of individuals), five of which were microsatellites linked to quantitative trait loci (QTLs) for upper temperature tolerance or spawning time in other salmonids (Tables 1 and S2, Supporting information).

The remaining four candidate loci were located in or near the MHC class I and class II regions. In contrast to other vertebrates, the two regions map to different chromosomes in teleost fishes and, hence, may respond independently to selection (Stet *et al.* 2003). In Atlantic salmon (*Salmo salar*), the class I region has been duplicated (locus IA and IB) but only the MHC gene within region IA is functional (Grimholt *et al.* 2002; Lukacs *et al.* 2007). Both regions seem to contain functional copies of the Transporter associated with Antigen Processing (TAP2a and TAP2b; Grimholt 1997; Grimholt *et al.* 2002).

In region IA, we genotyped markers UBA (3' untranslated tail of MHC I gene) and TAP2a (intron 5 of TAP2a) which, in the closely related Atlantic salmon, are separated by 28 kb (Genbank EF210363; BAC sequence from Lukacs *et al.* 2007). Marker TAP2b is in intron 5 of the second copy of TAP2 (MHC region IB) in a different linkage group (Grimholt *et al.* 2002). Note that we adhere to the nomenclature of Lukacs *et al.* (2007), with TAP2a corresponding to TAP2B and TAP2b to TAP2A of Grimholt *et al.* (2002). A microsatellite marker in very close physical proximity to the MHC class II region of Atlantic salmon (Ssa60; Gharbi *et al.* 2009) was genotyped in a subset of 461 individuals from 16 populations.

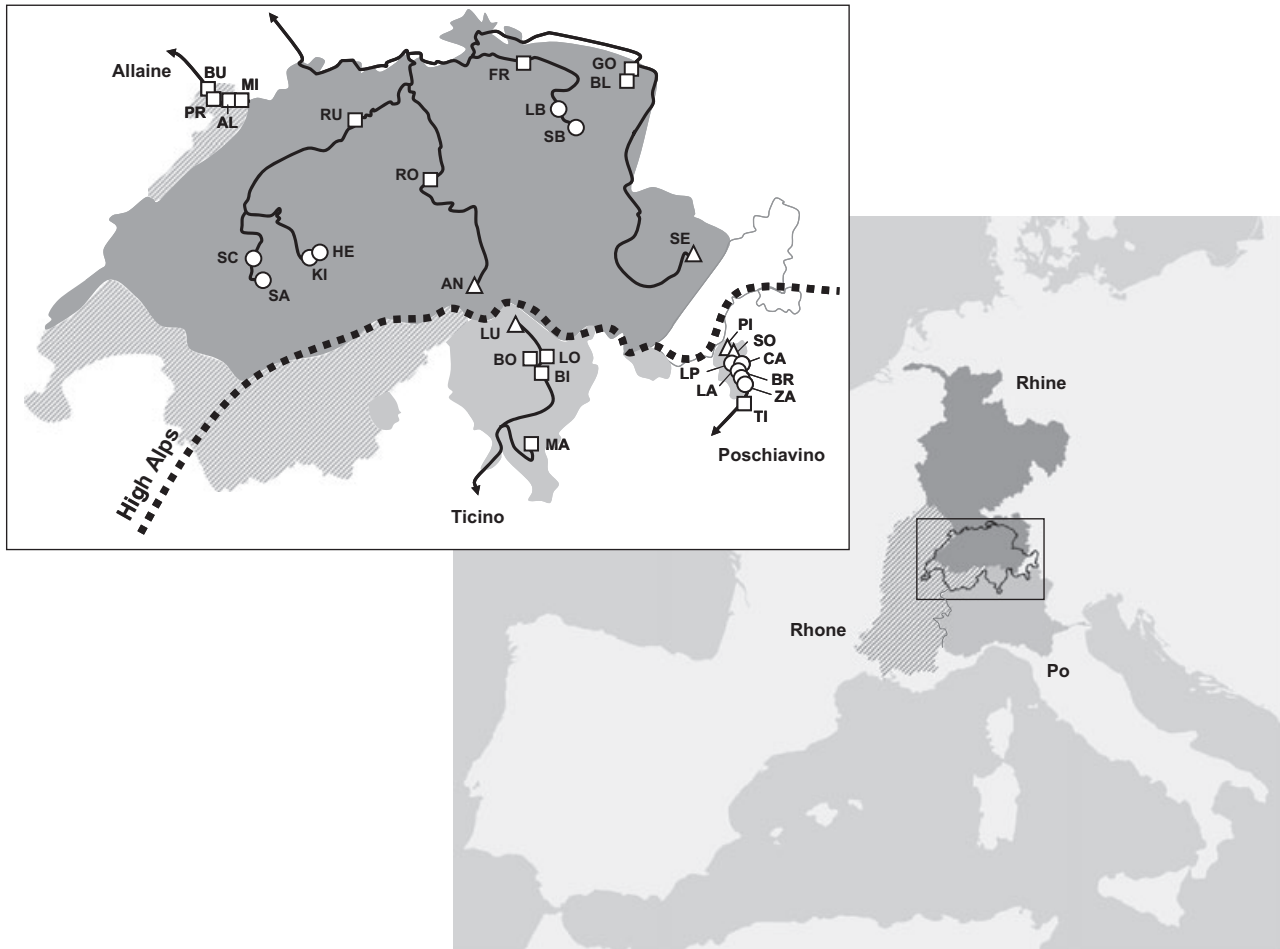


Fig. 1 Sampling sites in Switzerland and Northern Italy (TI) along altitudinal transects replicated in different rivers in three major European drainage systems. Only one stream (Allaine) was studied in the Rhone drainage. Within the Po system, samples were collected from the Poschiavino and the Ticino subdrainages. Different symbols indicate different altitudinal zones: squares <500 m, circles = 500–1000 m, triangles >1000 m. The investigated altitudinal range was 381–486 m asl in the Allaine, 398–1776 m asl in the Rhine, 430–1336 m asl in the Poschiavino, and 277–1802 m asl in the Ticino. The site labels are abbreviations of the names of nearby villages. The full names together with the altitude and coordinates of each sampling location are given in Table S1 (Supporting information).

Molecular methods

DNA was extracted on a robot with a BioSprint 96 DNA Blood Kit (Qiagen GmbH, Hilden, Germany) according to the instructions of the manufacturer and quantified on a NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, USA). The markers were amplified in three multiplex PCRs in a final volume of 12.5 μ L using a Techne TC412 thermal cycler (Witec AG). Each reaction contained 6.25 μ L multiplex PCR mastermix (Qiagen), 1.25 μ L primer mix with individual primer concentrations as indicated in Table S2 (Supporting information), and 50 ng of DNA. An initial denaturation step of 15 min at 95 $^{\circ}$ C was followed by 35 cycles of 30 s at 94 $^{\circ}$ C, 90 s at a multiplex specific

annealing temperature, 60 s at 72 $^{\circ}$ C, and a final elongation step of 30 min at 60 $^{\circ}$ C. Annealing temperatures were 60 $^{\circ}$ C for multiplex 1, 54 $^{\circ}$ C for multiplex 2 and 52 $^{\circ}$ C for multiplex 3. The PCR products were diluted 1:50 in water and resolved on a Beckman Coulter CEQ 8000 sequencer.

