Contrasting Responses to Selection in Class I and Class IIα

Major Histocompatibility Linked Markers in Salmon

S. CONSUEGRA1,2, E. DE EYTO3, P. MCGINNITY3,5, R. J. M. STET4† and W. C. JORDAN1

1Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK
2Aberystwyth University, Institute of Biological, Environmental & Rural Sciences, Aberystwyth SY23 3DA, UK
3Aquaculture and Catchment Management Services, Marine Institute, Newport, Co Mayo, Republic of Ireland
4University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, Scotland
5Present address: Cork Enterprise Centre, National University of Ireland Cork, Ireland

Correspondence: Sofia Consuegra. Tel: +44 (0)1970 628775, Fax +44 (0)1970 622350, Email: skc@aber.ac.uk

Running Head: Selection in salmon MH markers.

Keywords: MHC, microsatellites, natural selection, temporal variation, Atlantic salmon, Salmo salar.

Word count (main text): 5,487
†This paper is dedicated to the memory of Professor René Stet. Deceased, 12th September 2007.
Abstract

Comparison of levels and patterns of genetic variation in natural populations either across loci or against neutral expectation can yield insight into locus-specific differences in the strength and direction of evolutionary forces. We used both approaches to test hypotheses on patterns of selection on major histocompatibility (MH)-linked markers. We performed temporal analyses of class I and class IIα MH-linked markers and eight microsatellite loci in two Atlantic salmon populations in Ireland on two temporal scales: over six decades and nine years in the rivers Burrishoole and Delphi, respectively. We also compared contemporary Burrishoole and Delphi samples with nearby populations for the same loci. On comparing patterns of temporal and spatial differentiation among classes of loci, class IIα MH-linked marker was consistently identified as outlier compared to patterns at other microsatellite loci or neutral expectation. We found higher levels of temporal and spatial heterogeneity in heterozygosity (but not in allelic richness) for the class IIα MH-linked marker compared to microsatellites. Tests on both within and among population differentiation are consistent with directional selection acting on the class IIα-linked marker in both temporal and spatial comparisons but only in temporal comparisons for the class I-linked marker. Our results indicate a complex pattern of selection on MH-linked markers in natural populations of Atlantic salmon. These findings highlight the importance of considering selection on MH-linked
markers when using these markers for management and conservation purposes.
Introduction

Identifying loci under selection, and therefore potentially involved in adaptation, is a major aim in evolutionary ecology (Nielsen, 2005). Often, approaches to identifying selected loci entail comparing levels and patterns of variation in natural populations with an expectation under a neutral model of evolution (Beaumont and Balding, 2004; Ford, 2002; Goldringer and Bataillon, 2004; Kimura, 1985; Vitalis et al, 2001). Alternatively, contrasting variation among loci (or classes of loci) from the same individuals may reveal locus-specific selection (e.g. (Dhuyvetter et al, 2004; Dufresne et al, 2002; Jordan et al, 1997; Karl and Avise, 1992; McDonald, 1991; Pogson et al, 1995; Vasemagi et al, 2005).

As a result of using both approaches, the genes of the major histocompatibility complex (MHC) are widely believed to be subject to strong balancing selection (Apanius et al, 1997; Hedrick, 1994; Hughes and Yeager, 1998). MHC molecules play a central role in the T-cell-mediated specific immune response (Klein, 1986; Parham and Ohta, 1996), encoding molecules that bind small self or non-self peptides within the cell and then present them on the cell-surface to T cells (Hedrick, 1994). Although MHC genes are among the most studied loci in vertebrates, the mechanisms that maintain their high levels of polymorphism remain vigorously debated.
Sexual selection (Consuegra and García de Leániz, 2008; Jordan and Bruford, 1998; Landry et al, 2001; Potts and Wakeland, 1990; Potts and Wakeland, 1993; Reusch et al, 2001) and neutral forces (Landry and Bernatchez, 2001; Miller and Lambert, 2004; Seddon and Baverstock, 1999; Seddon and Ellegren, 2004) may have a role in determining levels of variability in MHC genes. However, pathogen-driven balancing selection (through overdominance, negative frequency dependence or temporal/spatial heterogeneity in pathogen phenotype) is thought by many to be the main force driving MHC evolution (Edwards and Hedrick, 1998; Hedrick and Kim, 2000; Jeffery and Bangham, 2000; Klein and O'HiUigin, 1994; Parham and Ohta, 1996). Evidence of selection on MHC genes has traditionally come from four sources (Hughes and Yeager, 1998): (a) long persistence times for MHC alleles compared to neutral expectation (often resulting in trans-specific polymorphism) (Figueroa et al, 1988; Klein et al, 2007; Lawlor et al, 1988; McConnell et al, 1988), (b) frequency distributions of MHC alleles in natural populations that are more even that expected under a neutral model (Hedrick and Thompson, 1983; Markow et al, 1993), (c) high levels of non-synonymous versus synonymous substitutions in codons for peptide binding residues (PBRs) (Hughes and Nei, 1988; Hughes and Nei, 1989; Hughes and Nei, 1990) and (d) homogenisation of introns with concurrent diversification of exons at MHC loci (Cereb et al, 1997; Reusch and Langefors, 2005).
More recently, an increasing number of descriptions of the geographic
distribution of MHC variation (or variation at markers tightly linked to
MHC loci) - often in conjunction with analysis at other, (putatively) neutral
loci - have provided further insight into the selective influences on MHC
loci in a range of species (Aguilar and Garza, 2006; Alcaide et al, 2008;
Bryja et al, 2007; Ekblom et al, 2007; Miller et al, 2001). In general, MHC
heterozygosity within populations is higher than that for neutral loci
(Aguilar et al, 2004; Huang and Yu, 2003); but see (Boyce et al, 1997).
However, there appears to be a great deal of species-to-species variation in
the relative levels of MHC and neutral variability among populations.
Among-population differentiation at MHC loci has been observed to range
from lower than (Aguilar et al, 2004; Sommer, 2003), similar to (Boyce et
al, 1997; Hedrick et al, 2001; Huang and Yu, 2003; Parker et al, 1999) to
higher than that at neutral loci (Beacham et al, 2004; Miller et al, 2001),
reflecting differences in the relative strengths of natural selection, genetic
drift and gene flow across species (Eizaguirre et al, 2010). Indeed, the
relative levels of neutral and MHC variation within and among populations
can vary across closely-related species (Hambuch and Lacey, 2002; Jarvi et
al, 2004), and even within a species depending on the spatial scale of
analysis (Landry and Bernatchez, 2001). In comparison to studies of
geographic patterns in MHC variation, there have been relatively few
studies on MHC which also include a temporal dimension, despite the
possibility that selective forces across time within a population may differ from those across populations (Beacham et al, 2004; Coughlan et al, 2006; Oliver et al, 2009; Seddon and Ellegren, 2004; Smulders et al, 2003; Sommer, 2003; Westerdahl et al, 2004). Moreover, most studies focus on a single class of MHC gene (most commonly class II $\beta$), while selection can affect differentially the genetic diversity of both genes (Bryja et al, 2007).

