MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 1730–1748 doi: 10.1111/mec.12684

Neutral and selective processes shape MHC gene diversity and expression in stocked brook charr populations (*Salvelinus fontinalis*)

FABIEN C. LAMAZE,* SCOTT A. PAVEY,* ERIC NORMANDEAU,* GABRIEL ROY,* DANY GARANT† and LOUIS BERNATCHEZ*

**Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, 1030 avenue de la Médecine, Québec, QC G1V 0A6, Canada, †Département de Biologie Faculté des Sciences, Université de Sherbrooke, 2500 Boul. de l'Université, Sherbrooke, QC J1K 2R1, Canada*

Abstract

The capacity of an individual to battle infection is an important fitness determinant in wild vertebrate populations. The major histocompatibility complex (MHC) genes are crucial for a host's adaptive immune system to detect pathogens. However, anthropogenic activities may disrupt natural cycles of co-evolution between hosts and pathogens. In this study, we investigated the dynamic sequence and expression variation of host parasite interactions in brook charr (*Salvelinus fontinalis*) in a context of past human disturbance via population supplementation from domestic individuals. To do so, we developed a new method to examine selection shaping MHC diversity within and between populations and found a complex interplay between neutral and selective processes that varied between lakes that were investigated. We provided evidence for a lower introgression rate of domestic alleles and found that parasite infection increased with domestic genomic background of individuals. We also documented an association between individual MHC alleles and parasite taxa. Finally, longer *cis*-regulatory minisatellites were positively correlated with MHC II down-regulation and domestic admixture, suggesting that inadvertent selection during domestication resulted in a lower immune response capacity, through a trade-off between growth and immunity, which explained the negative selection of domestic alleles at least under certain circumstances.

Keywords: 454 sequencing, brook charr, major histocompatibility complex, parasites, q-PCR

Received 14 October 2013; revision received 17 January 2014; accepted 22 January 2014

Introduction

An animal's ability to battle infection is one of the main attributes influencing fitness in wild vertebrate populations (Eizaguirre & Lenz 2010). As parasites can quickly develop adaptive resistance, they represent a moving threat for their host, but, as a result, hosts have evolved a battery of mechanisms to resist infection. The major histocompatibility complex (MHC) genes play a pivotal role in the interaction between pathogens and the host's adaptive immune system. MHC genes code for cell-surface glycoproteins that present peptides to the antigen receptor of T-cells, which then initiate the adaptive immune response in vertebrates (Klein 1986; Janeway *et al.* 2005). The MHC comprises the most polymorphic genes known in vertebrates (Robinson *et al.* 2009) and pathogen-driven balancing selection is one of the mechanisms maintaining MHC polymorphism (Bernatchez & Landry 2003; Spurgin & Richardson 2010). Three non-mutually exclusive theories have been proposed to explain the maintenance MHC diversity under pathogen-driven selection. At the individual level, overdominance means that heterozygotes are more fit than either

Correspondence: Fabien C. Lamaze, Fax: (1) 418 656 7176; E-mail: fabien.lamaze.1@ulaval.ca

homozygote as a wider range of antigens can be detected to initiate a specific immune response (Doherty & Zinkernagel 1975). At the population level, negative frequency dependent selection results in the selective advantage of resistant rare MHC alleles, after pathogens adapt to counter the more common alleles (Slade & McCallum 1992). At the metapopulation level, habitat heterogeneity promotes spatial variation in parasite selection regime driving diversifying selection of MHC alleles (Hedrick 2002).

Anthropogenic activities may disrupt natural cycles of co-evolution between hosts and pathogens (Poulin 2006; Pulkkinen *et al.* 2010). In particular, little attention has been paid to the threat of expanding global trade, animal transfer and hatcheries which increase the risk of pathogens and/or non-native pathogens encountered by wild populations (Harvell 1999; Krkošek *et al.* 2007; van Oosterhout *et al.* 2007). In parallel, the overexploitation of natural populations often leads to extensive stocking of domestic individuals as part of management strategies, raising concerns about the potential impacts of introgression on the evolutionary trajectories of natural populations (Rhymer & Simberloff 1996; Laikre *et al.* 2010). For example, the introgression of alleles underlying the expression of artificially selected traits may alter fitness-related traits in wild populations (Rhymer & Simberloff 1996; Allendorf *et al.* 2001; Edmands 2007; McClelland & Naish 2007), including immunogenic traits (Currens *et al.* 1997). Concerns about the consequences of hybridization have led to a number of complex issues that must be dealt with by conservation scientists (Allendorf *et al.* 2001; Edmands 2007; Fraser *et al.* 2008; Carroll 2011).

