Contrasting Responses to Selection in Class I and Class IIα Major Histocompatibility Linked Markers in Salmon

S. CONSUEGRA[1,2], E. DE EYTO[3], P. MCGINNITY[3,5], R. J. M. STET[4]† and W. C. JORDAN[1]

[1]Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

[2]Aberystwyth University, Institute of Biological, Environmental & Rural Sciences, Aberystwyth SY23 3DA, UK

[3]Aquaculture and Catchment Management Services, Marine Institute, Newport, Co Mayo, Republic of Ireland

[4]University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, Scotland

[5]Present 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

(Spurgin and Richardson, 2010). 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'HUigin, 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 β), 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 IIα (*Sasa-DAA*) and class IIβ (*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

140 residues (PBRs) in both class I and class IIα loci (Consuegra *et al*, 2005a;
141 Consuegra *et al*, 2005b). Fourth, a dinucleotide microsatellite repeat located
142 in the 3' untranslated region (UTR) of the MH class I locus (termed here
143 *Sasa-UBA*-3UTR) (Grimholt *et al*, 2002) and a 10-base pair minisatellite
144 repeat in the 3' UTR of the MH class II α locus (termed here *Sasa-DAA*-
145 3UTR) (Stet *et al*, 2002) can be used to rapidly assay variation at these loci
146 in Atlantic salmon. Fifth, the Atlantic salmon is one of the few species
147 where MH polymorphism has been experimentally shown to be associated
148 with resistance/susceptibility to specific pathogens (Grimholt *et al*, 2003).
149 Furthermore, historical samples are widely available from Atlantic salmon
150 populations, often as dried scales that are suitable for extraction of DNA,
151 allowing analysis of temporal variation within populations (Ciborowski *et*
152 *al*, 2007; Consuegra *et al*, 2002; Nielsen *et al*, 1999).
153
154 The aims of this study were to test the hypotheses that (a) MH-linked
155 markers in natural Atlantic salmon populations show temporal and spatial
156 levels of heterogeneity that differ from a sample of neutral microsatellite
157 loci and fall outside the range expected from neutral models and (b) MH
158 class I and class II-linked markers show differential responses to selective
159 forces as expected from the different evolutionary rates of the MH genes
160 (Consuegra *et al*, 2005a; Consuegra *et al*, 2005b). To test these hypotheses
161 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 182km (Burrishoole-Moy) and 57km (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

188 Genomic DNA was isolated from fin and scale samples (2 scales per
189 individual) using the Wizard SV96 Genomic Purification kit (Promega),
190 eluted in 500 µl (fin samples) or 100 µl (scale samples) elution buffer and
191 stored at 4ºC until use in PCR amplifications. All samples were amplified
192 for eight microsatellite loci: *SsoSL 85* (Slettan *et al*, 1995), *Ssa 171*, *Ssa*
193 *197*, *Ssa 202* (O'Reilly *et al*, 1996), *Ssa D144b*, *Ssa D170* (King *et al*,
194 2005), *Ssa 2215SP* and *Ssa 1G7SP* (Paterson *et al*, 2004). These unlinked
195 microsatellites were selected as they behaved in a neutral manner in
196 previous studies in Atlantic salmon (de Eyto *et al*, 2007). Samples were also
197 amplified for a microsatellite repeat located in the 3' untranslated region
198 (UTR) of the MHC class I locus (*Sasa-UBA-3UTR*) and a minisatellite
199 repeat located in the 3' UTR of the MHC class II α locus (*Sasa-DAA-3UTR*)
200 (Stet *et al*, 2002). Previous studies (e.g. (de Eyto *et al*, 2007; Grimholt *et al*,
201 2002; Stet *et al*, 2002) have shown that the micro/minisatellite markers used
202 in this study are uniquely- and tightly-linked to the expressed MH loci in
203 Atlantic salmon.
204
205 All loci were amplified with the Qiagen Multiplex Kit (Qiagen) in three
206 multiplex reactions with the following proportions of each 100 pM primer
207 combined in a final volume of 500 µl primer stock solution: (a) *SsoS 85* (20
208 µl), *Ssa2215SP* (5 µl), *Ssa171* (10 µl), *SsaD144b* (5 µl); (b) *Ssa197* (10 µl),

SsaD170 (10 µl), *Ssa1G7SP* (10 µl), *Ssa202* (10 µl); and (c) *Sasa-UBA-3UTR* (10 µl); *Sasa-DAA-3UTR* (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 95ºC for 15 min followed by 30/35 cycles (modern/historical samples): 94ºC 30 sec; 58ºC 90 sec; 72ºC 60 sec; and a final extension at 60ºC 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 heterozysity (H_O) 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_O 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

252 (Jost, 2008) based on allele identities, that also accounts for the fixation of
253 alleles in different populations. G'_{ST} and D_{est} were estimated in SMOGD
254 (Crawford, 2010).
255
256 Deviations of the distribution of genetic variability within populations from
257 that expected under neutrality were tested using the Ewens-Watterson
258 homozyosity test of neutrality (Ewens 1972; Watterson 1978) with
259 significance assessed by the Slatkin exact *P*-test (Slatkin 1994, 1996)
260 implemented in Arlequin v.3 (Excoffier *et al.* 2007). The Ewens-Watterson
261 test compares the observed homozygosity (F_o) with the equilibrium
262 homozygosity under the neutral theory (F_e) from a simulation of randomly
263 generated populations (1000 in this case). Significant negative values are
264 indicative of balancing selection while positive values indicate directional
265 selection.
266
267 *Inter-population genetic diversity*
268 Concordance to an isolation by distance spatial model of population
269 structure between all contemporary samples was estimated with IBDWS
270 (IBD Web Service at http://ibdws.sdsu.edu/) (Jensen *et al*, 2005) for
271 microsatellite loci and MH-linked markers separately using logarithm
272 transformations of genetic (F_{ST}, $G'_{ST,}$ D_{est}) and geographic distance (Km)
273 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 α (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

316 Mean H_O values for microsatellite loci were high (>0.70) in all samples, and
317 while values of H_O for *Sasa-UBA-3UTR* tended to fall within the 95%
318 confidence limits for the mean value for microsatellites, those for *Sasa-*
319 *DAA-3UTR* were generally lower than the lower 95% bound for the
320 distribution of H_O among microsatellites (Figure 2). Significant
321 heterogeneity in H_O was found among temporal samples in the BRH and
322 DPH case studies at the *Sasa-DAA-3UTR* locus (Table 3, Figure 2) and at
323 spatial samples at the microsatellite loci in the DPH spatial comparison.
324 *Post hoc* tests showed that most of the heterogeneity in mean H_O of
325 microsatellites in the DPH case study was due to the OWW sample having a
326 significantly lower value of H_O than the samples from the DPH and CAW.
327 Conversely, there was no evidence for significant heterogeneity in H_O
328 among spatial or temporal samples in both case studies at *Sasa-UBA-3UTR*
329 (Table 3, Figure 2).
330
331 There was relatively little fluctuation in values of N_A over space and time
332 compared to that for H_O (Figure 3), with no significant heterogeneity
333 detected at any locus or class of loci at any scale (Table 3). As for H_O,
334 values of N_A for *Sasa-DAA-3UTR* were generally lower than those for other
335 loci (Figure 3).
336

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'_{ST}: 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'_{ST}: 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'_{ST}: 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'_{ST}: *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 α=1.42; BRH spatial α=0.98; DPH temporal α=1.14; DPH spatial α=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 α=0.76; DPH temporal α=0.83).