The Atlantic salmon ($Salmo salar$) has proved to be an excellent model for the study of MHC evolution for a number of reasons. First, as in teleosts in general, MHC class I and class II loci are not physically linked in the species, allowing for independent evolution of these classes of genes (Grimholt et al, 2002). As MHC genes do not form a single complex they are therefore known simply as MH genes in teleosts (Stet et al, 2002). Second, while MH pseudogenes exist in the genome, and several non-classical class I genes have been described recently (Lukacs et al, 2010), the major classical MH genes expressed are class I ($Sasa-UBA$), class II$\alpha$ ($Sasa-DAA$) and class II$\beta$ ($Sasa-DAB$) (Grimholt et al, 2000; Grimholt et al, 1993; Stet et al, 2002), making analyses of functional MHC variation relatively simple. Third, the molecular structure of MH genes has been extensively studied in Atlantic salmon (Grimholt et al, 2002; Grimholt et al, 2000; Grimholt et al, 1993; Hordvik et al, 1993; Stet et al, 2002) and there is evidence of balancing selection acting on potential peptide binding
residues (PBRs) in both class I and class IIα loci (Consuegra et al, 2005a; Consuegra et al, 2005b). Fourth, a dinucleotide microsatellite repeat located in the 3’ untranslated region (UTR) of the MH class I locus (termed here Sasa-UBA-3UTR) (Grimholt et al, 2002) and a 10-base pair minisatellite repeat in the 3’ UTR of the MH class II α locus (termed here Sasa-DAA-3UTR) (Stet et al, 2002) can be used to rapidly assay variation at these loci in Atlantic salmon. Fifth, the Atlantic salmon is one of the few species where MH polymorphism has been experimentally shown to be associated with resistance/susceptibility to specific pathogens (Grimholt et al, 2003). Furthermore, historical samples are widely available from Atlantic salmon populations, often as dried scales that are suitable for extraction of DNA, allowing analysis of temporal variation within populations (Ciborowski et al, 2007; Consuegra et al, 2002; Nielsen et al, 1999).

The aims of this study were to test the hypotheses that (a) MH-linked markers in natural Atlantic salmon populations show temporal and spatial levels of heterogeneity that differ from a sample of neutral microsatellite loci and fall outside the range expected from neutral models and (b) MH class I and class II-linked markers show differential responses to selective forces as expected from the different evolutionary rates of the MH genes (Consuegra et al, 2005a; Consuegra et al, 2005b). To test these hypotheses two case studies involving Atlantic salmon populations in western Ireland
were used: (1) based on the Burrishoole (BRH) river system for temporal analysis and the nearby rivers Moy (MOY), Owenmore (OWM) and Owenduff (OWD) for spatial analysis and (2) based on the Delphi (DPH) river system for temporal analysis and the nearby rivers Owenwee (OWW) and Carrowinskey (CAW) for spatial analysis.
Materials and Methods

Samples

Samples were obtained from both contemporary and historical sources. For contemporary analysis, juveniles were sampled from all study rivers by electro-fishing of different sections, each approximately 200 m in length, in 2002 and fin clips stored in 95% ethanol for subsequent extraction of DNA. The maximum distance between the focal rivers and other spatial samples was 182 km (Burrishoole-Moy) and 57 km (Deplhi-Owenwee) (Figure 1).

Historical samples from the Burrishoole system were obtained from the Marine Institute (Newport) as dried scales from both angled fish and fish caught in a research fish trap in 1956, 1968, 1973, 1983 and 1995. All selected samples were from grilse salmon (adults returning to the river after a single winter in the sea) and, given the predominance of two year smolts in the Burrishoole, they mostly represented single cohorts. For the Delphi system, fin clips for DNA analyses were collected from juvenile salmon sampled by electro-fishing fish in 1993 and 2002, representing 4 different cohorts. Fish were assigned to either the 0+ or 1+ age class using length frequency analysis (cut off length was 7 cm).

DNA isolation and microsatellite genotyping
Genomic DNA was isolated from fin and scale samples (2 scales per individual) using the Wizard SV96 Genomic Purification kit (Promega), eluted in 500 µl (fin samples) or 100 µl (scale samples) elution buffer and stored at 4°C until use in PCR amplifications. All samples were amplified for eight microsatellite loci: SsoSL 85 (Slettan et al., 1995), Ssa 171, Ssa 197, Ssa 202 (O’Reilly et al., 1996), Ssa D144b, Ssa D170 (King et al., 2005), Ssa 2215SP and Ssa 1G7SP (Paterson et al., 2004). These unlinked microsatellites were selected as they behaved in a neutral manner in previous studies in Atlantic salmon (de Eyto et al., 2007). Samples were also amplified for a microsatellite repeat located in the 3’ untranslated region (UTR) of the MHC class I locus (Sasa-UBA-3UTR) and a minisatellite repeat located in the 3’ UTR of the MHC class II α locus (Sasa-DAA-3UTR) (Stet et al., 2002). Previous studies (e.g. (de Eyto et al., 2007; Grimholt et al., 2002; Stet et al., 2002) have shown that the micro/minisatellite markers used in this study are uniquely- and tightly-linked to the expressed MH loci in Atlantic salmon.