Statistical analyses

Input files for the different analysis programs were created with PGDSPIDER version 2.0.0.2 available from <http://www.cmpg.unibe.ch/software/PGDSpider/>. The frequency of null alleles was estimated for all loci in FREEANA (Chapuis & Estoup 2007). We tested all loci in all populations for deviations from Hardy–Weinberg

Table 1 Microsatellite markers used in this study. Loci are grouped as candidate, diagnostic or anonymous based on information from the literature. See Table S2 (Supporting information) for additional information

Type	Locus name	Putative function
Anonymous	SsoSL438	None
	Ssa100NVH	None
	Str15INRA	None
	Str60INRA	None
	Strutta58	None
	Str73INRA	None
	Str85INRA	None
	Str543INRA	None
Diagnostic	JMS1_M1	None; different allele size range in AT, MA and ME trout
	JMS2_M1	None; different allele size range in AT, MA and ME trout
Candidate	Ssa60NVH	Antigen presentation (linked to MHC II, locus DAA, in Atlantic salmon)
	Ssa-UBA	Antigen presentation (3' untranslated tail of MHC IA in Atlantic salmon)
	Ssa-TAP2a	Antigen presentation (intron 5 of Transporter associated with Antigen Processing, region MHC IA, of Atlantic salmon)
	Ssa-TAP2b	Antigen presentation (intron 5 of Transporter associated with Antigen Processing, region MHC IB, of Atlantic salmon)
	Omy325	QTL for UTT and body size in rainbow trout, QTL for growth rate and body size in coho salmon
	Ssa14	QTL for UTT in rainbow trout and Arctic charr
	Ssa85DU	QTL for UTT and spawning date in rainbow trout and Arctic charr
Ots515NWFSC	QTL for spawning date and body weight in rainbow trout	
SsoSL417	QTL for spawning data in rainbow trout	

AT, Atlantic; MA, marble; ME, Mediterranean lineage of Bernatchez (2001); QTL, quantitative trait locus; UTT, upper temperature tolerance; MHC, major histocompatibility complex.

equilibrium (HWE) in ARLEQUIN version 3.1 (Excoffier *et al.* 2005) using 100 000 steps in the Markov chain and 1000 dememorisation steps. For each locus, various measures of genetic diversity (number of alleles, allele and genotype frequencies) were calculated in FSTAT version 2.9.3.2 (Goudet 2001). We tested for deviations from linkage equilibrium between the two physically linked loci, UBA and TAP2a, in ARLEQUIN 3.1. with 10 000 permutations and 10 initial conditions for the EM algorithm.

Identification of loci potentially under selection. We used outlier scan approaches to identify loci showing exceptionally high or low genetic differentiation (measured as F_{ST}) compared to the genomic average. Analyses were performed for the global data set using an approach based on Beaumont & Nichols (1996) and implemented with an extension to hierarchically structured populations in ARLEQUIN v3.5 (Excoffier *et al.* 2009). If unaccounted for, hierarchical structure can lead to strongly elevated false-positive rates in outlier scans (Excoffier *et al.* 2009). The null distribution of F_{ST} as a function of heterozygosity was simulated under a finite island model with 100 islands, a stepwise mutation model (SMM) and using 20 000 realizations. The number of simulated groups was chosen to reflect the num-

ber of main river systems in a data set. The two geographically disjunct regions sampled within the Po drainage (Poschiavino and Ticino) were treated as separate groups, a subdivision consistent not only with geography but also with the observed genetic substructure (see below). If the average F_{ST} was heavily influenced by strong outliers, simulations were repeated without these loci to evaluate the observed F_{ST} values for the remaining loci. Significance was assessed at the 1% level.

Within individual drainages, analyses were performed in ARLEQUIN as described earlier but without hierarchical structure. Further, we used a Bayesian method implemented in BAYESCAN version 1.0, which directly estimates the posterior probability of a particular locus being under selection by comparing models with and without a selection term (Foll & Gaggiotti 2008). The analyses were run with the default settings and a locus was considered to be an outlier if \log_{10} (Bayes factor) ≥ 2 .

Neutral population structure across drainages. For the analysis of neutral population structure, five loci showing evidence of selection in the global analysis or within individual drainages were excluded from the data set (two candidate loci: Ssa14, UBA; both diagnostic loci:

Table 2 (a) Outlier loci detected in a global analysis accounting for hierarchical structure in ARLEQUIN. (b) Outlier loci detected within individual drainages using two different approaches implemented in BAYESCAN and ARLEQUIN. In all cases, F_{ST} was higher than average. Loci are grouped as candidate, diagnostic or anonymous based on information from the literature. A locus was considered a significant outlier if $P < 0.01$ (ARLEQUIN) or log Bayes factor > 2 (BAYESCAN)

(a) Global analysis									
ARLEQUIN with hierarchical structure									
F_{SC} (within regions)			F_{CT} (among regions)			F_{ST} (among all populations)			
Candidate	Diagnostic	Anonymous	Candidate	Diagnostic	Anonymous	Candidate	Diagnostic	Anonymous	
Ssa14			Ssa14			Ssa14			
UBA				JMS1M1			JMS1M1		
		Str60							
(b) Drainage-specific analyses									
Drainage	BAYESCAN			ARLEQUIN					
	Candidate	Diagnostic	Anonymous	Candidate	Diagnostic	Anonymous			
Allaine	Ssa14			Ssa14					
Rhine	UBA								
		JMS2M1			JMS2M1				
Ticino							Str60		
Poschiavino	UBA*							Str60	
			Str60					Str60	

*Evidence of selection reduced to substantial ($\log BF > 0.5$) if allele frequencies are adjusted for null alleles.

JMS1M1, JMS2M1; one anonymous locus: Str60; Table 2). Note that in the following we use the term ‘neutral loci’ to refer to the 13 non-outlier loci included in the analysis of neutral population structure. The data were analysed in STRUCTURE v. 2.3.1 (Pritchard *et al.* 2000; Falush *et al.* 2003) assuming an admixture model and correlated allele frequencies with 10 000 burn-in steps and 100 000 MCMC (Markov chain Monte Carlo) steps. This chain length was found to be sufficient to reach convergence. We investigated the likelihood of various numbers of groups, $K = 1-8$, with five independent runs per K . The support for different values of K was assessed from the likelihood distribution (STRUCTURE output) and plots of ΔK (Evanno *et al.* 2005) produced by STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest/).

The spatial structure at the 13 neutral loci was also visually displayed in a principal coordinates plot. We used GENDIST version 3.69 in the PHYLIP package (Felsenstein 2005) to calculate the Cavalli-Sforza chord distance between all pairs of populations. We used a genotype file corrected for null alleles using the INA method (Chapuis & Estoup 2007) implemented in FREENA. Note that evidence of a high-frequency null allele was mainly

apparent at the outlier locus UBA and not at the neutral markers (see Results). The null allele correction was included here only so that all principal coordinates analyses would be strictly comparable. The pairwise genetic distances were subjected to a principal coordinates analysis (PCoA) and the first two axes were plotted in GENALEX (Peakall & Smouse 2006).

An analysis of molecular variance (AMOVA) was used to assess the proportion of genetic variance between major drainages. Again, the two geographically disjunct regions sampled within the Po drainage (Poschiavino and Ticino) were treated as separate groups. The AMOVA was computed in Arlequin v. 3.1, and the significance tests were based on 16 000 permutations.

Neutral population structure within drainages. Putatively non-neutral loci were excluded if they were detected as significant outliers by BAYESCAN or ARLEQUIN within that particular drainage (Table 2b). Additionally, loci with null alleles at estimated frequencies above 10% in a particular drainage were omitted, leaving 15 loci in the Allaine (without UBA, Strutta58, Ssa14), 13 loci in the Poschiavino (without Str60, UBA, SL438, Ssa14, Omy325), 16 loci in the Ticino (without UBA, Str60)

and 16 loci in the Rhine (without UBA, JMS2M1). STRUCTURE analyses were carried out as described above except for the evaluated range of K which was limited to values meaningful in a particular drainage ($K = 1-4$ in Allaine, $K = 1-5$ in Ticino, $K = 1-8$ in Poschiavino and Rhine). Within the Rhine and the Poschiavino, the analyses were repeated with a model that utilized information about sampling locations (Hubisz *et al.* 2009) as no substructure could be detected with our default model.