379 *SasaD144b* showed evidence for balancing selection (α=-0.72) in the
380 Burrishole spatial study (Figure 4).
381

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.

447 Analysis of $F_{ST,}$ G'_{ST} and D_{est} among samples showed no correlation
448 between matrices of genetic differentiation based on microsatellites and the
449 MH class II-linked locus (*Sasa-DAA-3UTR*). However, there was a
450 correlation between matrices of genetic differentiation from microsatellites
451 and the MH class I-linked locus (*Sasa-UBA-3UTR*) in the BRH temporal
452 case study and in both DPH case studies (all the three distances). Results
453 using the three estimates of genetic diversity were in general consistent.
454 Only microsatellites conformed to an isolation by distance model of
455 population structure, suggesting that neutral forces such as gene flow and
456 migration played a more important role in shaping the variability of the
457 neutral microsatellites than of the MH-linked markers (Dionne *et al*, 2007).
458 Higher levels of population differentiation for MH-linked markers could be
459 the result of directional selection and local adaptation (Cohen, 2002; Heath
460 *et al*, 2006; Landry and Bernatchez, 2001).
461
462 *Comparison with neutral models*
463 We also hypothesized that MH-linked markers would lie outside the
464 expectation from neutral models of population structure. Indeed,
465 comparisons with neutral models provided additional evidence for
466 differences in response to selection of class I and class IIα-linked markers.
467 The Ewens-Watterson test confirmed that directional selection was acting
468 on the *Sasa-UBA-3UTR* with higher homozygosity than expected under

469 neutrality, albeit in only one of the populations (Owenduff). However,
470 variable results of this test are commonly found when several populations
471 are compared, probably as a result of differences in the strength of selection
472 in different environments (Bernatchez and Landry, 2003) and among
473 different subpopulations (Garrigan and Hedrick, 2003). In addition, the
474 power of the Ewens-Watterson test decreases rapidly when beneficial
475 mutations reach fixation, the time to fixation for positively selected alleles
476 being commonly short (Zhai *et al*, 2009).
477
478 Perhaps the most unequivocal evidence for selection on the MH-linked
479 markers comes from the Bayesian analysis of deviation from a neutral
480 model of genetic variation. In all tests *Sasa-DAA-3UTR* was a strong outlier
481 from the neutral expectation, while *Sasa-UBA-3UTR* was an outlier in the
482 two temporal tests. The strong evidence for selection on *Sasa-DAA-3UTR* in
483 this analysis is perhaps surprising given its relatively low level of
484 polymorphism which could reduce the power of the analysis to detect
485 selection (Foll and Gaggiotti, 2008). In contrast, microsatellites conformed
486 to the neutral model, with only one exception (*Sasa-D144b* in the
487 Burrishoole spatial analysis). Although assumed neutral, microsatellites can
488 appear as outliers as a consequence of hitchhiking selection, and this has
489 been seen previously in salmonids with microsatellites known to be
490 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 *Sasa-D144b* and given that this only occurred in one case, while the evidence for selection at *Sasa-UBA-3UTR* and *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 *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 *Sasa-*

513 *DAA-3UTR* (de Eyto *et al*, 2007). Moreover, the results suggested that
514 additive allelic effects were more important than heterozygote advantage for
515 individual survival (de Eyto *et al*, 2007). Sexual selection could also favour
516 individuals carrying specific resistance alleles against common parasites in
517 different populations, and differences in MHC composition due to divergent
518 parasite-mediated selection could be further maintained by assortative
519 mating of females with locally adapted males (Eizaguirre *et al*, 2010;
520 Eizaguirre *et al*, 2009).
521
522 *Differences between MH-linked markers*
523 Our second hypothesis was that MH-linked markers would show different
524 patterns of selection. To our knowledge, this is the first time that the
525 response to selection has been compared between MHC class I and class II-
526 linked markers at both temporal and spatial scales. Collectively, our results
527 indicate that the diversity of both MH-linked markers may be affected by
528 selection, while undoubtedly neutral forces (genetic drift) also play a role.
529 We found strong evidence for directional selection acting on *Sasa-DAA-*
530 *3UTR*; we also found some evidence for directional selection on *Sasa-UBA-*
531 *3UTR* but at a different scale. While *Sasa-DAA-3UTR* showed evidence for
532 directional selection at both temporal and spatial studies, we only found
533 evidence of selection acting on *Sasa-UBA-3UTR* mostly in the temporal

534 comparisons and only in one of the populations for the Burrishoole spatial
535 study.
536
537 Previous studies in salmonids found temporal stability at MHC class I
538 (Hansen *et al*, 2007) and class II (Miller *et al*, 2001) genes and higher levels
539 of population structuring at MHC class II (Heath *et al*, 2006; Landry and
540 Bernatchez, 2001), while guppies seem to show low levels of differentiation
541 on MH class II genes (Van Oosterhout *et al*, 2006) and evidence from
542 mammals and birds suggests that there is temporal variation in selection
543 acting at MHC class II and class I genes (Charbonnel and Pemberton, 2005;
544 Westerdahl *et al*, 2004). Our finding of different responses to selection of an
545 MH class IIα-linked marker and a class I-linked marker might reflect an
546 association between level of environmental heterogeneity (within and
547 among rivers) and possibly differences in pathogen-driven selective
548 pressures acting on variability at each class of MH locus (Paterson, 1998;
549 Wegner *et al*, 2008; Wegner *et al*, 2003). Although we cannot automatically
550 assume that the genetic variation at the MHC loci is reflected in the linked
551 microsatellite loci, previous studies (e.g. (Grimholt *et al*, 2002; Stet *et al*,
552 2002)) have shown that the micro/minisatellite markers used in our current
553 study are uniquely- and tightly- linked to the expressed MH loci in Atlantic
554 salmon, and can be used as good proxies for functional variation in MH
555 genes. Class I and class II genes have different patterns of variability in

556 teleosts. For example, (Consuegra *et al*, 2005a) observed greater divergence
557 of alleles in class I in contrast with an overlap of most of the allelic
558 composition of class II in two isolated Atlantic salmon populations. In
559 contrast, (Kruiswijk *et al*, 2005) found a complete divergence in class II but
560 not in class I of a barbs species flock. Differences in evolutionary rates and
561 the response to selection between both genes have been observed not only in
562 fish but also in mammals and birds (Bryja *et al*, 2007; Go *et al*, 2003). Class
563 I and class II molecules differ in the nature of the peptides that they bind, in
564 how they bind and process them (Castellino *et al*, 1997; Go *et al*, 2003;
565 Grommé and Neefjes, 2002; Kaufman *et al*, 1999) and in the T cells they
566 react with (Housset and Malissen, 2003; Huseby *et al*, 2003). Our results
567 suggest that differences in the response to selection of class I and class II
568 markers could reflect structural and functional differences in the genes to
569 which they are linked. Given the increasing importance of using MHC-
570 linked markers for conservation and management purposes (Bos *et al*, 2008;
571 Hedrick *et al*, 2001) our results highlight the need to consider both loci in
572 population and evolutionary studies.
573

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

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* Walbaum). *Mol Ecol* **15**(4)**:** 923-937.