All loci were amplified with the Qiagen Multiplex Kit (Qiagen) in three multiplex reactions with the following proportions of each 100 pM primer combined in a final volume of 500 µl primer stock solution: (a) SsoS 85 (20 µl), Ssa2215SP (5 µl), Ssa171 (10 µl), SsaD144b (5 µl); (b) Ssa197 (10 µl),
Each reaction consisted of 4 µl of the multiplex mix (containing hot-start Taq polymerase, buffer and dNTPs), 0.8 µl of the primer stock; 1.2 µl of nuclease-free water and 2-3 µl of template DNA. PCR conditions were as follows: denaturation step at 95°C for 15 min followed by 30/35 cycles (modern/historical samples): 94°C 30 sec; 58°C 90 sec; 72°C 60 sec; and a final extension at 60°C for 30 min. Fragment sizes were then analysed on an Applied Biosystems ABI377 automatic sequencer and estimated with the aid of GeneScan and Genotyper software (Applied Biosystems) using an internal molecular size marker (TAMRA 350/500) as a reference standard.

All historical samples and half of the modern samples were replicated at least once and only repeatable peaks were counted as real alleles. Error rates (allelic dropouts-ADO and false alleles-FA) were estimated using GIMLET v1.3.3 (Valière, 2002), which was also used to construct consensus genotypes from the PCR replicates of each sample.

**Statistical Analysis**

**Intra-population genetic diversity**
Concordance with Hardy-Weinberg expectation (significance of $F_{IS}$ values) and linkage disequilibrium between pairs of loci were tested for each locus in all samples with GENEPOP 3.2 (Raymond and Rousset, 1995). Observed heterozygosity ($H_0$) was also calculated in GENEPOP 3.2 while allelic richness ($N_a$) was calculated using FSTAT (Goudet, 1995).

After testing for deviation from normality and heterogeneity of variances, temporal and spatial heterogeneity in mean $H_0$ and $N_a$ for microsatellites was tested using one-way ANOVA, with post hoc tests of differences between sample means where appropriate, using SYSTAT v.10. Spatial and temporal heterogeneity in frequency of heterozygotes and deviation from mean allelic richness at MH-linked loci was assessed using $G$ tests.

For temporal and spatial samples from each case study Mantel tests were performed between matrices of genetic differentiation ($F_{ST}$) at microsatellite and MH-linked markers and the significance of the test statistic was assessed by performing 10,000 permutations of the data using GENETIX (Belkhir et al, 2001). Given that $F_{ST}$ values can underestimate the differentiation between populations with highly polymorphic microsatellites and when variability differs between marker classes, we also estimated Hedrick’s standardized $G'_{ST}$ (Hedrick, 2005) that provide more robust estimates of population difference and the $D_{est}$ estimate of differentiation.
(Jost, 2008) based on allele identities, that also accounts for the fixation of alleles in different populations. $G'_{ST}$ and $D_{est}$ were estimated in SMOGD (Crawford, 2010).

Deviations of the distribution of genetic variability within populations from that expected under neutrality were tested using the Ewens-Watterson homozygosity test of neutrality (Ewens 1972; Watterson 1978) with significance assessed by the Slatkin exact $P$-test (Slatkin 1994, 1996) implemented in Arlequin v.3 (Excoffier et al. 2007). The Ewens-Watterson test compares the observed homozygosity ($F_o$) with the equilibrium homozygosity under the neutral theory ($F_e$) from a simulation of randomly generated populations (1000 in this case). Significant negative values are indicative of balancing selection while positive values indicate directional selection.

Inter-population genetic diversity

Concordance to an isolation by distance spatial model of population structure between all contemporary samples was estimated with IBDWS (IBD Web Service at http://ibdws.sdsu.edu/) (Jensen et al., 2005) for microsatellite loci and MH-linked markers separately using logarithm transformations of genetic ($F_{ST}$, $G'_{ST}$, $D_{est}$) and geographic distance (Km) between rivers. We also performed partial Mantel tests using $zt$ (Bonnet and
Van de Peer, 2002) to estimate the correlation between pairwise population
differentiation ($F_{ST}$, $G’_{ST}$) at MH-linked markers and geographical distance
while keeping differentiation at microsatellites constant, in order to test for
significant positive correlation between geographical distance and MH
differentiation independent of demographic and stochastic factors.

Further temporal and spatial analysis of evidence for selection at MH-linked
loci was performed using BayeScan (Foll & Gaggiotti 2008). BayeScan uses
an extension of the (Beaumont and Balding, 2004) method to detect outlier
loci by employing a Bayesian likelihood method. To identify loci under
selection we used 10 pilot runs of 5000 iterations to estimate the distribution
of $\alpha$ (the locus-specific component of $F_{ST}$), followed by a burn-in period of
50,000 iterations and 150,000 iterations, with sample size of 5,000 and a
thinning interval of 20 between samples. The identification of loci under
selection is based on the Bayes factor (BF), the ratio between the posterior
probabilities of two models (i.e. with and without selection). According to
the scale of evidence for BF, a BF above 100 ($\log10 > 2$; posterior
probabilities ranging between 0.99 and 1) is interpreted as decisive evidence
for selection.
Results

Intra-population variability

Numbers of alleles in the two case studies were similar: mean number per locus for the BRH study was 23.9 (range 12-52) with that for the DPH study 21.0 (range 9-33). Allele frequencies for each sample are given in Supporting Information.

There were a number of significant deviations from Hardy Weinberg equilibrium in the samples, even after correcting for the number of tests carried out using a Bonferroni procedure (Tables 1 and 2). These significant deviations were almost exclusively associated with a positive value for $F_{IS}$, indicating an excess of homozygotes in the sample. In general, the significant deviations were not associated with any particular locus. However, the earlier sample (1956) from the BRH case study (Table 1) and the OWW sample from the DPH case study (Table 2) displayed relatively high levels of deviation from Hardy Weinberg equilibrium across many loci. Only thirty three of 184 tests for linkage disequilibrium gave significant results, fifteen of them in the DPH spatial study, mostly involving Ssa171 (5 tests) and SasaDAA (5 tests) (Supporting Information). The average value of ADO was 3.1% and of FA 0.1% suggesting genotyping error rate was low.