Domestication is predicted to impact selection shaping MHC diversity for a number of reasons: first, because domesticated individuals are often strongly selected for enhanced growth. Trade-offs between growth and the maintenance of an effective immune system exist, because the activation of the immune system is energetically costly and may expose cells to oxidative stress (Kurtz *et al.* 2006; van der Most *et al.* 2011). Second, stochastic events in inbred captive populations may result in a reduction in allelic richness for the entire genome including the MHC genes, which could have drastic consequences for survivorship in natural environment (van Oosterhout *et al.* 2007; but see Wynne *et al.* 2007). Third, relaxed natural and sexual selection on MHC genes may occur because highly effective disease control (e.g. medication) is applied on domesticated populations to avoid epizootic events and by artificially controlling mating (Ramstad *et al.* 2007; Neff *et al.* 2008; Plumb & Hanson 2011). Beside disease control, breeding programmes have recently aimed at enhancing disease resistance through genomically assisted programmes (Leeds *et al.* 2010; Villanueva *et al.* 2011). As a consequence, immunogenic genes have been shown to be differentially expressed between domesticated and wild populations in laboratory experiments after only three to seven generations of artificial selection, suggesting extremely strong selection in the hatchery environment resulting in rapid evolution of the immune system (Roberge *et al.* 2006; Vandersteen Tymchuk *et al.* 2010; Bougas *et al.* 2010).

Salmonids are of great interest for the study of MHC evolution due to their economic importance for aquaculture as well as their importance as a model group for studies of ecological divergence and local adaptation (Hendry & Stearns 2003). For instance, there is a vast amount of effort devoted to the understanding of disease control with association studies between MHC variation and several major salmonid diseases and the monitoring of the immune response following infection (Langefors *et al.* 2001; Lohm *et al.* 2002; Grimholt *et al.* 2003; Roberge *et al.* 2007; Croisetière *et al.* 2008). This has led to an accumulation of knowledge of molecular structure and expression regulation of the MHC genes in salmonids (e.g. Lukacs *et al.* 2007; Harstad *et al.* 2008; Croisetière *et al.* 2010). Also, analyses of MHC evolution are simplified as most of the salmonid species express a single locus of classical MHC classes I and II loci, which are not physically linked (Hansen *et al.* 1999; Grimholt *et al.* 2000; Stet *et al.* 2002; Conejeros *et al.* 2008; Croisetière *et al.* 2008; but see Pavey *et al.* 2013).

In contrast to the plethora of studies on MHC gene diversity, our current understanding of the MHC gene expression is almost exclusively derived from a handful of studies performed in the laboratory (e.g. Stet *et al.* 2002; Grimholt *et al.* 2002; Kurtz *et al.* 2006; but see Schwensow *et al.* 2011). There is also growing evidence of epigenetic regulations through a complex three-dimensional chromatin organization (Handunnetthi *et al.* 2010). Moreover, introns are known to be associated with diverse gene regulatory functions (e.g. Rose 2008). In higher eukaryotes, tandem repeated sequences such as microsatellite and minisatellite are frequently found in introns. The minisatellite extension or contraction may modulate the chromatin organization and is known to modulate MHC gene expression, including in brook charr (*Salvelinus fontinalis*; Croisetière *et al.* 2010).