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc Natl Acad Sci U S A* **101:** 3490-3494.

Alcaide M, Edwards SV, Negro JJ, Serrano D, Tella JOS, L (2008). Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (Falco naumanni). *Mol Ecol* **17:** 2652-2665.

Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997). The nature of selection on the major histocompatibility complex. *Crit Rev Immunol* **17**(2)**:** 179 - 224.

Beacham TD, Lapointe M, Candy JR, McIntosh B, MacConnachie C, Tabata A *et al* (2004). Stock Identification of Fraser River Sockeye Salmon

606 Using Microsatellites and Major Histocompatibility Complex Variation.
607 *Trans Am Fish Soc* **133**(5)**:** 1117-1137.
608
609 Beaumont MA, Balding DJ (2004). Identifying adaptive genetic divergence
610 among populations from genome scans. *Mol Ecol* **13:** 969-980.
611
612 Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. (2001).
613 Laboratoire Génome, Populations, Interactions, CNRS UMR.
614
615 Bernatchez L, Landry C (2003). MHC studies in non-model vertebrates:
616 What have we learned about natural selection in 15 years? *J Evol Biol* **16:**
617 363 - 377.
618
619 Bonnet E, Van de Peer Y (2002). zt: a software tool for simple and partial
620 Mantel tests. *J Stat Soft* **7**(10)**:** 1-12.
621
622 Bos DH, Gopurenko D, Williams RN, DeWoody AJ, Knowles L (2008).
623 Inferring Population History and Demography Using Microsatellites,
624 Mitochondrial DNA, and Major Histocompatibility Complex (MHC) Genes.
625 *Evolution* **62**(6)**:** 1458-1468.
626

627 Boyce WM, Hedrick PW, Muggli-Cockett NE, Kalinowski S, Penedo MCT,
628 Ramey-Ii RR (1997). Genetic Variation of Major Histocompatibility
629 Complex and Microsatellite Loci: A Comparison in Bighorn Sheep.
630 *Genetics* **145**(2)**:** 421-433.
631
632 Bryja J, Charbonnel N, Berthier K, Galan M, Cosson JF (2007). Density-
633 related changes in selection pattern for major histocompatibility complex
634 genes in fluctuating populations of voles. *Mol Ecol* **16:** 5084-5097.
635
636 Castellino F, Zhong G, Germain RN (1997). Antigen Presentation by MHC
637 Class II Molecules: Invariant Chain Function, Protein Trafficking, and the
638 Molecular Basis of Diverse Determinant Capture. *Hum Immunol* **54**(2)**:**
639 159-169.
640
641 Cereb N, Hughes AL, Yang SY (1997). Locus-specific conservation of the
642 HLA class I introns by intra-locus homogenization. *Immunogenetics* **47**(1)**:**
643 30-36.
644
645 Charbonnel N, Pemberton J (2005). A long-term genetic survey of an
646 ungulate population reveals balancing selection acting on MHC through
647 spatial and temporal fluctuations in selection. *Heredity* **95**(5)**:** 377-388.
648

649 Ciborowski K, Consuegra S, García de Leániz C, Wang J, Beaumont M,
650 Jordan W (2007). Stocking may increase mitochondrial DNA diversity but
651 fails to halt the decline of endangered Atlantic salmon populations. *Conserv*
652 *Genet* **8**(6)**:** 1355-1367.
653
654 Cohen S (2002). Strong Positive Selection and Habitat-Specific Amino Acid
655 Substitution Patterns in Mhc from an Estuarine Fish Under Intense Pollution
656 Stress. *Mol Biol Evol* **19**(11)**:** 1870-1880.
657
658 Comps B, Gomory D, Letouzey J, Thiebaut B, Petit RJ (2001). Diverging
659 Trends Between Heterozygosity and Allelic Richness During Postglacial
660 Colonization in the European Beech. *Genetics* **157**(1)**:** 389-397.
661
662 Consuegra S, García de Leániz C (2008). MHC-mediated mate choice
663 increases parasite resistance in salmon. *Proc R Soc Biol Sci Ser B*
664 **275**(1641)**:** 1397-1403.
665
666 Consuegra S, GarcíaDeLeániz C, Serdio A, Morales MG, Straus LG, Knox
667 D *et al* (2002). Mitochondrial DNA variation in Pleistocene and modern
668 Atlantic salmon from the Iberian glacial refugium. *Mol Ecol* **11**(10)**:** 2037-
669 2048.
670

671 Consuegra S, Megens HJ, Leon K, Stet RJM, Jordan WC (2005a). Patterns
672 of variability at the major histocompatibility class II alpha locus in Atlantic
673 salmon contrast with those at the class I locus. *Immunogenetics* **57**(1)**:** 16-
674 24.
675
676 Consuegra S, Megens HJ, Schaschl H, Leon K, Stet RJM, Jordan WC
677 (2005b). Rapid Evolution of the MH Class I Locus Results in Different
678 Allelic Compositions in Recently Diverged Populations of Atlantic Salmon.
679 *Mol Biol Evol* **22**(4)**:** 1095-1106.
680
681 Coughlan J, McGinnity P, O'Farrell B, Dillane E, Diserud O, de Eyto E *et al*
682 (2006). Temporal variation in an immune response gene (MHC I) in
683 anadromous Salmo trutta in an Irish river before and during aquaculture
684 activities. *ICES J Mar Sci* **63**(7)**:** 1248-1255.
685
686 Crawford NG (2010). smogd: software for the measurement of genetic
687 diversity. *Mol Ecol Res* **10**(3)**:** 556-557.
688
689 de Eyto E, McGinnity P, Consuegra S, Coughlan J, Tufto J, Farrell K *et al*
690 (2007). Natural selection acts on Atlantic salmon major histocompatibility
691 (MH) variability in the wild. *Proc R Soc Biol Sci Ser B* **274**(1611)**:** 861-869.
692

693 Dhuyvetter H, Gaublomme E, Desender K (2004). Genetic differentiation
694 and local adaptation in the salt-marsh beetle *Pogonus chalceus*: a
695 comparison between allozyme and microsatellite loci. *Mol Ecol* **13**(5)**:**
696 1065-1074.
697
698 Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L (2007). Clinal
699 variation in mhc diversity with temperature: Evidence for the role of host-
700 pathogen interaction on local adaptation in Atlantic salmon. *Evolution*
701 **61**(9)**:** 2154 - 2164.
702
703 Dufresne F, Bourget E, Bernatchez L (2002). Differential patterns of spatial
704 divergence in microsatellite and allozyme alleles: further evidence for locus-
705 specific selection in the acorn barnacle, *Semibalanus balanoides*? *Mol Ecol*
706 **11**(1)**:** 113-123.
707
708 Edwards SV, Hedrick PW (1998). Evolution and ecology of MHC
709 molecules: from genomics to sexual selection. *Trends Ecol Evol* **13**(8)**:** 305-
710 311.
711
712 Eizaguirre C, Lenz T, Sommerfeld R, Harrod C, Kalbe M, Milinski M
713 (2010). Parasite diversity, patterns of MHC II variation and olfactory based
714 mate choice in diverging three-spined stickleback ecotypes. *Evol Ecol***:** 1-18.