Comparisons across loci
Mean $H_O$ values for microsatellite loci were high (>0.70) in all samples, and while values of $H_O$ for $Sasa-UBA-3UTR$ tended to fall within the 95% confidence limits for the mean value for microsatellites, those for $Sasa-DAA-3UTR$ were generally lower than the lower 95% bound for the distribution of $H_O$ among microsatellites (Figure 2). Significant heterogeneity in $H_O$ was found among temporal samples in the BRH and DPH case studies at the $Sasa-DAA-3UTR$ locus (Table 3, Figure 2) and at spatial samples at the microsatellite loci in the DPH spatial comparison. Post hoc tests showed that most of the heterogeneity in mean $H_O$ of microsatellites in the DPH case study was due to the OWW sample having a significantly lower value of $H_O$ than the samples from the DPH and CAW. Conversely, there was no evidence for significant heterogeneity in $H_O$ among spatial or temporal samples in both case studies at $Sasa-UBA-3UTR$ (Table 3, Figure 2).

There was relatively little fluctuation in values of $N_A$ over space and time compared to that for $H_O$ (Figure 3), with no significant heterogeneity detected at any locus or class of loci at any scale (Table 3). As for $H_O$, values of $N_A$ for $Sasa-DAA-3UTR$ were generally lower than those for other loci (Figure 3).
Matrices of genetic distance measured as $F_{ST}$, $G'_{ST}$ and $D_{est}$ among temporal and spatial samples in both case studies were not correlated with each other in the comparisons between microsatellites and Sasa-DAA 3UTR (BRH temporal $F_{ST}$: $Z = 0.04, P = 0.074$; $G'_{ST}$: $Z = 0.649, P = 0.056$; $D_{est}$: $Z = 0.080, P = 0.074$; BRH spatial $F_{ST}$: $Z = 0.019, P = 0.432$; $G'_{ST}$: $Z = 0.515, P = 0.450$; $D_{est}$: $Z = 0.426, P = 0.432$; DPH temporal $F_{ST}$: $Z = 0.01, P = 0.074; G'_{ST}$: $Z = 0.014, P = 0.220$; $D_{est}$: $Z = 0.014, P = 0.208$; DPH spatial $F_{ST}$: $Z = 0.001, P = 0.866; G'_{ST}$: $Z = 0.012, P = 0.886$; $D_{est}$: $Z = 0.012, P = 0.867$). In contrast, most matrices of genetic distances measured with microsatellites and Sasa-UBA-3UTR were correlated (BRH temporal $F_{ST}$: $Z = 0.010, P = 0.043$; $G'_{ST}$: $Z = 0.649, P = 0.038$; $D_{est}$: $Z = 0.584, P = 0.016$; BRH spatial $F_{ST}$: $Z = 0.024, P = 0.141; G'_{ST}$: $Z = 1.272, P = 0.450$; $D_{est}$: $Z = 1.160, P = 0.383$; DPH temporal $F_{ST}$: $Z = 0.012, P = 0.039; G'_{ST}$: $Z = 0.292, P = 0.049$; $D_{est}$: $Z = 0.265, P = 0.049$; DPH spatial $F_{ST}$: $Z = 0.002, P = 0.036; G'_{ST}$: $Z = 0.882, P = 0.025$; $D_{est}$: $Z = 1.160, P = 0.025$).

Evidence for selection on MH-linked markers - comparison with neutral models

The Ewens-Watterson test rejected the null hypothesis of neutrality for Sasa-UBA-3UTR in the OWD sample ($F_o = 0.219, F_e = 0.122$ Slatkin Exact $P = 0.014$).
The geographical distribution of variation at microsatellite loci conformed to an isolation by distance model of population structure ($F_{ST}$: $Z = -13.84$, $P = 0.049$; $G'S_T$: $Z = -18.99$, $P = 0.032$; $D_{est}$: $Z = -477.82$, $P = 0.042$), whereas this was not true of either Sasa-DAA-3UTR ($F_{ST}$: $Z = -50.05$, $P = 0.382$; $G'S_T$: $Z = -50.66$, $P = 0.331$; $D_{est}$: $Z = -1327.82$, $P = 0.355$) or Sasa-UBA-3UTR ($F_{ST}$: $Z = -49.62$, $P = 0.101$; $G'S_T$: $Z = -25.69$, $P = 0.098$; $D_{est}$: $Z = -663.46$, $P = 0.099$).

Partial Mantel tests controlling for differentiation in microsatellites did not reveal any significant correlation between differentiation at MH-linked markers between populations and geographical distance ($F_{ST}$: Sasa-DAA-3UTR: $r = -0.0095$ $P = 0.5029$ (one tailed); Sasa-UBA-3UTR: $r = 0.2406$ $P = 0.1113$ (one tailed); $G'S_T$: Sasa-DAA-3UTR: $r = 0.1627$ $P = 0.2610$ (one tailed); Sasa-UBA-3UTR: $r = 0.3113$ $P = 0.0970$ (one tailed)).

Sasa-DAA-3UTR displayed evidence of selection with high log10BF with values above 2 in spatial and temporal samples in both BRH and DPH case studies (BRH temporal $\alpha=1.42$; BRH spatial $\alpha=0.98$; DPH temporal $\alpha=1.14$; DPH spatial $\alpha=0.83$) (Figure 4). Sasa-UBA-3UTR also showed evidence of selection but only in the temporal samples of the BRH and DPH case studies, respectively (BRH temporal $\alpha=0.76$; DPH temporal $\alpha=0.83$).
SasaD144b showed evidence for balancing selection ($\alpha=-0.72$) in the Burrishole spatial study (Figure 4).
Discussion

Previous studies had shown that both neutral forces and natural selection can act in shaping variability of MHC genes (Charbonnel and Pemberton, 2005; de Eyto et al, 2007; Landry and Bernatchez, 2001; Oliver et al, 2009; Seddon and Ellegren, 2004). To our knowledge this is the first study that compares the effects of selection and neutral forces at both class I and class IIα MH-linked markers spatial and temporally. Here, using two case studies involving Atlantic salmon populations in western Ireland, we tested two hypotheses on patterns of selection on MH-linked makers. To do this we compared temporal and spatial heterogeneity at microsatellite loci, which were presumed to be neutral to the effects of natural selection, with variation at MHC-linked markers using three measures of genetic variability: heterozygosity, allelic richness and variance in allele frequency among samples. We also tested levels of temporal and spatial differentiation at MH-linked markers and microsatellites against neutral models of genetic differentiation. Finally, we compared patterns of variability between classes of MH-linked markers to look for evidence of differential selection.