Within each drainage, the overall F_{ST} value at neutral loci was calculated in Arlequin v. 3.1 and the significance assessed based on 16 000 permutations. We tested for associations between pairwise genetic, geographical and altitudinal distances between populations using Mantel tests and partial Mantel tests in FSTAT v. 2.9.3.2 using 10 000 randomizations. Altitudinal distance was the absolute value of the elevational difference between two sites and geographical distance was measured along the course of rivers in ArcGIS 9.

Spatial and altitudinal patterns at outlier loci. We investigated the role of altitude in shaping the genetic patterns observed at outlier loci by testing for associations between particular outlier alleles and altitude based on (i) the full data set and (ii) only the Rhine and Poschiavino drainages where at least eight sites had been sampled spaced out well across a wide altitudinal range. We considered 23 alleles with overall frequencies above 5% at the five loci detected as high outliers either in the global analysis or within individual drainages (Ssa14, UBA, JMS1M1, JMS2M1, Str60; Table 2). In each population, we recorded the number of individuals in which a particular allele was present (both heterozygotes and homozygotes) and the number in which it was absent. These data were then analysed using a generalized linear model with binomial errors in R v. 2.9.2 (R Development Core Team 2007). The full model included altitude as a covariate and drainage as a factor to allow for hierarchical structure. Terms were dropped from the model if their deletion did not cause a significant increase in deviance, starting with the interaction followed by the covariate. A Bonferroni corrected significance level of 0.0022 was used to account for the 23 performed analyses. The final model was checked for overdispersion of residual deviance and, if necessary, the analysis was repeated using quasibinomial errors and F -tests for model comparison (e.g. Crawley 2002).

Further, we plotted the first two axes from a PCoA to visually compare the genetic structure at individual outlier loci with the neutral pattern at the geographical scale where outlier behaviour had been detected. The analyses were performed as described earlier for the neutral loci.

Patterns of genetic diversity at MHC- and TAP-linked loci. Balancing selection either through rare allele or heterozygote advantage is expected to lead to relatively even frequencies of the different alleles at a particular locus (e.g. Bernatchez & Landry 2003). As a consequence, the observed Hardy–Weinberg heterozygosity (H_e) should be elevated compared to the expectation at mutation drift equilibrium (H_{eq}) for a given number of alleles (Watterson 1978). We investigated this prediction for all MHC-linked loci by comparing H_e to estimates of H_{eq} obtained with the software BOTTLENECK v. 1.2.02 (Cornuet & Luikart 1996). Note that BOTTLENECK allows the estimation of H_{eq} under mutation models other than the infinite alleles model underlying the classical Ewens–Watterson test (Ewens 1972; Watterson 1978). Here, the equilibrium heterozygosity was estimated under a SMM using 1000 iterations. As the distribution of allele frequencies can be strongly affected by demographic events, the results for the MHC-linked loci were compared to 11 neutral microsatellites (omitting all outliers and all MHC-linked loci).

Results

Genetic and genotypic diversity within populations

All loci were polymorphic in the majority of samples with the number of alleles ranging between 3 and 49 (Table S2, Supporting information). Ssa14 and JMS1M1 were monomorphic in ten and six samples from the Rhine and Ticino drainages respectively, and a third locus (Ots515) in sample BU (Allaine). After sequential Bonferroni correction within populations (Rice 1989), four loci showed significant deviations from HWE, but in two cases this was limited to one (Ssa100) or two populations (SL438) only. The MHC I linked locus UBA, on the other hand, showed a highly significant heterozygote deficit in 24 of 30 population samples. Jensen *et al.* (2008) have previously reported null alleles for this locus in *S. trutta*, which also seem to be common in our populations at an estimated average frequency of 0.32 (Table S2, Supporting information). Locus Ssa14 showed highly significant deviations from HWE in all samples where it was polymorphic (frequency of minor allele $\geq 5\%$). The deficit of heterozygous genotypes was particularly pronounced in the four Allaine samples where only 1 of 116 individuals was heterozygous in spite of expected heterozygosities between 0.18 and 0.51. However, in contrast to the UBA locus, not a single missing genotype was observed in the full data set ($N = 853$) which rules out a high-frequency null allele at locus Ssa14.

At the two diagnostic loci, the alleles typical of Atlantic trout dominated in the Rhine but were also found at

high frequencies in all other drainages (>36% in all sites; Table S5, Supporting information). Pronounced allele frequency clines were observed in the Allaine, with the alleles specific to the Mediterranean (ME) lineage decreasing from average frequencies of 56% (JMS1M1)/30% (JMS2M1) in the two downstream sites to 23%/14% in the two upstream sites (Table S5, Supporting information). The two rivers from the Po system (Ticino and Poschiavino) contained alleles typical of marble (MA) trout (population frequencies 0–26%) but also ME-like alleles (0–22%; Table S5, Supporting information). Note, however, that the usefulness of the markers is limited in this region where Adriatic and, in the Poschiavino, also introduced Danubian trout may occur for which the expected allele size range is unknown. ME- and MA-like alleles were also observed in the Rhine but were generally very rare except in population SE.

The two loci physically linked to MHC I, UBA and TAP2a showed significant deviations from linkage equilibrium in all but two populations ($P < 0.05$). The effect was particularly strong in SE, GO, BL, FR, MI and five of the eight Poschiavino samples where the deviation remained significant after sequential Bonferroni correction.

Significant outliers across and within drainages

In the global analysis, the diagnostic locus JMS1M1 and the temperature tolerance-candidate locus Ssa14 were detected as highly significant positive outliers both among populations (F_{ST}) and among regions (F_{CT} ; Table 2a). Within regions, UBA, Str60 and, again, Ssa14 showed elevated levels of genetic differentiation (F_{SC} ; Table 2a).

A separate analysis was performed on a subset of 16 populations from three main river systems (without Allaine) where the MHCII linked marker Ssa60 had been genotyped. The only outlier was the candidate marker Ssa14, which showed evidence of divergent selection both among populations (F_{ST}) and among regions (F_{CT} ; not shown).

The loci showing high outlier behaviour in individual drainages were mostly a subset of those significant in the global analysis: Ssa14 in the Allaine, UBA and, additionally, JMS2M1 in the Rhine and, again, UBA in the Poschiavino (Table 2b). The anonymous locus Str60 was identified as a high outlier in both rivers within the Po drainage, Ticino and Poschiavino.

Locus UBA shows a high frequency of null alleles in our populations which could affect the results of the outlier scans. We used MICROCHECKER (Van Oosterhout *et al.* 2004) to recalculate allele frequencies at UBA assuming the presence of null alleles and repeated the

BAYESCAN analysis for the two drainages where UBA was detected as a significant outlier. In the Rhine, the outlier status of UBA remained unchanged (log Bayesfactor >2), while in the Poschiavino, the signal of divergent selection was reduced from previously decisive to substantial (log Bayesfactor >0.5; see BAYESCAN manual).