715

716 Eizaguirre C, Yeates SE, Lenz TL, Kalbe M, Milinski M (2009). MHC-
717 based mate choice combines good genes and maintenance of MHC
718 polymorphism. *Mol Ecol* **18**(15)**:** 3316-3329.

719

720 Ekblom R, SÆTher SA, Jacobsson P, Fiske P, Sahlman T, Grahn M *et al*
721 (2007). Spatial pattern of MHC class II variation in the great snipe
722 (*Gallinago media*). *Mol Ecol* **16**(7)**:** 1439-1451.

723

724 Figueroa F, Gunther E, Klein J (1988). MHC polymorphism pre-dating
725 speciation. *Nature* **335**(6187)**:** 265-267.

726

727 Foll M, Gaggiotti O (2008). A Genome-Scan Method to Identify Selected
728 Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian
729 Perspective. *Genetics* **180**(2)**:** 977-993.

730

731 Ford MJ (2002). Applications of selective neutrality tests to molecular
732 ecology. *Mol Ecol* **11**(8)**:** 1245-1262.

733

734 Garrigan D, Hedrick PW (2003). Perspective: Detecting Adaptive Molecular
735 Polymorphism: Lessons from the MHC. *Evolution* **57**(8)**:** 1707-1722.

736

Go Y, Satta Y, Kawamoto Y, Rakotoarisoa G, Randrianjafy A, Koyama N *et al* (2003). Frequent segmental sequence exchanges and rapid gene duplication characterize the MHC class I genes in lemurs. *Immunogenetics* **55**(7)**:** 450-461.

Goldringer I, Bataillon T (2004). On the Distribution of Temporal Variations in Allele Frequency: Consequences for the Estimation of Effective Population Size and the Detection of Loci Undergoing Selection. *Genetics* **168**(1)**:** 563-568.

Goudet J (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *J Hered* **86**(6)**:** 485-486.

Grimholt U, Drablos F, Jorgensen SM, Hoyheim B, Stet RJ (2002). The major histocompatibility class I locus in Atlantic salmon (Salmo salar L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics* **54:** 570 - 581.

Grimholt U, Getahun A, Hermsen T, Stet RJM (2000). The major histocompatibility class II alpha chain in salmonid fishes. *Dev Comp Immunol* **24**(8)**:** 751-763.

759 Grimholt U, Hordvik I, Fosse VM, Olsaker I, Endresen C, Lie Ø (1993).
760 Molecular cloning of major histocompatibility complex class I cDNAs from
761 Atlantic salmon (*Salmo salar*). *Immunogenetics* **37**(6)**:** 469-473.
762
763 Grimholt U, Larsen S, Nordmo R, Midtlyng P, Kjoeglum S, Storset A *et al*
764 (2003). MHC polymorphism and disease resistance in Atlantic salmon
765 (*Salmo salar*); facing pathogens with single expressed major
766 histocompatibility class I and class II loci. *Immunogenetics* **55**(4)**:** 210 -
767 219.
768
769 Grommé M, Neefjes J (2002). Antigen degradation or presentation by MHC
770 class I molecules via classical and non-classical pathways. *Mol Immunol*
771 **39**(3-4)**:** 181-202.
772
773 Hambuch TM, Lacey EA (2002). Enhanced selection for MHC diversity in
774 social tuco-tucos. *Evolution* **56**(4)**:** 841-845.
775
776 Hansen MM, Nielsen EE, Mensberg KLD (1997). The problem of sampling
777 families rather than populations: relatedness among individuals in samples
778 of juvenile brown trout *Salmo trutta* L. *Mol Ecol* **6**(5)**:** 469-474.
779

780 Hansen MM, Skaala Ø, Jensen LF, Bekkevold D, Mensberg K-LD (2007).
781 Gene flow, effective population size and selection at major
782 histocompatibility complex genes: brown trout in the Hardanger Fjord,
783 Norway. *Mol Ecol* **16**(7)**:** 1413-1425.
784
785 Heath DD, Shrimpton JM, Hepburn RI, Jamieson SK, Brode SK, Docker
786 MF (2006). Population structure and divergence using microsatellite and
787 gene locus markers in Chinook salmon (Oncorhynchus tshawytscha)
788 populations. *Can J Fish Aquat Sci* **63**(6)**:** 1370–1383
789
790 Hedrick P, Kim T (2000). *Genetics of complex polymorphisms: parasites*
791 *and maintenance of MHC variation*. Harvard University Press: Cambridge,
792 MA, 204–234 pp.
793
794 Hedrick P, W. (2005). A standardized genetic differentiation measure
795 *Evolution* **59**(8)**:** 1633-1638.
796
797 Hedrick PW (1994). Evolutionary Genetics of the Major Histocompatibility
798 Complex. *Am Nat* **143**(6)**:** 945-964.
799

800 Hedrick PW, Parker KM, Lee RN (2001). Using microsatellite and MHC
801 variation to identify species, ESUs, and MUs in the endangered Sonoran
802 topminnow. *Mol Ecol* **10**(6)**:** 1399-1412.
803
804 Hedrick PW, Thompson G (1983). Evidence for balancing selection at
805 HLA. *Genetics* **104:** 449-456.
806
807 Hordvik I, Grimholt U, Fosse VM, Lie Ø, Endresen C (1993). Cloning and
808 sequence analysis of cDNAs encoding the MHC class II β chain in Atlantic
809 salmon (*Salmo salar*). *Immunogenetics* **37**(6)**:** 437-441.
810
811 Housset D, Malissen B (2003). What do TCR-pMHC crystal structures
812 teach us about MHC restriction and alloreactivity? *Trends Immunol* **24**(8)**:**
813 429-437.
814
815 Huang S-W, Yu H-T (2003). Genetic Variation of Microsatellite Loci in the
816 Major Histocompatibility Complex (MHC) Region in the Southeast Asian
817 House Mouse (Mus musculus castaneus). *Genetica* **119**(2)**:** 201-218.
818
819 Hughes AL, Nei M (1988). Pattern of nucleotide substitution at major
820 histocompatibility complex class I loci reveals overdominant selection.
821 *Nature* **335:** 167-170.

822

823 Hughes AL, Nei M (1989). Nucleotide substitution at major
824 histocompatibility complex class II loci: evidence for overdominant
825 selection. *Proc Natl Acad Sci U S A* **86**(3)**:** 958-962.

826

827 Hughes AL, Nei M (1990). Evolutionary relationships of class II MHC
828 genes in mammals. *Mol Biol Evol* **7:** 491.

829

830 Hughes AL, Yeager M (1998). Natural selection at the major
831 histocompatibility complex loci of vertebrates. *Annu Rev Genet* **32:** 415-
832 435.