Comparison among loci

Our first hypothesis was that MH-linked markers would show higher levels of heterogeneity over temporal and spatial scales than microsatellites. We found differences among loci in two out of three measures of variability.
Comparison of levels of heterozygosity across time suggested that the class IIα MH-linked locus (*Sasa-DAA-3UTR*) showed more heterogeneity than microsatellite loci. However, this contention requires careful scrutiny and justification as heterogeneity in heterozygosity could arise as a technical or sampling artefact. Analysis of the pattern of deviation from Hardy-Weinberg equilibrium suggested that older samples from the Burrishoole and samples from the Owenwee were exceptional in that they displayed significantly higher levels of homozygosity across several loci (although not in *Sasa-DAA-3UTR*). In particular the Owenwee was identified by post hoc tests as significantly lower in microsatellite heterozygosity than other samples from the same population or case study. Given the known problems with PCR amplification of DNA extracted from old scale samples (Nielsen *et al.*, 1997) and the potential for sampling a small number of families from salmonid populations (Hansen *et al.*, 1997), it may be that allelic dropout (in older Burrishoole samples) or sampling effects (in the Owenwee sample) were responsible for some of the observed temporal and spatial heterogeneity at microsatellite and MH-linked loci. However, repeated genotyping (with low levels of genotyping error rate) and sampling across several kilometres of river were specifically employed in this study to avoid these potential problems. Moreover, there is no *a priori* reason to believe, and no evidence from our data, that MH-linked markers were more susceptible to these problems than other loci. On this basis, it appears that
heterogeneity in heterozygosity was indeed greater at Sasa-DAA-3UTR than at microsatellites in our case studies.

In contrast with heterozygosity, the lack of heterogeneity observed in allelic richness is perhaps surprising given that allelic richness is often assumed to be a more sensitive measure of demographic and selective effects than heterozygosity (Leberg, 1992). Correlations between allelic richness and heterozygosity are scale dependent, and tend to be higher in nearby populations that share similar environments and selective pressures (Comps et al., 2001) but in general allelic richness correlates better with demographic patterns as a consequence of the effect of genetic drift in rare alleles, that contribute little to heterozygosity values but are more easily lost by random mating (Leberg, 2002). In fact, demographic effects such as post-glacial history and bottlenecks can result in important population structuring at MHC genes (e.g. Miller et al. (1997)) and selection on particular alleles also result in important allele differences between populations (e.g. MHC class II in Fundulus heteroclitus; Cohen (2002)). In our study, we did not observe allelic number reductions but we found some degree of population structuring at MH-linked markers suggesting that neutral forces are also important in shaping MH variability.
Analysis of $F_{ST}$, $G_{ST}$ and $D_{est}$ among samples showed no correlation between matrices of genetic differentiation based on microsatellites and the MH class II-linked locus ($Sasa-DAA-3UTR$). However, there was a correlation between matrices of genetic differentiation from microsatellites and the MH class I-linked locus ($Sasa-UBA-3UTR$) in the BRH temporal case study and in both DPH case studies (all the three distances). Results using the three estimates of genetic diversity were in general consistent. Only microsatellites conformed to an isolation by distance model of population structure, suggesting that neutral forces such as gene flow and migration played a more important role in shaping the variability of the neutral microsatellites than of the MH-linked markers (Dionne et al., 2007). Higher levels of population differentiation for MH-linked markers could be the result of directional selection and local adaptation (Cohen, 2002; Heath et al., 2006; Landry and Bernatchez, 2001).

**Comparison with neutral models**

We also hypothesized that MH-linked markers would lie outside the expectation from neutral models of population structure. Indeed, comparisons with neutral models provided additional evidence for differences in response to selection of class I and class IIα-linked markers. The Ewens-Watterson test confirmed that directional selection was acting on the $Sasa-UBA-3UTR$ with higher homozygosity than expected under
neutrality, albeit in only one of the populations (Owenduff). However, variable results of this test are commonly found when several populations are compared, probably as a result of differences in the strength of selection in different environments (Bernatchez and Landry, 2003) and among different subpopulations (Garrigan and Hedrick, 2003). In addition, the power of the Ewens-Watterson test decreases rapidly when beneficial mutations reach fixation, the time to fixation for positively selected alleles being commonly short (Zhai et al, 2009).

Perhaps the most unequivocal evidence for selection on the MH-linked markers comes from the Bayesian analysis of deviation from a neutral model of genetic variation. In all tests Sasa-DAA-3UTR was a strong outlier from the neutral expectation, while Sasa-UBA-3UTR was an outlier in the two temporal tests. The strong evidence for selection on Sasa-DAA-3UTR in this analysis is perhaps surprising given its relatively low level of polymorphism which could reduce the power of the analysis to detect selection (Foll and Gaggiotti, 2008). In contrast, microsatellites conformed to the neutral model, with only one exception (Sasa-D144b in the Burrishoole spatial analysis). Although assumed neutral, microsatellites can appear as outliers as a consequence of hitchhiking selection, and this has been seen previously in salmonids with microsatellites known to be associated to life history traits (Aguilar and Garza, 2006). However, we did
not find evidence in the literature suggesting that this could be the case for 
\textit{Sasa-D144b} and given that this only occurred in one case, while the 
evidence for selection at \textit{Sasa-UBA-3UTR} and \textit{Sasa-DAA-3UTR} was 
supported by results of several tests, we consider that this result does not 
invalidate the evidence found here for selection acting on the MHC linked 
markers.