Altogether, there is a need for a better understanding of host–parasite co-evolution in natural systems and by considering the modulating effects of biotic and abiotic factors (Evans & Neff 2009; Schwensow *et al.* 2011; Pavey *et al.* 2013). For example, it has been recently documented that temperature regime impacts the spatial distribution of pathogenic bacterial communities and MHC II alleles (Dionne *et al.* 2007, 2009). Also, there is a need for considering human impacts on

host–pathogen interaction as well as the population dynamics as relatively few studies have assessed the stocking impact on the immunogenic traits (Currens *et al.* 1997; van Oosterhout *et al.* 2007) compared with studies focussed on translocation and indirect effects of hatcheries (Coughlan *et al.* 2006; Krkošek *et al.* 2007; Evans *et al.* 2010b; de Eyto *et al.* 2011; Eizaguirre *et al.* 2012).

Brook charr is one of the most economically important species for freshwater aquaculture in eastern North America, largely based on producing fish to supplement re-creational fisheries (Ministère des Ressources Naturelles et de la Faune du Québec 2008). Previous studies on brook charr investigated the potential impact on genetic, transcriptional and physiological consequences of stocking practices on natural populations (Marie *et al.* 2010; Lamaze *et al.* 2012, 2013). Compared with other salmonids, there have been relatively few studies on MHC genes of brook charr (Croisetière *et al.* 2008, 2010; Pavey *et al.* 2011). Here, we used a holistic approach to investigate the dynamics of host–parasite interaction in a context of stocking. We pursued three main objectives: we first documented patterns of MHC variation in natural populations to assess the impact of stocking on MHC variation and assess the role of neutral vs. selective processes in shaping patterns observed in stocked lakes. Second, we tested for an association between individual MHC genotypes with parasite infection, controlling for the genetic background of individuals. Third, we documented the levels of expression at the two unlinked MHC Iα and IIβ genes to test whether they correlated with the extent of domestic admixture resulting from stocking in a natural context, while taking environmental, genetic, physiological and pathogenic variables into account. Finally, we conducted a laboratory experiment to confirm the cis-regulatory effect of a minisatellite length variation within the second intron of the MHC IIβ gene in order to take this variable into account in the natural context.

Materials and methods

Sampling

Sampling was conducted on nine different populations ($n = 336$) in the Portneuf wildlife Reserve in Québec, Canada (47°09′N, 72°17′W), in June 2007 and 2008 (for details see Table S1, Supporting information) divided in three stocking groups: heavy (HS: AMA, BEL, MET), moderate (MS: ARC, RIV, VEI) and never stocked (NS: CAR, MAI, SOR; see Marie *et al.* 2010 for the rationale underlying the separation of the groups and Lamaze *et al.* 2012 for details of sampling procedures). Two additional populations ($n = 90$) were sampled specifically for the detailed analysis of MHC gene expression in the Portneuf Wildlife Reserve (HS: MET9) and in the Mastigouche Wildlife Reserve (MS: BER9; 46°40′N, 73°30′W) in May and June 2009 (Table S1, Supporting information; see Lamaze *et al.* 2013 for details of sampling procedures). The adipose fins were sampled and preserved in 95% ethanol until DNA extraction. Fork length (cm), weight (g) (only in 2007–2008) and sex were recorded. In 2009, the anterior kidney tissue of each individual was dissected within one minute following euthanasia and preserved in liquid nitrogen for subsequent RNA analysis. Finally, adipose fins from 131 individuals from two reference domestic strains (JC, LDE) and a wild strain (BOU) used for stocking were sampled (Table S1, Supporting information; see Lamaze *et al.* 2012, 2013 for details). The Animal Care and Use Committee of Université Laval approved all protocols used when handling the fish.

Laboratory experiment on intronic length variation and gene expression

Twelve domestic individuals were used to produce six full-sib families that were homozygous for the minisatellite length variant in intron 2 of the MHC IIβ gene. Three homozygous families possessed either short (19–25) or long (46–59) repeats of a 32 nucleotide minisatellite motif characterized by Croisetière *et al.* (2010). The six families were reared at 12 °C for 1 year at the Laboratoire de Recherche en Sciences Aquatiques (LARSA, Université Laval, Canada) and fed *ad libitum*. As temperature is known to influence the level of expression of MHC genes (Bowden 2008; Croisetière *et al.* 2010), and in order to simulate important environmental variation, each family was split into two groups of 20 individuals and gradually acclimated for 1 month at either 8 or 21 °C. The head kidney was sampled within a minute of death and snap-frozen in liquid nitrogen for subsequent MHC IIβ gene expression analysis for 20 fish from each of the six families. Adipose fins were stored in 95% ethanol for subsequent genomic DNA extraction.