Some genetic differences between major drainages have persisted

The most pronounced genetic subdivision coincided with major drainage boundaries, with a split between Allaine, Poschiavino and Rhine, and the Ticino populations intermediate between the latter two. This subdivision was supported both by the principal coordinates analysis (Fig. 2a) and the STRUCTURE analysis (Fig. 3)

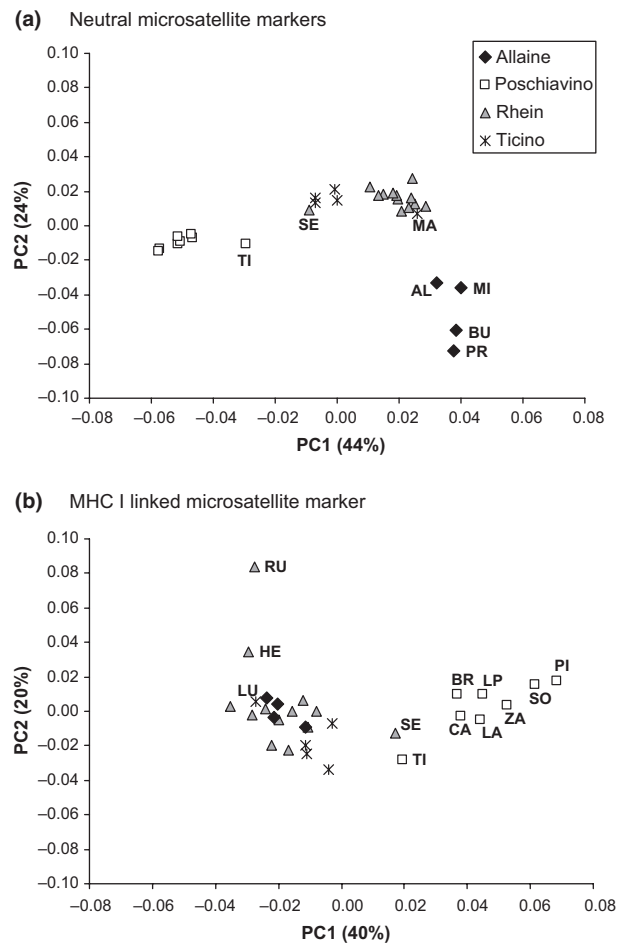


Fig. 2 Plots of the first two axes from a principal coordinates analysis of Cavalli-Sforza chord distances between pairs of populations based on (a) 13 neutral microsatellite markers (non-outliers) and (b) the major histocompatibility complex I linked marker UBA. Different symbols denote populations from different drainages. Labels are given for populations referred to in the text and correspond to Fig. 1.

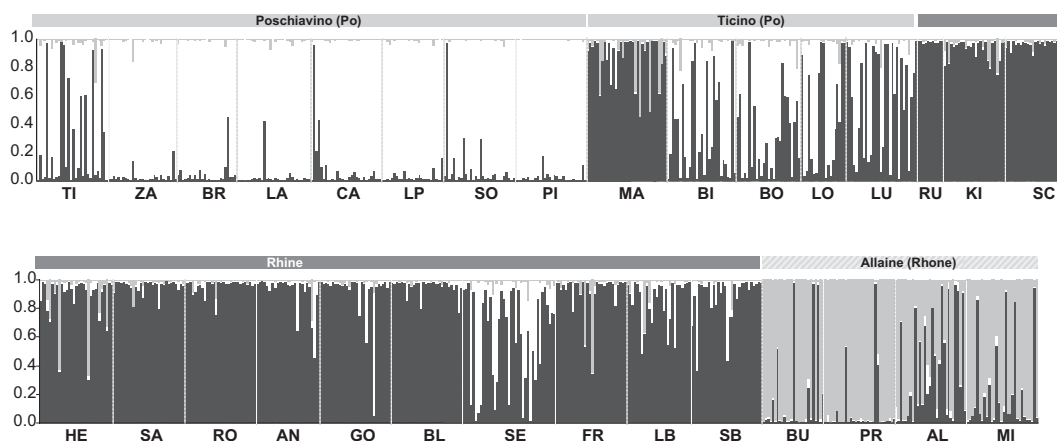


Fig. 3 Results of STRUCTURE analysis based on 13 neutral microsatellite markers and samples from 30 locations in three major European drainage systems ($K = 3$). Population labels correspond to Fig. 1 and Table S1 (Supporting information).

based on neutral loci. Some populations deviated from this general pattern: MA (Ticino) appeared genetically identical to most Rhine populations, while SE (Rhine) resembled the Ticino samples (PCoA: Fig. 2a; STRUCTURE: Fig. 3).

Across all populations, the average F_{ST} estimated from the 13 neutral loci was 0.056 ($P < 0.001$). When populations were grouped based on the four main river systems (Allaine, Rhine, Poschiavino and Ticino), 3.5% of the variation was observed among groups compared with 2.1% between populations within groups (AMOVA).

Geography explains much of the genetic structure within drainages

Significant neutral structure was also observed within each of the four drainages with $F_{ST} = 0.007$ in the Poschiavino, $F_{ST} = 0.025$ in the Rhine, $F_{ST} = 0.032$ in the Ticino and $F_{ST} = 0.034$ in the Allaine ($P < 0.001$ in all cases; AMOVA). A pattern of isolation by distance was detected among the Rhine ($R^2 = 19.9\%$, $P < 0.01$; Mantel test) and the Ticino populations ($R^2 = 84.1\%$, $P < 0.01$) but not in the Allaine ($P = 0.85$) and the Poschiavino ($P = 0.54$). Within the Rhine, but not in the other three rivers, we observed a significant pattern of isolation by altitude after accounting for the effect of geographical distance ($R^2 = 32.1\%$, $P(\text{altitude}) < 0.01$; partial Mantel test). Note that in all other rivers only a single gradient had been sampled (Fig. 1) and, as a consequence, altitudinal and geographical distance were highly correlated. A table of pairwise F_{ST} values between all populations is provided in the electronic appendix (Table S3, Supporting information).

In the Ticino, the STRUCTURE analysis revealed a very clear geographical and altitudinal signal with a split between the high (LU), middle (BO/LO/BI) and low

elevation sites (MA; Fig. S1a, Supporting information). However, each population also contained a few individuals with genotypes more typical of a different region many of which looked like first-generation immigrants. In the Poschiavino and Rhine, substructure was generally weak and could only be resolved when information about the sampling locations was included. The clearest split in the Poschiavino was observed between the lowest site in Italy (TI) and the seven Swiss locations (Fig. S1b, Supporting information). In the Rhine, nearby populations tended to be similar, while the two isolated high elevation sites (SE, and to a lesser extent AN) were quite distinct (Fig. S1c, Supporting information). Finally, within the Allaine, STRUCTURE detected strong support for two genetic clusters (Fig. S1d, Supporting information). One cluster dominated at the two upstream sites (AL, MI), while both were observed at the two downstream sites (BU, PR). In fact, the presence of two genetic clusters was still strongly supported in an analysis based only on BU/PR, in which case the split was clearly not associated with geography (i.e. sampling location).

Altitudinal and geographical patterns at outlier loci

Significant associations between allele frequency and altitude were detected for three of 23 investigated alleles at outlier loci. Allele UBA-301 showed a concordant and highly significant positive correlation between allele frequency and altitude in both the Poschiavino and the Rhine (glm, $P < 0.001$; note that this allele is extremely rare or absent in the other two drainages; Fig. 4a), while a negative correlation was observed at a second allele, UBA-307 (glm, $P < 0.001$). This latter allele showed a concordant pattern in the Allaine but not in the Ticino (Fig. 4b) and, in the full analysis, the