833

834 Huseby ES, Crawford F, White J, Kappler J, Marrack P (2003). Negative
835 selection imparts peptide specificity to the mature T cell repertoire. *Proc*
836 *Natl Acad Sci U S A* **100**(20)**:** 11565-11570.

837

838 Jarvi SI, Tarr CL, Mcintosh CE, Atkinson CT, Fleischer RC (2004). Natural
839 selection of the major histocompatibility complex (*Mhc*) in Hawaiian
840 honeycreepers (Drepanidinae). *Mol Ecol* **13**(8)**:** 2157-2168.

841

842 Jeffery KJM, Bangham CRM (2000). Do infectious diseases drive MHC
843 diversity? *Microbes Infect* **2:** 1335-1341.

844

845 Jensen JL, Bohonak AJ, Kelley ST (2005). Isolation by distance, web
846 service. *BMC Genet* **6:** 13. v.13.16
847

848 Jordan WC, Bruford MW (1998). New perspectives on mate choice and the
849 MHC. *Heredity* **81**(2)**:** 127-133.
850

851 Jordan WC, Verspoor E, Youngson AF (1997). The effect of natural
852 selection on estimates of genetic divergence among populations of the
853 Atlantic salmon. *J Fish Biol* **51**(3)**:** 546-560.
854

855 Jost L (2008). G_{ST} and its relatives do not measure differentiation. *Mol Ecol*
856 **17**(18)**:** 4015-4026.
857

858 Karl SA, Avise JC (1992). Balancing selection at allozyme loci in oysters:
859 implications from nuclear RFLPs. *Science* **256**(5053)**:** 100-102.
860

861 Kaufman J, Milne S, Gobel TWF, Walker BA, Jacob JP, Auffray C *et al*
862 (1999). The chicken B locus is a minimal essential major histocompatibility
863 complex. *Nature* **401**(6756)**:** 923-925.
864

865 Kimura M (1985). *The Neutral Theory of Molecular Evolution*. Cambridge
866 University Press New York, 384pp.
867
868 King TL, Eackles MS, Letcher BH (2005). Microsatellite DNA markers for
869 the study of Atlantic salmon (*Salmo salar*) kinship, population structure,
870 and mixed-fishery analyses. *Mol Ecol Notes* **5**(1)**:** 130-132.
871
872 Klein J (1986). *Natural history of the major histocompatibility complex*.
873 John Wiley & Sons: New York, 775pp.
874
875 Klein J, O'HUigin C (1994). The conundrum of nonclassical major
876 histocompatibility complex genes. *Proc Natl Acad Sci U S A* **91**(14)**:** 6251-
877 6252.
878
879 Klein J, Sato A, Nikolaidis N (2007). MHC, TSP, and the Origin of Species:
880 From Immunogenetics to Evolutionary Genetics. *Annu Rev Genet* **41**(1)**:**
881 281-304.
882
883 Kruiswijk CP, Hermsen T, van Heerwaarden J, Dixon B, Savelkoul HFJ,
884 Stet RJM (2005). Major histocompatibility genes in the Lake Tana African
885 large barb species flock: evidence for complete partitioning of class II B, but

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

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*). *Mol Ecol* **10:** 2525-2539.

Landry C, Garant D, Duchesne P, Bernatchez L (2001). ˜Good genes as heterozygosity": the major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proc R Soc Biol Sci Ser B* **268**(1473)**:** 1279-1285.

Lawlor DA, Ward FF, Ennis PD, Jackson AP, Parham P (1988). HLA-A,-B polymorphisms predate the divergence of humans and chimpanzees. *Nature* **335:** 268.

Leberg PL (1992). Effects of Population Bottlenecks on Genetic Diversity as Measured by Allozyme Electrophoresis. *Evolution* **46**(2)**:** 477-494.

Leberg PL (2002). Estimating allelic richness: Effects of sample size and bottlenecks. *Mol Ecol* **11**(11)**:** 2445-2449.

908

909 Lukacs M, Harstad H, Bakke H, Beetz-Sargent M, McKinnel L, Lubieniecki
910 K *et al* (2010). Comprehensive analysis of MHC class I genes from the U-,
911 S-, and Z-lineages in Atlantic salmon. *BMC Genomics* **11**(1)**:** 154.
912

913 Markow T, Hedrick PW, Armstrong C, Martin J, Zuerlein K, Vyvial T *et al*
914 (1993). HLA polymorphism in the Havasupai: Evidence for balancing
915 selection. *Am J Hum Genet* **53**(4)**:** 943-952.
916

917 McConnell TJ, Talbot WS, McIndoe RA, Wakeland EK (1988). The origin
918 of MHC class II gene polymorphism within the genus Mus. *Nature*
919 **332**(6165)**:** 651-654.
920

921 McDonald JH (1991). Contrasting amounts of geographical variation as
922 evidence for direct selection: the Mpi and Pgm loci in eight crustacean
923 species. *Heredity* **67**(2)**:** 215-219.
924

925 Miller HC, Lambert DM (2004). Genetic drift outweighs balancing selection
926 in shaping post-bottleneck major histocompatibility complex variation in
927 New Zealand robins (Petroicidae). *Mol Ecol* **13**(12)**:** 3709-3721.
928

929 Miller KM, Kaukinen KH, Beacham TD, Withler ME (2001). Geographic
930 heterogeneity in natural selection on an MHC locus in sockeye salmon.
931 *Genetica* **111:** 237-257.
932
933 Miller KM, Withler RE, Beacham TD (1997). Molecular evolution at *Mhc*
934 genes in two populations of chinook salmon *Oncorhynchus tshawytscha*.
935 *Mol Ecol* **6**(10)**:** 937-954.
936
937 Nielsen EE, Hansen MM, Loeschcke V (1997). Analysis of microsatellite
938 DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison
939 of genetic composition over 60 years. *Mol Ecol* **6**(5)**:** 487-492.
940
941 Nielsen EE, Hansen MM, Loeschcke V (1999). Genetic Variation in Time
942 and Space: Microsatellite Analysis of Extinct and Extant Populations of
943 Atlantic Salmon. *Evolution* **53**(1)**:** 261-268.
944
945 Nielsen R (2005). Molecular signatures of natural selection. *Annu Rev Genet*
946 **39**(1)**:** 197-218.
947
948 O'Reilly P, Hamilton L, McConnell S, Wright J (1996). Rapid analysis of
949 genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of

950 dinucleotide and tetranucleotide microsatellites. *Can J Fish Aquat Sci* **53:**
951 2292-2298.
952
953 Oliver MK, Lambin X, Cornulier T, Piertney SB (2009). Spatio-temporal
954 variation in the strength and mode of selection acting on major
955 histocompatibility complex diversity in water vole (*Arvicola terrestris*)
956 metapopulations. *Mol Ecol* **18**(1)**:** 80-92.
957
958 Parham P, Ohta T (1996). Population Biology of Antigen Presentation by
959 MHC Class I Molecules. *Science* **272**(5258)**:** 67-74.
960
961 Parker KM, Ruby JS, Hedrick PW (1999). Molecular Variation and
962 Evolutionarily Significant Units in the Endangered Gila Topminnow.
963 *Conserv Biol* **13**(1)**:** 108-116.
964
965 Paterson S (1998). Evidence for balancing selection at the major
966 histocompatibility complex in a free-living ruminant. *J Hered* **89:** 289-294.
967
968 Paterson S, Piertney SB, Knox D, Gilbey J, Verspoor E (2004).
969 Characterization and PCR multiplexing of novel highly variable
970 tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Mol Ecol*
971 *Notes* **4**(2)**:** 160-162.