Directional selection acting on MH class II has been previously observed in 
Atlantic salmon within a river drainage, and was attributed to local 
adaptation to environmental conditions or pathogen composition (Landry 
and Bernatchez, 2001). An extensive study in populations of sockeye 
salmon found evidence for directional and balancing selection acting on MH 
class II (Miller \textit{et al}, 2001), with a high degree of population differentiation 
at the MH locus that may have resulted from directional selection spatial 
and/or temporal variability in pathogen composition. The directional 
selection we observed could also be the result of spatial and temporal 
variation in pathogen composition, if the selective advantage of MH alleles 
differed among environments (Bernatchez and Landry, 2003) and this was 
reflected in the linked markers studied here. In the Burrishoole system a 
previous study linking MH variability to survival and fitness in Atlantic 
salmon under natural conditions found evidence for disease-mediated 
selection acting on both the MH class II locus and the linked marker \textit{Sasa-}
Moreover, the results suggested that additive allelic effects were more important than heterozygote advantage for individual survival (de Eyto et al., 2007). Sexual selection could also favour individuals carrying specific resistance alleles against common parasites in different populations, and differences in MHC composition due to divergent parasite-mediated selection could be further maintained by assortative mating of females with locally adapted males (Eizaguirre et al., 2010; Eizaguirre et al., 2009).

**Differences between MH-linked markers**

Our second hypothesis was that MH-linked markers would show different patterns of selection. To our knowledge, this is the first time that the response to selection has been compared between MHC class I and class II-linked markers at both temporal and spatial scales. Collectively, our results indicate that the diversity of both MH-linked markers may be affected by selection, while undoubtedly neutral forces (genetic drift) also play a role. We found strong evidence for directional selection acting on Sasa-DAA-3UTR; we also found some evidence for directional selection on Sasa-UBA-3UTR but at a different scale. While Sasa-DAA-3UTR showed evidence for directional selection at both temporal and spatial studies, we only found evidence of selection acting on Sasa-UBA-3UTR mostly in the temporal
comparisons and only in one of the populations for the Burrishoole spatial study.

Previous studies in salmonids found temporal stability at MHC class I (Hansen et al., 2007) and class II (Miller et al., 2001) genes and higher levels of population structuring at MHC class II (Heath et al., 2006; Landry and Bernatchez, 2001), while guppies seem to show low levels of differentiation on MH class II genes (Van Oosterhout et al., 2006) and evidence from mammals and birds suggests that there is temporal variation in selection acting at MHC class II and class I genes (Charbonnel and Pemberton, 2005; Westerdahl et al., 2004). Our finding of different responses to selection of an MH class IIα-linked marker and a class I-linked marker might reflect an association between level of environmental heterogeneity (within and among rivers) and possibly differences in pathogen-driven selective pressures acting on variability at each class of MH locus (Paterson, 1998; Wegner et al., 2008; Wegner et al., 2003). Although we cannot automatically assume that the genetic variation at the MHC loci is reflected in the linked microsatellite loci, previous studies (e.g. (Grimholt et al., 2002; Stet et al., 2002)) have shown that the micro/minisatellite markers used in our current study are uniquely- and tightly- linked to the expressed MH loci in Atlantic salmon, and can be used as good proxies for functional variation in MH genes. Class I and class II genes have different patterns of variability in
teleosts. For example, (Consuegra et al, 2005a) observed greater divergence of alleles in class I in contrast with an overlap of most of the allelic composition of class II in two isolated Atlantic salmon populations. In contrast, (Kruiswijk et al, 2005) found a complete divergence in class II but not in class I of a barbs species flock. Differences in evolutionary rates and the response to selection between both genes have been observed not only in fish but also in mammals and birds (Bryja et al, 2007; Go et al, 2003). Class I and class II molecules differ in the nature of the peptides that they bind, in how they bind and process them (Castellino et al, 1997; Go et al, 2003; Grommé and Neefjes, 2002; Kaufman et al, 1999) and in the T cells they react with (Housset and Malissen, 2003; Huseby et al, 2003). Our results suggest that differences in the response to selection of class I and class II markers could reflect structural and functional differences in the genes to which they are linked. Given the increasing importance of using MHC-linked markers for conservation and management purposes (Bos et al, 2008; Hedrick et al, 2001) our results highlight the need to consider both loci in population and evolutionary studies.
Acknowledgements

Funding for this study was supplied by a contract from the European Commission (Salimpact: QLRT-2000-01185). We want to thank the field staff of the Marine Institute Newport for assistance with sampling and Carlos García de Leániz, Christophe Eizaguirre and an anonymous referee for their useful comments on the manuscript. PMcG was part supported by the Beaufort Marine Research award in Fish Population Genetics funded by the Irish Government under the Sea Change Programme. The authors declare no conflict of interest.
References


Using Microsatellites and Major Histocompatibility Complex Variation.


not class I, genes among different species. *Immunogenetics* 56(12): 894-908.


**Figure Legends**

**Fig. 1** Map showing the location of the study rivers for the Burrishoole and Delphi case studies.

**Fig. 2** Patterns of observed heterozygosity ($H_O$) at microsatellite loci (mean ± 95% confidence intervals) and MH-linked markers across time in the (a) Burrishoole and (b) Delphi river systems and across space for samples in the (c) Burrishoole and (d) Delphi case studies.

**Fig. 3** Patterns of allelic richness ($N_A$) at microsatellite loci (mean ± 95% confidence intervals) and MH-linked markers across time in the (a) Burrishoole and (b) Delphi river systems and across space for samples in the (c) Burrishoole and (d) Delphi case studies.

**Fig. 4** Spatial and temporal analyses for identification of loci potentially subject to differential selection. a) Burrishole temporal case, b) Delphi temporal case c) Burrishole spatial case and d) Delphi spatial case. Vertical lines mark the Log10 of the Bayes factors estimated using BAYESCAN, Log10(BF)=2 corresponding to posterior probabilities of locus effects of 0.99 (decisive). Only loci identified as under selection are labelled.
Table 1 Values of $F_{IS}$ (Weir & Cockerham 1984) for samples from the Burrishoole case study and the significance of deviation of genotype frequencies from Hardy Weinberg equilibrium ($P$). Values in bold are those that remained significant after Bonferroni correction for number of tests in this series ($\alpha=0.05/90=0.0006$). River abbreviations: Burrishoole (BRH), Moy (MOY), Owenmore (OWM), Owenduff (OWD), Delphi (DPH), Owenwee (OWW) and Carrowinskey (CAW).