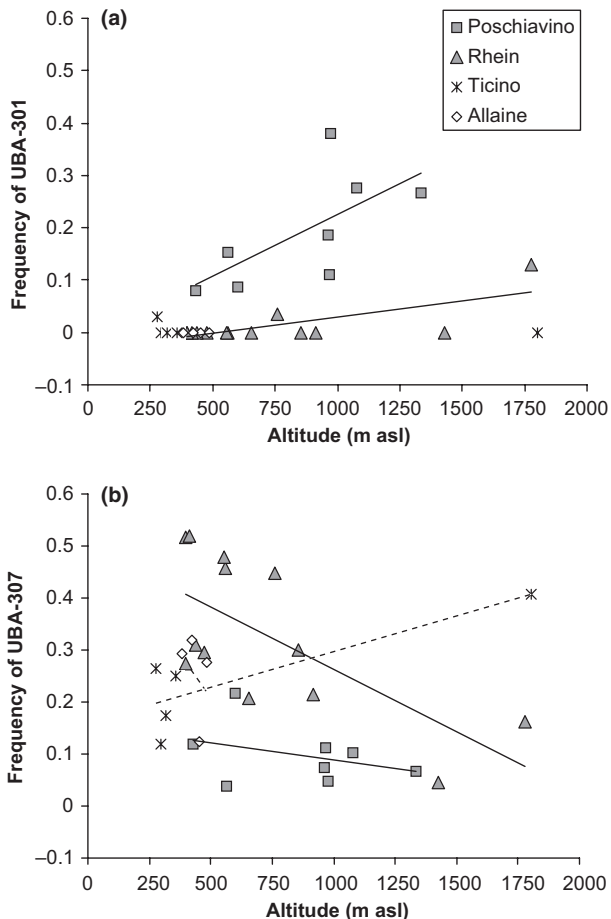


Fig. 4 Association between altitude (m above sea level) and the frequency of particular alleles at the major histocompatibility complex I linked locus UBA in the four drainages. Solid trend lines are shown for Rhine and Poschiavino and dashed for Allaine and Ticino. Note the clumped altitudinal distribution of the four Allaine samples and four of the five Ticino samples. (a) allele 301, (b) allele 307.

effect of altitude was not significant (glm, $P = 0.18$). Finally, an allele at the diagnostic locus JMS1M1 (allele 146 specific to Mediterranean trout) showed a significant interaction between drainage and altitude in the analysis based on all samples (glm, $P < 0.001$). The allele decreased in frequency from downstream towards upstream along the course of the Allaine, while a (much weaker) trend in the opposite direction was observed in the Poschiavino (plot not shown; the allele was rare in the other drainages).

The overall geographical pattern at the two global outliers, Ssa14 and JMS1M1, was consistent with that observed at non-outlier loci, but with, on average, larger differences between and, in some cases, also within drainages. The elevated differentiation detected at JMS2M1 and Str60 within individual drainages

seemed to be driven largely by one or two populations which, at the respective locus, were genetically very different from the other populations.

The spatial pattern at the repeated outlier locus UBA, however, differed markedly from that at neutral loci. Most prominently, the genetic distances between the Poschiavino populations were elevated at UBA (Fig. 2b). Further, we observed at least a weak association between the altitude of a site and its position along PC1 (Fig. 2b), with the lowest site (TI) to the left and the two highest sites (SO/PI) to the right of the Poschiavino cluster. Most of the remaining populations showed little divergence at UBA and, contrary to the neutral pattern, we observed no clear subdivision between Rhine, Ticino and Allaine (Fig. 2b). The only exceptions were three Rhine populations with distinct UBA genotypes (SE, RU and HE; Fig. 2b).

Patterns of intra-population diversity at MHC- and TAP-linked vs. neutral markers

Across neutral loci, one population (RU) showed a significant heterozygosity excess, and five populations (SO, MA, BI, BO, LU) a significant heterozygosity deficit (Wilcoxon test; Table 3). Across populations, one sample t -tests performed for each locus separately found that the mean difference between H_e and H_{eq} was significantly negative for four neutral and two MHC-linked markers (TAP2b; Ssa60), and significantly positive for two neutral and one MHC-linked marker (TAP2a; Table S4, Supporting information). The observed Hardy-Weinberg heterozygosity was most strongly elevated at TAP2a with a 95% confidence interval for the mean $H_e - H_{eq}$ of 0.073–0.102. The next highest value was observed at the neutral locus SL438 (95% c.i. 0.029–0.064).

Discussion

Phylogeographic structure has not been completely erased by stocking

In spite of large-scale anthropogenic movements of trouts between different drainages, some genetic differences between these major river systems have persisted. The populations from the Allaine, the Poschiavino and the Rhine form distinct genetic clusters, while the Ticino samples are intermediate between the latter two. Much of this divergence between drainages can probably be ascribed to differences in the contribution of distinct evolutionary trout lineages to the local gene pools. Microsatellite alleles diagnostic of Mediterranean trout are present in the Allaine (see also next section), and genetic variants typical of Adriatic and marble trout are

Table 3 Comparison of observed Hardy–Weinberg heterozygosity (H_e) and the expectation at mutation-drift equilibrium (H_{eq}) under a stepwise mutation model. For the neutral markers (non-outliers), the number of loci showing a heterozygosity excess ($H_e > H_{eq}$) or deficit ($H_e < H_{eq}$) is indicated (for full results see Table S4)

Drainage	Neutral				MHC linked			
	Site	Excess	Deficit	<i>P</i> -value	UBA	TAP2a	TAP2b	Ssa60
Poschiavino	TI	7	4	0.577	0.029	0.075	−0.014	−0.004
Poschiavino	ZA	3	8	0.083	−0.014	0.052	−0.047	0.003
Poschiavino	BR	4	7	0.175	−0.008	0.083	−0.008	NA
Poschiavino	LA	5	6	0.240	0.009	0.065	0.020	−0.007
Poschiavino	CA	8	3	0.831	0.016	0.079	0.027	−0.018
Poschiavino	LP	3	8	0.206	−0.017	0.099	0.025	NA
Poschiavino	SO	4	7	0.021	−0.026	0.202	0.032	NA
Poschiavino	PI	6	5	0.638	−0.015	0.039	0.014	0.001
Ticino	MA	2	9	0.002	−0.009	0.093	−0.053	−0.015
Ticino	BI	1	10	0.001	−0.135	0.103	0.017	NA
Ticino	BO	2	9	0.007	−0.073	0.100	0.009	NA
Ticino	LO	4	7	0.240	0.039	0.086	−0.022	−0.018
Ticino	LU	3	8	0.016	−0.058	0.100	0.003	NA
Rhine	RU	10	1	0.002	0.061	0.070	−0.035	−0.048
Rhine	KI	4	7	0.206	−0.016	0.053	−0.043	NA
Rhine	SC	6	5	0.898	0.063	0.093	−0.020	0.004
Rhine	HE	7	4	0.966	−0.064	0.120	−0.050	0.009
Rhine	SA	5	6	0.278	0.008	0.106	−0.080	−0.025
Rhine	RO	5	6	0.240	−0.079	0.110	−0.068	−0.003
Rhine	AN	4	7	0.123	−0.005	0.136	0.024	NA
Rhine	GO	7	4	0.765	0.014	0.096	−0.025	−0.023
Rhine	BL	5	6	0.966	0.005	0.100	−0.037	−0.017
Rhine	SE	4	7	0.067	−0.047	0.097	−0.040	−0.011
Rhine	FR	4	7	0.365	−0.019	0.106	−0.141	−0.008
Rhine	LB	3	8	0.067	−0.007	0.128	−0.028	NA
Rhine	SB	3	8	0.175	−0.063	0.041	−0.102	NA
Allaine	BU	6	4	0.625	0.011	0.025	−0.046	NA
Allaine	PR	5	6	0.278	0.009	−0.029	0.000	NA
Allaine	AL	6	5	0.638	0.009	0.100	0.045	NA
Allaine	MI	6	5	0.765	0.018	0.098	0.017	NA

MHC, major histocompatibility complex.

P-value = two-sided *P*-value for heterozygosity excess or deficit from a Wilcoxon test based on all neutral loci. For the MHC-linked loci, $H_e - H_{eq}$ is given. Values significantly different from zero ($P < 0.05$) are shown in bold.

observed in the Poschiavino and, to a lesser extent, the Ticino (this study, Largiadèr & Scholl 1995).