Pogson GH, Mesa KA, Boutilier RG (1995). Genetic Population Structure and Gene Flow in the Atlantic Cod Gadus morhua: A Comparison of Allozyme and Nuclear RFLP Loci. *Genetics* **139**(1)**:** 375-385.

Potts WK, Wakeland EK (1990). Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol* **5:** 181-187.

Potts WK, Wakeland EK (1993). The evolution of MHC genetic diversity: a tale of incest, pestilence, and sexual preference. *Trends Genet* **9:** 408-412.

Raymond M, Rousset F (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J Hered* **86**(3)**:** 248-249.

Reusch T, Langefors Å (2005). Inter- and Intralocus Recombination Drive MHC Class IIB Gene Diversification in a Teleost, the Three-Spined Stickleback <i>Gasterosteus aculeatus</i&gt. *J Mol Evol* **61**(4)**:** 531-541.

992 Reusch TBH, Aeschlimann PB, Haberli MA, Milinski M (2001). Female
993 sticklebacks count alleles in a strategy of sexual selection explaining MHC-
994 polymorphism. *Nature* **414:** 300 - 302.
995
996 Seddon JM, Baverstock PR (1999). Variation on islands: major
997 histocompatibility complex (*Mhc*) polymorphism in populations of the
998 Australian bush rat. *Mol Ecol* **8**(12)**:** 2071-2079.
999
1000 Seddon JM, Ellegren H (2004). A temporal analysis shows major
1001 histocompatibility complex loci in the Scandinavian wolf population are
1002 consistent with neutral evolution. *Proc R Soc Biol Sci Ser B* **271**(1554)**:**
1003 2283-2291.
1004
1005 Slettan A, Olsaker I, Lie Ø (1995). Atlantic salmon, Salmo salar,
1006 microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci.
1007 *Anim Genet* **26**(4)**:** 281-282.
1008
1009 Smulders MJM, Snoek LB, Booy G, Vosman B (2003). Complete loss of
1010 MHC genetic diversity in the Common Hamster (*Cricetus cricetus*)
1011 population in The Netherlands. Consequences for conservation strategies.
1012 *Conserv Genet* **4**(4)**:** 441-451.
1013

1014 Sommer S (2003). Effects of habitat fragmentation and changes of dispersal
1015 behaviour after a recent population decline on the genetic variability of
1016 noncoding and coding DNA of a monogamous Malagasy rodent. *Mol Ecol*
1017 **12**(10)**:** 2845-2851.
1018
1019 Spurgin LG, Richardson DS (2010). How pathogens drive genetic diversity:
1020 MHC, mechanisms and misunderstandings. *Proc R Soc Biol Sci Ser B*
1021 **277**(1684)**:** 979-988.
1022
1023 Stet R, de Vries B, Mudde K, Hermsen T, van Heerwaarden J, Shum B *et al*
1024 (2002). Unique haplotypes of co-segregating major histocompatibility class
1025 II A and class II B alleles in Atlantic salmon (*Salmo salar*) give rise to
1026 diverse class II genotypes. *Immunogenetics* **54**(5)**:** 320-331.
1027
1028 Valière N (2002). gimlet: a computer program for analysing genetic
1029 individual identification data. *Mol Ecol Notes* **2**(3)**:** 377-379.
1030
1031 Van Oosterhout C, Joyce DA, Cummings SM, Blais J, Barson NJ,
1032 Ramnarine IW *et al* (2006). Balancing selection, random genetic drift, and
1033 genetic variation at the major histocompatibility complex in two wild
1034 populations of guppies (*Poecilia reticulata*). *Evolution* **60**(12)**:** 2562-2574.
1035

1036 Vasemagi A, Nilsson J, Primmer CR (2005). Expressed Sequence Tag-
1037 Linked Microsatellites as a Source of Gene-Associated Polymorphisms for
1038 Detecting Signatures of Divergent Selection in Atlantic Salmon (*Salmo*
1039 *salar* L.). *Mol Biol Evol* **22**(4)**:** 1067-1076.
1040
1041 Vitalis R, Dawson K, Boursot P (2001). Interpretation of variation across
1042 marker loci as evidence of selection. *Genetics* **158**(4)**:** 1811 - 1823.
1043
1044 Wegner KM, Kalbe M, Milinski M, Reusch T (2008). Mortality selection
1045 during the 2003 European heat wave in three-spined sticklebacks: effects of
1046 parasites and MHC genotype. *BMC Evol Biol* **8**(1)**:** 124.
1047
1048 Wegner KM, Reusch TBH, Kalbe M (2003). Multiple parasites are driving
1049 major histocompatibility complex polymorphism in the wild. *J Evol Biol*
1050 **16**(2)**:** 224-232.
1051
1052 Westerdahl H, Hansson B, Bensch S, Hasselquist D (2004). Between-year
1053 variation of MHC allele frequencies in great reed warblers: selection or
1054 drift? *J Evol Biol* **17:** 485-492.
1055

Zhai W, Nielsen R, Slatkin M (2009). An Investigation of the Statistical Power of Neutrality Tests Based on Comparative and Population Genetic Data. *Mol Biol Evol* **26**(2)**:** 273-283.

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 (α=0.05/90=0.0006). River abbreviations: Burrishoole (BRH), Moy (MOY), Owenmore (OWM), Owenduff (OWD), Delphi (DPH), Owenwee (OWW) and Carrowinskey (CAW).

5

		BRH						OWD (N=57)	OWM (N=57)	MOY (N=48)
		1956 (N=50)	1968 (N=44)	1973 (N=48)	1983 (N=47)	1995 (N=40)	2002 (N=57)			
Ssa D144b	F_{IS}	**+0.245**	+0.277	**+0.170**	**+0.183**	+0.160	+0.028	+0.023	+0.081	+0.150
	P	**<0.0001**	0.0007	**<0.0001**	**<0.0001**	0.0179	0.0476	0.2994	0.0805	0.0417
Ssa 171	F_{IS}	+0.134	**+0.120**	-0.026	-0.117	+0.012	**+0.067**	+0.026	+0.089	**+0.299**
	P	0.1450	**<0.0001**	0.8374	0.4654	0.2267	**<0.0001**	0.2880	0.0512	**<0.0001**
Ssa 2215SP	F_{IS}	-0.026	-0.009	+0.076	+0.116	**+0.326**	+0.205	-0.001	+0.092	+0.130
	P	0.1413	0.0520	0.0655	0.0119	**<0.0001**	0.0014	0.1354	0.0621	0.0145
SsoSL 85	F_{IS}	**+0.275**	+0.189	+0.087	-0.049	+0.158	**+0.248**	+0.119	**+0.142**	**+0.407**
	P	**<0.0001**	0.0094	0.0160	0.2018	0.0142	**<0.0001**	0.1858	**<0.0001**	**<0.0001**