DNA and RNA extraction

Total DNA was extracted from 20 mg of adipose fins tissue using a salt extraction protocol (Aljanabi & Martinez 1997). DNA quality was evaluated by electrophoresis on a 1% agarose gel, and DNA concentrations and purity were determined using a Multiskan® Spectrum spectrophotometer (Thermo labsystems, Waltham, MA, USA). Total RNA was extracted from 20 mg of the head kidney to evaluate the expression of MHC Iα and IIβ, using the PureLink™ RNA mini kit protocol (Invitrogen, Burlington, VT) with DNase I treatment applied

directly on the membrane of the column according to the manufacturer's protocol.

Microsatellite, minisatellite and SNP data sets

We used the microsatellite data set of Lamaze *et al.* (2013) for individuals sampled in 2009 and from the three reference populations. A total of 23 microsatellites were used to estimates admixture of individuals in Lamaze *et al.* (2013), and these estimates are incorporated as an independent variable in gene expression analysis (see below). For the individuals sampled in 2007, 2008 and JC (Jacques Cartier Hatchery facility), we used 231 SNPs of Lamaze *et al.* (2012). This data set will be used to estimate and compare the population structure between SNPs and MHC (see below). Individuals sampled in 2009 in the wild and from the laboratory experiment were also genotyped for length variation at the minisatellite within MHC IIβ intron 2 with the P399 and P400 primers according to Croisetière *et al.* (2010). The final reaction volume was 25 μL using the GoTaq® Flexi DNA polymerase (Promega), as recommended by the manufacturer. This data set will be incorporated as independent variable within gene expression analysis (see below).

MHC sequencing

The amplicon library preparation followed a previously described method (Pavey *et al.* 2013) and followed the manufacturers' protocols for the GS FLX titanium series 454 (Roche). Briefly, this protocol involves four different steps: (i) amplicon preparation, (ii) purification, (iii) library construction and (iv) 454 sequencing. A total of 426 individuals were sequenced. Each genomic DNA sample (5–20 ng) from 2007, 2008, 2009 and JC was PCR-amplified using Platinum® Taq high-fidelity DNA polymerase (Invitrogen) following the manufacturer's instructions. The primers used for PCR amplification have been shown to efficiently amplify the complete β1 domain of the MHC IIβ gene for numerous species of salmonid fishes (Pavey *et al.* 2011). Further information about the procedure and quality control is available in the Supporting information/Appendix S6 section and from Pavey *et al.* (2013).

MHC genotyping

The number of MHC class II alleles in haplotype inference and genotyping of population samples with barcoded individuals were assessed according to Pavey *et al.* (2013). The input files are from the 454 output with the adaptors removed and split into one file per individual based on the barcode. Each file then goes through an iterative cleaning algorithm and alignment with MUSCLE (Edgar 2004). Briefly, a genotyping pipeline in both R and Bash languages used an iterative procedure in three successive steps. The first step generates putative allele sequences for each individual. The second step combines and strengthens the putative allele output of the first step into a single global alleles data set for all individuals. Then, the third step genotypes each individual. The pipeline as well as the current documentation and scripts are available at: https://github.com/enormandeau/ngs_genotyping. Further information about the procedure is available in the Supporting information Appendix S6 section and from Pavey *et al.* (2013).

cDNA synthesis and quantitative PCR

First-strand cDNA was synthesized from 5 μg of total RNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems), according to the manufacturer's instructions. To design qPCR primers, we selected MHC genes from *S. fontinalis* MHC class IIβ (EU478851.1), *Salmo salar* MHC class Iα (AF504019.1) and *Oncorhynchus mykiss* beta-actin (NM_001124235.1) as reference gene primer, and the sequences for specific primers and probes are reported in Table S2 (Supporting information). Amplicon lengths were between 60 and 80 base pairs. qPCR was performed with an ABI 7500 (Applied Biosystems) using TaqMan® Gene expression Master Mix (Applied Biosystems) in a reaction volume of 25 μL and a standard amplification protocol of 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of a denaturation step at 95 °C for 15 s and an annealing/elongation step at 60 °C for 1 min. Relative quantification of gene expression levels was performed by comparing differences in crossing threshold (C_T) values by the $\Delta\Delta C_T$ method, because the PCR efficiency was nearly equivalent for pairs of primers (Livak & Schmittgen 2001).