<table>
<thead>
<tr>
<th></th>
<th>BRH</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>OWD</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$Ssa \ D144b$</td>
<td>$F_{IS}$</td>
<td>+0.245</td>
<td>+0.277</td>
<td>+0.170</td>
<td>+0.183</td>
<td>+0.160</td>
<td>+0.028</td>
<td>+0.023</td>
<td>+0.081</td>
<td>+0.150</td>
</tr>
<tr>
<td></td>
<td>$P$</td>
<td>&lt;0.0001</td>
<td>0.0007</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.0179</td>
<td>0.0476</td>
<td>0.2994</td>
<td>0.0805</td>
<td>0.0417</td>
</tr>
<tr>
<td>$Ssa \ 171$</td>
<td>$F_{IS}$</td>
<td>+0.134</td>
<td>+0.120</td>
<td>-0.026</td>
<td>-0.117</td>
<td>+0.012</td>
<td>+0.067</td>
<td>+0.026</td>
<td>+0.089</td>
<td>+0.299</td>
</tr>
<tr>
<td></td>
<td>$P$</td>
<td>0.1450</td>
<td>&lt;0.0001</td>
<td>0.8374</td>
<td>0.4654</td>
<td>0.2267</td>
<td>&lt;0.0001</td>
<td>0.2880</td>
<td>0.0512</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>$Ssa \ 2215SP$</td>
<td>$F_{IS}$</td>
<td>-0.026</td>
<td>-0.009</td>
<td>+0.076</td>
<td>+0.116</td>
<td>+0.326</td>
<td>+0.205</td>
<td>-0.001</td>
<td>+0.092</td>
<td>+0.130</td>
</tr>
<tr>
<td></td>
<td>$P$</td>
<td>0.1413</td>
<td>0.0520</td>
<td>0.0655</td>
<td>0.0119</td>
<td>&lt;0.0001</td>
<td>0.0014</td>
<td>0.1354</td>
<td>0.0621</td>
<td>0.0145</td>
</tr>
<tr>
<td>$SsoSL \ 85$</td>
<td>$F_{IS}$</td>
<td>+0.275</td>
<td>+0.189</td>
<td>+0.087</td>
<td>-0.049</td>
<td>+0.158</td>
<td>+0.248</td>
<td>+0.119</td>
<td>+0.142</td>
<td>+0.407</td>
</tr>
<tr>
<td></td>
<td>$P$</td>
<td>&lt;0.0001</td>
<td>0.0094</td>
<td>0.0160</td>
<td>0.2018</td>
<td>0.0142</td>
<td>&lt;0.0001</td>
<td>0.1858</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Species</td>
<td>F&lt;sub&gt;IS&lt;/sub&gt;</td>
<td>P</td>
<td>F&lt;sub&gt;IS&lt;/sub&gt;</td>
<td>P</td>
<td>F&lt;sub&gt;IS&lt;/sub&gt;</td>
<td>P</td>
<td>F&lt;sub&gt;IS&lt;/sub&gt;</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa 197</td>
<td>+0.113</td>
<td>0.0010</td>
<td>+0.085</td>
<td>0.0025</td>
<td>+0.162</td>
<td>0.0324</td>
<td>-0.067</td>
<td>0.4183</td>
<td>+0.002</td>
<td>0.3111</td>
</tr>
<tr>
<td>Ssa 1G7SP</td>
<td>+0.079</td>
<td>0.0018</td>
<td>+0.259</td>
<td>0.0211</td>
<td>+0.292</td>
<td>0.0113</td>
<td>-0.101</td>
<td>0.559</td>
<td>+0.075</td>
<td>0.5648</td>
</tr>
<tr>
<td>Ssa 202</td>
<td>+0.447</td>
<td>&lt;0.0001</td>
<td>+0.230</td>
<td>0.0010</td>
<td>+0.251</td>
<td>0.0234</td>
<td>+0.139</td>
<td>0.0023</td>
<td>+0.158</td>
<td>0.1124</td>
</tr>
<tr>
<td>Ssa D170</td>
<td>+0.319</td>
<td>&lt;0.0001</td>
<td>+0.112</td>
<td>&lt;0.0001</td>
<td>+0.189</td>
<td>&lt;0.0001</td>
<td>+0.213</td>
<td>&lt;0.0001</td>
<td>+0.086</td>
<td>0.0334</td>
</tr>
<tr>
<td>Sasa-DAA-3UTR</td>
<td>+0.082</td>
<td>0.5158</td>
<td>+0.125</td>
<td>0.4199</td>
<td>-0.044</td>
<td>0.1343</td>
<td>+0.051</td>
<td>0.8610</td>
<td>+0.148</td>
<td>0.0873</td>
</tr>
<tr>
<td>Sasa-UBA-3UTR</td>
<td>+0.052</td>
<td>0.1460</td>
<td>-0.102</td>
<td>0.2183</td>
<td>+0.161</td>
<td>0.2725</td>
<td>+0.297</td>
<td>&lt;0.0001</td>
<td>+0.028</td>
<td>0.0746</td>
</tr>
</tbody>
</table>
Table 2 Values of $F_{IS}$ (Weir & Cockerham 1984) for samples from the Delphi case study and the significance of deviation of genotype frequencies from Hardy Weinberg equilibrium ($P$). Values in bold are those that remained significant after Bonferroni correction for number of tests in this series ($\alpha=0.05/60=0.0008$). River abbreviations: Burrishoole (BRH), Moy (MOY), Owenmore (OWM), Owenduff (OWD), Delphi (DPH), Owenwee (OWW) and Carrowinskey (CAW).