Ssa 197	F_{IS}	+0.113	+0.085	+0.162	-0.067	+0.002	+0.034	+0.123	+0.099	+0.090
	P	0.0010	0.0025	0.0324	0.4183	0.3111	0.0034	0.1167	0.1488	0.5526
Ssa 1G7SP	F_{IS}	+0.079	+0.259	+0.292	-0.101	+0.075	+0.025	**+0.293**	+0.128	+0.156
	P	0.0018	0.0211	0.0113	0.559	0.5648	0.0789	**<0.0001**	0.1692	0.0238
Ssa 202	F_{IS}	**+0.447**	+0.230	+0.251	+0.139	+0.158	+0.004	-0.027	+0.066	-0.025
	P	**<0.0001**	0.0010	0.0234	0.0023	0.1124	0.0298	0.4732	0.5981	0.6107
Ssa D170	F_{IS}	**+0.319**	**+0.112**	**+0.189**	**+0.213**	+0.086	-0.031	+0.137	+0.098	-0.040
	P	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	0.0334	0.3977	0.0243	0.3857	0.8574
Sasa-DAA-3UTR	F_{IS}	+0.082	+0.125	-0.044	+0.051	+0.148	-0.065	+0.203	+0.123	+0.062
	P	0.5158	0.4199	0.1343	0.8610	0.0873	0.3966	0.0013	0.4553	0.3670
Sasa-UBA-3UTR	F_{IS}	+0.052	-0.102	+0.161	**+0.297**	+0.028	-0.039	+0.111	**+0.182**	**+0.294**
	P	0.1460	0.2183	0.2725	**<0.0001**	0.0746	0.1530	0.0078	**<0.0001**	**<0.0001**

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 (α=0.05/60=0.0008). River abbreviations: Burrishoole (BRH), Moy (MOY), Owenmore (OWM), Owenduff (OWD), Delphi (DPH), Owenwee (OWW) and Carrowinskey (CAW).

5

		DPH				OWW (N=55)	CAW (N=48)
		1993 0+ (N=45)	1993 1+ (N=50)	2002 0+ (N=45)	2002 1+ (N=51)		
Ssa D144b	F_{IS}	-0.007	+0.114	+0.121	+0.157	**+0.031**	+0.029
	P	0.2089	0.0055	0.1098	0.0190	**>0.0001**	0.0545
Ssa 171	F_{IS}	+0.013	-0.012	-0.147	+0.028	+0.028	+0.051
	P	0.0460	0.5265	0.9571	0.2574	0.0062	0.0357
Ssa 2215SP	F_{IS}	-0.009	+0.070	+0.113	+0.074	**+0.281**	-0.032
	P	0.8879	0.3332	0.0044	0.2369	**0.0002**	0.1871
SsoSL 85	F_{IS}	+0.025	+0.082	+0.004	-0.003	+0.063	+0.183
	P	0.1160	0.0158	0.5275	0.5162	0.0066	0.0037

Ssa 197	F_{IS}	+0.012	+0.036	-0.049	-0.002	**+0.575**	-0.028
	P	0.5630	0.2662	0.4270	0.6161	**>0.0001**	0.7733
Ssa 1G7SP	F_{IS}	+0.069	+0.013	+0.067	**+0.172**	**+0.304**	+0.032
	P	0.7914	0.0505	0.3465	**>0.0001**	**0.0001**	0.0592
Ssa 202	F_{IS}	+0.176	**+0.365**	-0.040	+0.012	+0.271	**+0.210**
	P	0.0013	**>0.0001**	0.5938	0.1613	0.0033	**>0.0001**
Ssa D170	F_{IS}	**+0.279**	+0.130	**+0.059**	+0.129	**+0.240**	**+0.007**
	P	**>0.0001**	0.0036	**>0.0001**	0.0011	**>0.0001**	**>0.0001**
Sasa-DAA-3UTR	F_{IS}	+0.132	+0.303	+0.104	-0.005	+0.118	-0.223
	P	0.0038	0.0029	0.0674	0.6681	0.6304	0.0013
Sasa-UBA-3UTR	F_{IS}	+0.113	+0.114	+0.115	**+0.262**	**+0.313**	+0.027
	P	0.3261	0.0055	0.1193	**>0.0001**	**>0.0001**	0.0390

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 α=0.05 level.

		Temporal		Spatial	
		Burrishoole	Delphi	Burrishoole	Delphi
Heterozygosity (H_O)	microsatellites	$F_{5,42}$ = 1.24, P = 0.309	$F_{3,28}$ = 0.292, P = 0.786	$F_{3,28}$ = 0.98, P = 0.415	**$F_{2,21}$ = 3.80, P = 0.039**
	Sasa-DAA-3UTR	**G_5= 23.07, P <0.001**	**G_3= 8.17, P = 0.043**	G_3= 1.35, P = 0.715	G_2= 0.172, P = 0.917
	Sasa-UBA-3UTR	G_5= 3.595, P = 0.606	G_3= 0.719, P = 0.868	G_3= 3.06, P = 0.380	G_2= 3.07, P = 0.214
Allelic Richness (N_A)	microsatellites	$F_{5,42}$ = 0.95, P = 0.461	$F_{3,28}$ = 0.218, P = 0.883	$F_{3,28}$ = 0.71, P = 0.553	$F_{2,21}$ = 0.13, P = 0.876
	Sasa-DAA-3UTR	G_5= 1.59, P = 0.896	G_3=0.99, P = 0.306	G_3= 0.67, P = 0.879	G_2= 0.29, P = 0.866
	Sasa-UBA-3UTR	G_5= 1.61, P = 0.893	G_3= 0.61, P = 0.892	G_3= 4.48, P = 0.208	G_2= 0.18, P = 0.913

Figure 1. Location of the study rivers.

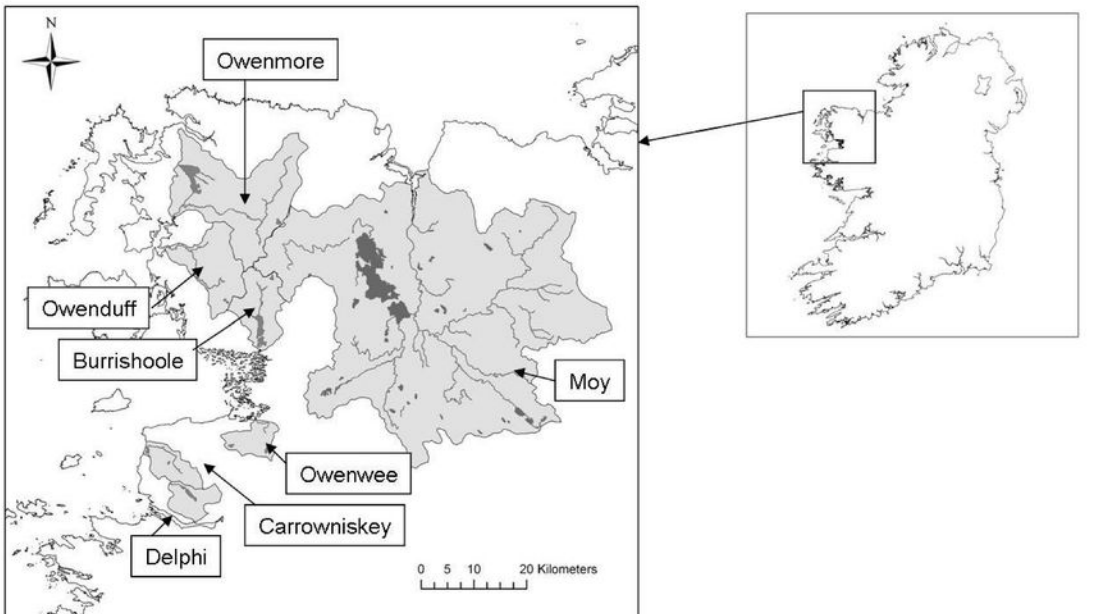

Figure 2. Heterozygosity at MHC and microsatellites.