However, it is also evident that nonindigenous genotypes have persisted in all Swiss drainages since the ban of between-drainage translocations in 1991 (Bundesgesetz über die Fischerei). Historically, large numbers of Atlantic trout were introduced throughout Switzerland and, on a more local scale, the Poschiavino was also stocked with fish of Danubian origin (Largiadèr & Scholl 1995). Such mostly asymmetric gene flow from the Rhine (i.e. Atlantic trout) into the different Mediterranean drainages (Allaine, Ticino, Poschiavino) is consistent with the molecular data which show that genetic variants typical of Atlantic trout occur in all Swiss drainages (Largiadèr & Scholl 1995, 1996; this study and our own unpublished data using other markers). An exception to

this general pattern is provided by our population SE from the Rhine which appears genetically similar to the Ticino populations. Elevated frequencies (10%; Table S5, Supporting information) of non-Atlantic alleles at one of the diagnostic markers and the presence of mitochondrial haplotypes typical of Danubian and marble trouts (our own unpublished data) suggest that, in the past, the site was stocked with fish from more than one source outside the Rhine system.

Genetic differentiation within drainages at neutral markers

Within drainages, populations were often significantly differentiated and geographical proximity was generally a good predictor of genetic similarity at neutral

markers. These results suggest that, also within rivers, the expected natural population structure may not have been erased completely by the effects of artificial stocking. An exception may be the Poschiavino river where genetic differentiation is generally low in spite of significant physical barriers at least between some of our sampling locations. This is surprising because besides the Allaine, the Poschiavino harbours the largest phenotypic variation between and within populations.

Overall, the patterns of genetic diversity at neutral markers suggest that the rates of gene flow in trout are probably high but still not high enough to completely homogenize allele frequencies along the river continuum.

Indications for sympatric coexistence of two trout species in the Allaine

The neutral microsatellite data suggest particularly intriguing genetic patterns in the river Allaine. This river is part of the Doubs subdrainage within the Rhone system and, historically, it would have contained trout of the Mediterranean lineage, which are considered a distinct species in the recent taxonomic reassessment of the European trout complex (*S. rhodanensis*; Kottelat & Freyhof 2007). *S. rhodanensis* have a very distinct phenotype and are also known as zebra trout because of four characteristic broad black bars on the flanks of adult fish (e.g. Mezzera *et al.* 1997). Like all Mediterranean drainages in Switzerland, the Doubs region was heavily stocked with hatchery fish of Atlantic origin in the past (e.g. Largiadèr & Scholl 1996).

Our molecular analyses, indeed, detected two genetic clusters within the Allaine whose relative frequency changed between the upstream and downstream section of the river. Within sites, we observed fairly nonadmixed genotypes of both groups but also some admixed individuals (Figs 3 and S1d, Supporting information). Together with observed allele frequency clines at both diagnostic loci (Table S5, Supporting information), this result suggests a decrease in the relative frequency of Mediterranean-like alleles and genotypes from the downstream towards the upstream populations. This interpretation is also in line with the phenotypic diversity observed in our samples. A subset of 43 individuals from all four sites could be grouped as Atlantic-like (79%) or Mediterranean-like (21%) based on their phenotypes, with all Mediterranean-like individuals restricted to the lower reaches of the Allaine (sites BU and PR). Individuals with a Mediterranean-like phenotype tended to be assigned almost exclusively to one of the clusters identified by STRUCTURE, while the admixture proportions in individuals with an Atlantic-like phenotype were more variable.

A very striking pattern was observed at the outlier locus Ssa14, which showed not only elevated levels of genetic differentiation but also exceptionally strong deviations from HWE. The almost complete absence of heterozygous genotypes is almost certainly not caused by a high-frequency null allele as not a single missing genotype (i.e. null homozygote) was observed. Possible explanations for the observed pattern include perfect assortative mating between two lineages fixed for alternative alleles and/or very strong selection against heterozygotes. Lineage-specific differences at this locus are indeed suggested by the observation that Ssa14 is often fixed or nearly fixed in Atlantic populations (Table S5, Supporting information; see also supplementary data of Hansen *et al.* 2010), and polymorphic in populations with a more diverse ancestry (e.g. Allaine, Poschiavino; Table S5, Supporting information). Within the Allaine, the frequency of allele 140 that is otherwise typical of Atlantic (i.e. Rhine) populations declines from 58% in the two upstream sites to 18% in the two downstream sites, concordant with the decrease of AT-like genotypes.

Together, our results indicate partial reproductive isolation between trout of Atlantic and Mediterranean origin, as suggested also by previous studies (Largiadèr & Scholl 1996; Mezzera & Largiadèr 2001). The extent of reproductive barriers between the two forms and the traits that confer such barriers will need to be further examined. Similar to our results, other authors have found large between-site variation in the proportion of Mediterranean-like genotypes in populations heavily stocked with fish of Atlantic origin and have hypothesized that these patterns could be caused by differences in local environments (Largiadèr & Scholl 1996; Poteaux *et al.* 1998; Berrebi *et al.* 2000). However, the ecological factors that permit or prevent coexistence will need to be further examined.

Contrary to the Allaine, we do not observe similarly distinct groups of genotypes in the rivers draining into the Po where, historically, at least two trout forms would have co-occurred. However, we are cautious not to exclude the possibility that more subtle substructure may exist in these drainages. For further investigation, we are currently analysing the same samples at a larger number of random genomic loci.

Potentially adaptive population differentiation

The different outlier scans detected unusual levels of population differentiation at several loci from all three categories (candidate, diagnostic and anonymous). Between-drainage differentiation was elevated at two loci, for which we know (in the case of JMS1M1; Estoup *et al.* 2000) or suspect (Ssa14) that different alleles were

originally associated with certain Pleistocene trout lineages. Such species-specificity of particular genetic variants may be a direct consequence of different selection pressures experienced by different evolutionary lineages (Minder & Widmer 2008). The fact that above-average levels of differentiation are maintained could suggest that these selection pressures are still operating between drainages today.

It was not consistently the same loci which showed outlier behaviour in the different drainages. The power to detect outliers almost certainly varies between drainages but the loci showing the strongest evidence of outlier behaviour were clearly not always the same. Such a pattern could result from between-drainage differences in the amount and type of (adaptive) genetic variation available to selection. This could apply for instance to the divergence at *Ssa14*, a locus lacking variation in the Rhine and Ticino. Alternatively, or additionally, the relative intensity of divergent selection on particular traits may differ between regions.

At this point, it is unclear which environmental differences could underlie population divergence at the detected outlier loci. Overall, altitudinal differences appear to play a limited role with the possible exception of the MHC I linked marker UBA.

Conclusions about genome-wide patterns of adaptive divergence in these populations are clearly not possible based on the limited number of markers used in this study. However, this does not invalidate our conclusion that individual loci show levels of genetic differentiation considerably higher than the neutral average which may suggest the action of divergent selection. A methodological concern in smaller data sets may be that this average neutral differentiation is less accurately estimated and more strongly influenced by individual loci. Against this background, and to reduce the number of false positives, we used stringent significance thresholds in all analyses. Further, all outlier scans in ARLEQUIN were carried out in an iterative manner where average F_{ST} was recalculated after excluding outlier loci until no further outlier loci were found.

Evidence of selection acting on immune loci

The patterns of diversity at locus UBA suggest that, sometimes, different MHC I variants (or sets of variants) may be favoured in different populations as previously observed in a range of other species (see e.g. Muirhead 2001). In the Poschiavino, UBA showed consistently larger divergence between populations compared to neutral loci. In the Rhine, on the other hand, we observed three populations (RU, HE and SE) that were clearly distinct while all others grouped together and, surprisingly, even overlapped with the Allaine

and Ticino samples. This suggests that, while some populations have very distinct UBA allele frequencies, others are more similar than predicted based on the neutral genetic structure. Consistent with this result, other studies have found that the type (i.e. balancing or directional) and intensity of selection acting on the MHC can be highly variable both in space and/or time (reviewed in Bernatchez & Landry 2003; Piertney & Oliver 2006). The correlation of some UBA alleles with altitude in several drainages may further point to replicated changes in the composition of local pathogen/parasite communities along these altitudinal gradients.