<table>
<thead>
<tr>
<th></th>
<th>DPH</th>
<th></th>
<th></th>
<th></th>
<th>OWW (N=55)</th>
<th>CAW (N=48)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1993 0+ (N=45)</td>
<td>1993 1+ (N=50)</td>
<td>2002 0+ (N=45)</td>
<td>2002 1+ (N=51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ssa D144b</strong></td>
<td>$F_{IS}$ -0.007</td>
<td>+0.114</td>
<td>+0.121</td>
<td>+0.157</td>
<td>+0.031</td>
<td>+0.029</td>
</tr>
<tr>
<td></td>
<td>$P$ 0.2089</td>
<td>0.0055</td>
<td>0.1098</td>
<td>0.0190</td>
<td>&gt;0.0001</td>
<td>0.0545</td>
</tr>
<tr>
<td><strong>Ssa 171</strong></td>
<td>$F_{IS}$ +0.013</td>
<td>-0.012</td>
<td>-0.147</td>
<td>+0.028</td>
<td>+0.028</td>
<td>+0.051</td>
</tr>
<tr>
<td></td>
<td>$P$ 0.0460</td>
<td>0.5265</td>
<td>0.9571</td>
<td>0.2574</td>
<td>0.0062</td>
<td>0.0357</td>
</tr>
<tr>
<td><strong>Ssa 2215SP</strong></td>
<td>$F_{IS}$ -0.009</td>
<td>+0.070</td>
<td>+0.113</td>
<td>+0.074</td>
<td>+0.281</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>$P$ 0.8879</td>
<td>0.3332</td>
<td>0.0044</td>
<td>0.2369</td>
<td>0.0002</td>
<td>0.1871</td>
</tr>
<tr>
<td><strong>SsoSL 85</strong></td>
<td>$F_{IS}$ +0.025</td>
<td>+0.082</td>
<td>+0.004</td>
<td>-0.003</td>
<td>+0.063</td>
<td>+0.183</td>
</tr>
<tr>
<td></td>
<td>$P$ 0.1160</td>
<td>0.0158</td>
<td>0.5275</td>
<td>0.5162</td>
<td>0.0066</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Ssa 197</td>
<td>+0.012</td>
<td>+0.036</td>
<td>-0.049</td>
<td>-0.002</td>
<td>+0.575</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>0.5630</td>
<td>0.2662</td>
<td>0.4270</td>
<td>0.6161</td>
<td>&gt;0.0001</td>
<td>0.7733</td>
</tr>
<tr>
<td>Ssa 1G7SP</td>
<td>+0.069</td>
<td>+0.013</td>
<td>+0.067</td>
<td>+0.172</td>
<td>+0.304</td>
<td>+0.032</td>
</tr>
<tr>
<td></td>
<td>0.7914</td>
<td>0.0505</td>
<td>0.3465</td>
<td>&gt;0.0001</td>
<td>0.0001</td>
<td>0.0592</td>
</tr>
<tr>
<td>Ssa 202</td>
<td>+0.176</td>
<td>+0.365</td>
<td>-0.040</td>
<td>+0.012</td>
<td>+0.271</td>
<td>+0.210</td>
</tr>
<tr>
<td></td>
<td>0.0013</td>
<td>&gt;0.0001</td>
<td>0.5938</td>
<td>0.1613</td>
<td>0.0033</td>
<td>&gt;0.0001</td>
</tr>
<tr>
<td>Ssa D170</td>
<td>+0.279</td>
<td>+0.130</td>
<td>+0.059</td>
<td>+0.129</td>
<td>+0.240</td>
<td>+0.007</td>
</tr>
<tr>
<td></td>
<td>&gt;0.0001</td>
<td>0.0036</td>
<td>&gt;0.0001</td>
<td>0.0011</td>
<td>&gt;0.0001</td>
<td>&gt;0.0001</td>
</tr>
<tr>
<td>Sasa-DAA-3UTR</td>
<td>+0.132</td>
<td>+0.303</td>
<td>+0.104</td>
<td>-0.005</td>
<td>+0.118</td>
<td>-0.223</td>
</tr>
<tr>
<td></td>
<td>0.0038</td>
<td>0.0029</td>
<td>0.0674</td>
<td>0.6681</td>
<td>0.6304</td>
<td>0.0013</td>
</tr>
<tr>
<td>Sasa-UBA-3UTR</td>
<td>+0.113</td>
<td>+0.114</td>
<td>+0.115</td>
<td>+0.262</td>
<td>+0.313</td>
<td>+0.027</td>
</tr>
<tr>
<td></td>
<td>0.3261</td>
<td>0.0055</td>
<td>0.1193</td>
<td>&gt;0.0001</td>
<td>&gt;0.0001</td>
<td>0.0390</td>
</tr>
</tbody>
</table>
Table 3 Results of tests of temporal and spatial heterogeneity in mean $H_O$ and $N_A$ at microsatellite loci, the frequency of heterozygotes at MH-linked makers and deviation from mean $N_A$ at MH-linked makers. Numbers in subscript indicate degrees of freedom associated with each test. Values in bold were significant at the $\alpha=0.05$ level.

<table>
<thead>
<tr>
<th></th>
<th>Temporal</th>
<th>Spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Burrishoole</td>
<td>Delphi</td>
</tr>
<tr>
<td>microsatellites</td>
<td>$F_{5,42} = 1.24, P = 0.309$</td>
<td>$F_{3,28} = 0.292, P = 0.786$</td>
</tr>
<tr>
<td>Heterozygosity $(H_O)$</td>
<td>Sasa-DAA-3UTR $G_5= 23.07, P &lt;0.001$</td>
<td>Sasa-UBA-3UTR $G_5= 3.595, P = 0.606$</td>
</tr>
<tr>
<td>Allelic Richness $(N_A)$</td>
<td>Sasa-DAA-3UTR $G_5= 1.59, P = 0.896$</td>
<td>Sasa-UBA-3UTR $G_5= 1.61, P = 0.893$</td>
</tr>
</tbody>
</table>
Figure 1. Location of the study rivers.
Figure 2. Heterozygosity at MHC and microsatellites.
Figure 3. Allelic richness at MHC and microsatellites.
Figure 4. Results from bayesian analysis of selection.