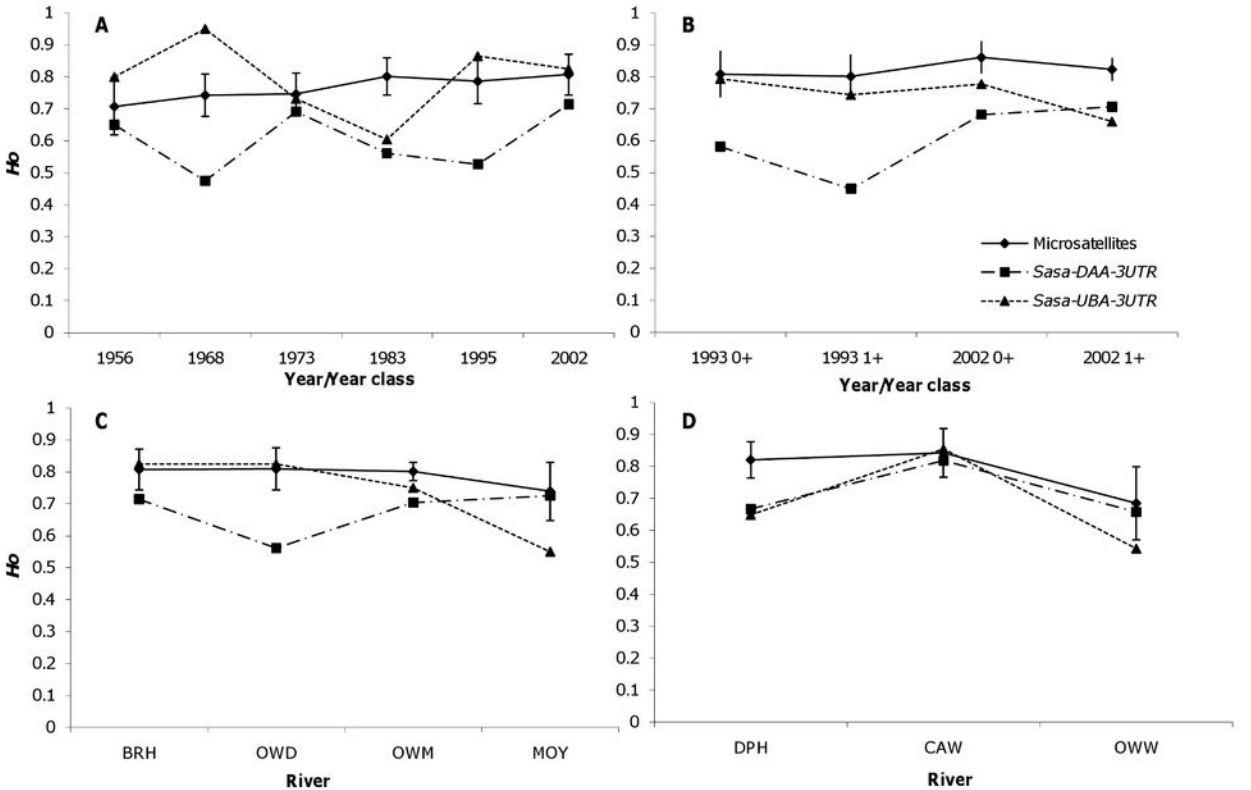

Figure 3. Allelic richness at MHC and microsatellites.

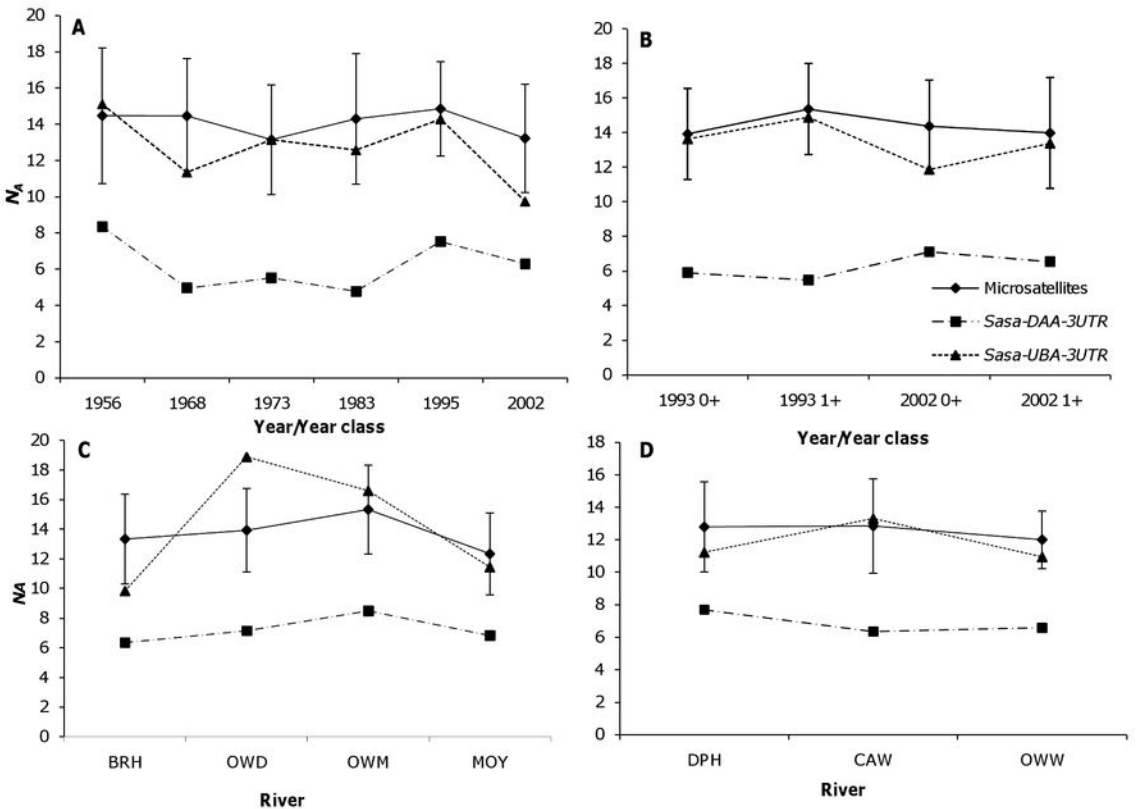

Figure 4. Results from bayesian analysis of selection.