Contrary to UBA, the pattern of genetic diversity at the second locus within the MHC region IA, TAP2a, appeared more consistent with the action of balancing selection. The overall level of genetic differentiation observed at TAP2a was relatively low ($F_{ST} = 0.039$), but none of the outlier scans found this to be inconsistent with the genomic average (average neutral $F_{ST} = 0.059$). What was remarkable, however, was the extremely even frequencies of the different TAP2a alleles, with exactly the same four alleles present in 29 of the 30 populations. In line with this result, the relative frequencies of the four alleles were remarkably even also within populations leading to systematically higher observed Hardy–Weinberg heterozygosities (H_e) than expected at mutation-drift equilibrium (H_{eq}) as predicted under balancing selection (Maruyama & Nei 1981).

Clearly, the evidence of balancing selection suggested by this analysis is tentative. The main concern is that estimates of equilibrium heterozygosity depend heavily on the mutation model, and that each locus probably deviates to a different extent from the assumed SMM. If, for example, TAP2a followed a strict SMM, while the neutral loci did not, the heterozygosity excess at TAP2a would no longer be outside the range observed at neutral loci. Such a scenario seems rather unlikely, however, given that TAP2a is an interrupted microsatellite and is, therefore, expected to deviate more from an SMM (Cornuet & Luikart 1996) than the neutral microsatellites which are mostly pure dinucleotide repeats (9 of 11 loci; Table S2, Supporting information).

TAP2a was also investigated in brown trout from different Danish river systems and exhibited similar or lower levels of genetic differentiation between populations than neutral markers (Jensen *et al.* 2008; note that our locus TAP2a corresponds to TAP2B of Jensen *et al.*). The functional role of TAP2 polymorphism is poorly understood even in humans (McCluskey *et al.* 2004) but, in some species, strong linkage disequilibria (LD) between MHC I and TAP2 alleles have been observed

(Ohta *et al.* 2003), which may suggest a selective advantage of particular allelic combinations. Despite some LD between UBA and TAP2a in our populations consistent with their close physical proximity (28 kb in the related Atlantic salmon; Lukacs *et al.* 2010), the two markers show clearly distinct patterns of genetic divergence which suggests recombination rates high enough for the two loci to respond, to a large extent, independently to selection.

It is clear that differences in locus-specific mutation rates and recombination can lead to an imperfect association between microsatellite alleles and particular functional variants. Still, our results demonstrate that the analysis of microsatellite markers within the MHC region may be a rapid and cost-effective way to assess the promise of more detailed and time-consuming molecular analyses of functional polymorphisms, which would certainly be warranted in this system.

Conclusions

Our survey of different categories of microsatellite markers in Swiss trout populations revealed that considerable regional genetic differences have persisted in spite of massive artificial stocking, which, in the past, was performed without considering the geographical origin of the introduced fish. Further, we find evidence of adaptive genetic differences between populations which are often drainage specific. Our results tentatively implicate loci historically associated with particular evolutionary trout lineages as well as genes of the immune system in local adaptation.

Acknowledgements

We thank everyone who helped with fieldwork, in particular B. Almasi, C. Baumgartner, L. Costa, J.-M. Fierz, B. Germann, M. Grünenfelder, S. Haertel, J. Koegler, A. Lièvre, A. Peter, B. Polli, C. Rau, E. Schager, J. Schuler, D. Senn, B. Tschirren, H. Walther, P. Warnier, and A. Westram. We are grateful to the fisheries authorities of Bern, Graubünden, Jura, Luzern, St. Gallen, Solothurn, Thurgau, Ticino, Uri and the province of Sondrio, Italy, for their support of this project. We further thank E. Bezault, L. Excoffier, M. Foll, S. Haertel, J. Hines, J. Jokela, K. Klappert, K. Kopp, S. Mwaiko, G. Périat and A. Peter for help with molecular work and/or helpful discussions. We are grateful to the subject editor and three anonymous reviewers for their detailed and constructive comments on an earlier version of this manuscript. This study was part of BioChange financed through ETH-CCES.

References

Abele R, Tampé R (2006) Modulation of the antigen transport machinery TAP by friends and enemies. *FEBS Letters*, **580**, 1156–1163.

- Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology & Evolution*, **21**, 433–438.
- Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss*). *Molecular Ecology*, **15**, 923–937.
- Angilletta MJ (2009) *Thermal Adaptation – A Theoretical and Empirical Synthesis*. Oxford University Press, Oxford.
- Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38–44.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.
- Beaumont M, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.
- Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351–379.
- Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, **16**, 363–377.
- Bernatchez L, Renaut S, Whiteley AR *et al.* (2010) On the origin of species: insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1783–1800.
- Berrebi P, Poteaux C, Fissier M, Cattaneo-Berrebi G (2000) Stocking impact and allozyme diversity in brown trout from Mediterranean southern France. *Journal of Fish Biology*, **56**, 949–960.
- Carlson SM, Seamons TR (2008) A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. *Evolutionary Applications*, **1**, 222–238.
- Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.
- Colosimo PF, Hosemann KE, Balabhadra S *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, **307**, 1928–1933.
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
- Crawley MJ (2002) *Statistical Computing – An Introduction to Data Analysis Using S-Plus*, 1st edn. John Wiley & Sons Ltd, Chichester.
- Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L (2007) Clinal variation in MHC diversity with temperature: evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. *Evolution*, **61**, 2154–2164.
- 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.
- Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.
- Estoup A, Largiadèr CR, Cornuet JM *et al.* (2000) Juxtaposed microsatellite systems as diagnostic markers for admixture:

- an empirical evaluation with brown trout (*Salmo trutta*) as model organism. *Molecular Ecology*, **9**, 1873–1886.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Ewens WJ (1972) Sampling theory of selectively neutral alleles. *Theoretical Population Biology*, **3**, 87–112.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
- Felsenstein J (2005) *PHYLIP (Phylogeny Inference Package) Version 3.69*. Department of Genome Sciences, University of Washington, Seattle, WA.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.
- Garrigan D, Hedrick PW (2003) Detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–1722.
- Gharbi K, Gautier A, Danzmann RG *et al.* (2006) A linkage map for brown trout (*Salmo trutta*): chromosome homeologies and comparative genome organization with other salmonid fish. *Genetics*, **172**, 2405–2419.
- Gharbi K, Glover K, Stone L *et al.* (2009) Genetic dissection of MHC-associated susceptibility to *Lepeophtheirus salmonis* in Atlantic salmon. *BMC Genetics*, **10**, 20.
- Giuffra E, Bernatchez L, Guyomard R (1994) Mitochondrial control region and protein coding genes sequence variation among phenotypic forms of brown trout *Salmo trutta* from northern Italy. *Molecular Ecology*, **3**, 161–171.
- Giuffra E, Guyomard R, Forneris G (1996) Phylogenetic relationships and introgression patterns between incipient parapatric species of Italian brown trout (*Salmo trutta* L. complex). *Molecular Ecology*, **5**, 207–220.
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).
- Grimholt U (1997) Transport-associated proteins in Atlantic salmon (*Salmo salar*). *Immunogenetics*, **46**, 213–221.
- Grimholt U, Drabløs F, Jørgensen S, Høyheim B, Stet R (2002) The major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics*, **54**, 570–581.
- Hansen MM, Skaala O, Jensen LF, Bekkevold D, Mensberg K-LD (2007) Gene flow, effective population size and selection at major histocompatibility complex genes: brown trout in the Hardanger Fjord, Norway. *Molecular Ecology*, **16**, 1413–1425.
- Hansen MM, Meier K, Mensberg K-LD (2010) Identifying footprints of selection in stocked brown trout populations: a spatio-temporal approach. *Molecular Ecology*, **19**, 1787–1800.
- Hari RE, Livingstone DM, Siber R, Burkhardt-Holm P, Güttinger H (2006) Consequences of climatic change for water temperature and brown trout populations in Alpine rivers and streams. *Global Change Biology*, **12**, 10–26.
- Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, **173**, 579–588.
- Hodgson S, Quinn TP (2002) The timing of adult sockeye salmon migration into fresh water: adaptations by populations to prevailing thermal regimes. *Canadian Journal of Zoology*, **80**, 542–555.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.
- Jackson TR, Ferguson MM, Danzmann RG *et al.* (1998) Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. *Heredity*, **80**, 143–151.
- Jensen LF, Hansen MM, Mensberg KLD, Loeschcke V (2008) Spatially and temporally fluctuating selection at non-MHC immune genes: evidence from TAP polymorphism in populations of brown trout (*Salmo trutta*, L.). *Heredity*, **100**, 79–91.
- Kingsolver JG (2009) The well-temperated biologist. *American Naturalist*, **174**, 755–768.
- Klein J (1986) *Natural History of the Major Histocompatibility Complex*. Wiley, New York.
- Kottelat M, Freyhof J (2007) *Handbook of European Freshwater Fishes*. Kottelat and Freyhof, Cornol, Switzerland and Berlin, Germany.
- Landry C, Bernatchez L (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **10**, 2525–2539.
- Largiadèr CR, Scholl A (1995) Effects of stocking on the genetic diversity of brown trout populations of the Adriatic and Danubian drainages in Switzerland. *Journal of Fish Biology*, **47**(Suppl. A), 209–225.
- Largiadèr CR, Scholl A (1996) Genetic introgression between native and introduced brown trout *Salmo trutta* L. populations in the Rhone River Basin. *Molecular Ecology*, **5**, 417–426.
- Lukacs MF, Harstad H, Grimholt U *et al.* (2007) Genomic organization of duplicated major histocompatibility complex class I region in Atlantic salmon (*Salmo salar*). *BMC Genomics*, **8**, 251.
- Lukacs M, Harstad H, Bakke H *et al.* (2010) Comprehensive analysis of MHC class I genes from the U-, S-, and Z-lineages in Atlantic salmon. *BMC Genomics*, **11**, 154.
- Marcogliese DJ (2008) The impact of climate change on the parasites and infectious diseases of aquatic animals. *Scientific and Technical Review of the World Organisation for Animal Health OIE*, **27**, 467–484.
- Maruyama T, Nei M (1981) Genetic variability maintained by mutation and overdominant selection in finite populations. *Genetics*, **98**, 441–459.
- McCluskey J, Rossjohn J, Purcell AW (2004) TAP genes and immunity. *Current Opinion in Immunology*, **16**, 651–659.
- Meraner A, Baric S, Pelster B, Dalla Via J (2007) Trout (*Salmo trutta*) mitochondrial DNA polymorphism in the centre of the marble trout distribution area. *Hydrobiologia*, **579**, 337–349.
- Mezzera M, Largiadèr CR (2001) Comparative analysis of introgression at three marker classes: a case study in a

- stocked population of brown trout. *Journal of Fish Biology*, **59**, 289–305.
- Mezzera M, Largiadèr CR, Scholl A (1997) Discrimination of native and introduced brown trout in the River Doubs (Rhône drainage) by number and shape of parr marks. *Journal of Fish Biology*, **50**, 672–677.
- Miller K, Winton J, Schulze A, Purcell M, Ming T (2004) Major histocompatibility complex loci are associated with susceptibility of Atlantic salmon to infectious hematopoietic necrosis virus. *Environmental Biology of Fishes*, **69**, 307–316.
- Minder AM, Widmer A (2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. *Molecular Ecology*, **17**, 1552–1563.
- Muirhead CA (2001) Consequences of population structure on genes under balancing selection. *Evolution*, **55**, 1532–1541.
- Ohlberger J, Staaks G, Petzoldt T, Mehner T, Hoelker F (2008) Physiological specialization by thermal adaptation drives ecological divergence in a sympatric fish species pair. *Evolutionary Ecology Research*, **10**, 1173–1185.
- Ohta Y, Powis SJ, Lohr RL *et al.* (2003) Two highly divergent ancient allelic lineages of the transporter associated with antigen processing *TAP* gene in *Xenopus*: further evidence for co-evolution among MHC class I region genes. *European Journal of Immunology*, **33**, 3017–3027.
- O'Malley KG, Sakamoto T, Danzmann RG, Ferguson MM (2003) Quantitative trait loci for spawning date and body weight in rainbow trout: testing for conserved effects across ancestrally duplicated chromosomes. *Journal of Heredity*, **94**, 273–284.
- Palti Y, Nichols KM, Waller KI, Parsons JE, Thorgaard GH (2001) Association between DNA polymorphisms tightly linked to MHC class II genes and IHN virus resistance in backcrosses of rainbow and cutthroat trout. *Aquaculture*, **194**, 283–289.
- Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.
- Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity*, **96**, 7–21.
- Poteaux C, Beaudou D, Berrebi P (1998) Temporal variations of genetic introgression in stocked brown trout populations. *Journal of Fish Biology*, **53**, 701–713.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- R Development Core Team (2007) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- Seehausen O, Terai Y, Magalhaes IS *et al.* (2008) Speciation through sensory drive in cichlid fish. *Nature*, **455**, 620–626.
- Somorjai IML, Danzmann RG, Ferguson MM (2003) Distribution of temperature tolerance quantitative trait loci in Arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*). *Genetics*, **165**, 1443–1456.
- Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 979–988.
- Steiner CC, Rompler H, Boettger LM, Schoneberg T, Hoekstra HE (2009) The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Molecular Biology and Evolution*, **26**, 35–45.
- Stet RJM, Kruiswijk CP, Dixon B (2003) Major histocompatibility lineages and immune gene function in teleost fishes: the road not taken. *Critical Reviews in Immunology*, **23**, 441–471.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.
- Tonteri A, Vasemägi A, Lumme J, Primmer CR (2010) Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. *Molecular Ecology*, **19**, 1273–1282.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Vasemägi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic Salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, **22**, 1067–1076.
- Wahli T, Bernet D, Steiner P, Schmidt-Posthaus H (2007) Geographic distribution of *Tetracapsuloides bryosalmonae* infected fish in Swiss rivers: an update. *Aquatic Sciences – Research Across Boundaries*, **69**, 3–10.
- Watterson GA (1978) The homozygosity test of neutrality. *Genetics*, **88**, 405–417.
- Wegner KM (2008) Historical and contemporary selection of teleost MHC genes: did we leave the past behind? *Journal of Fish Biology*, **73**, 2110–2132.

I.K. is a molecular population geneticist with particular interests in adaptation, speciation and conservation genetics. A.T. is a research assistant working on a wide range of fish taxa. O.S. is interested in processes and mechanisms implicated in the origins, maintenance and loss of species diversity and adaptive diversity.

Data accessibility

Data deposited at Dryad: doi:10.5061/dryad.8606.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig S1 Genetic structure within individual drainages. Unless otherwise indicated, STRUCTURE analyses were based on neutral microsatellite loci using the default settings outlined in the manuscript. All location names correspond to Fig. 1 and Table S1. The best supported number of groups (*K*) was inferred from plots of the likelihood distribution or, if no clear mode was detected, based on ΔK proposed by Evanno *et al.* (2005).

Table S1 Sampling locations in three major European drainages.

Table S2 Microsatellite markers used in this study.

Table S3 Pairwise F_{ST} values calculated based on ten neutral loci.

Table S4 Difference between the observed Hardy–Weinberg heterozygosity (H_e) and the expectation at mutation-drift equilibrium (H_{eq}) under a stepwise mutation model.

Table S5 Frequency of Atlantic-like (AT), marble trout-like (MA) and Mediterranean-like (ME) alleles (see footnote) at the two diagnostic microsatellite markers JMS1M1 and JMS2M1 for all populations.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.