Signatures of selection in MHC sequences

To test for global selection at the β1 domain of MHC, we used MEGA v5.10 (Tamura *et al.* 2011). The best mutational model as well as evolutionary rate model among sites was investigated with maximum likelihood for nucleotide and amino acid substitution patterns. Nonsynonymous (d*n*) and synonymous (d*s*) ratio was investigated for the putative antigen binding sites (ABS) and non-ABS based on the human MHC IIβ crystal structure (Brown *et al.* 1993; Bondinas *et al.* 2007). We used a *Z*-test assuming positive selection (d*n* > d*s*) with the method of (Nei & Gojobori 1986) with a Jukes–Cantor correction for nucleotide substitution rate with Gamma

(Γ). We used a NeighborNet algorithm for phylogenetic β1 domain network representation, with a Whelan and Goldman (WAG + Γ) protein mutational model, implemented in SPLITTREE v4.12.6 (Huson & Bryant 2006). There is no three-dimensional structure information available for salmonid MHC IIβ protein. Therefore, confidence in identification of sites involved in peptide binding is currently very limited. To identify putative ABS, we first determined which codons were evolving under positive selection ($\omega > 1$). We used a mixed effects model of evolution (MEME) for detecting sites subject to episodic diversifying selection using the DATAMONKEY Web server (http://www.datamonkey.org/; Kosakovsky Pond 2005). The MEME model had superior performance over previous models under a broad range of scenarios (Murrell *et al.* 2012). Results were then compared with the codons corresponding to human ABS (Brown *et al.* 1993; Bondinas *et al.* 2007).

3D modelling

We generated the putative MHC IIβ three-dimensional structure on the I-TASSER Web server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/; Zhang 2008; Roy *et al.* 2010). The most common allele (allele 29, see Results) was chosen to predict the three-dimensional protein structure using an automated protein structure and function prediction consisting in four different stages, namely threading, structural assembly, model selection and refinement (Roy *et al.* 2010). The prediction accuracy was estimated through *Z*-score.

Genome scans

To identify potentially neutrally evolving SNPs (non-outliers) among all populations, we used the island simulation method implemented in ARLEQUIN v3.5 (Beaumont & Nichols 1996; Excoffier & Lischer 2010). Null distributions were generated assuming an island model consisting of 100 populations, and we based the tests on 1×10^5 simulations across loci as a function of scaled population heterozygosity. We applied a significance cut-off of $P < 0.05$ to exclude loci putatively under selection. The method used was deliberately chosen to be permissive as we were interested in better discriminating potentially neutrally evolving SNPs to compare with MHC.

Population allele frequencies

GENETIX v4.05.2 was used to calculate observed and expected heterozygosity, F_{IS} (Weir & Cockerham 1984), and allelic MHC IIβ frequencies (Belkhir *et al.* 1996–2004). Allelic richness standardized to the smallest population size (SOR, $n = 17$) was also calculated, using FSTAT (Goudet 1995). Differences between stocking groups were evaluated for expected heterozygosity and allelic richness using ANOVAs, with the statistical environment R version 3.0.1 (R Core Team 2013). A logarithmic transformation was applied on the expected heterozygosity. Conformity of MHC IIβ to Hardy–Weinberg equilibrium (HWE) expectations was tested using GENEPOP v4.2 (Raymond & Rousset 1995; Rousset 2008).

To study the evolutionary forces shaping patterns of MHC diversity, we developed a novel approach to compare MHC and neutral population differentiation. To do so, we selected 'potentially neutrally evolving SNPs' with a genome scan, that is, by excluding outlier SNPs, to build a neutral data set. Then, for each population, we interpreted the length of the vector that separates the MHC and neutral SNP data sets as the amount of departure from neutral expectation. The correlation between these two data sets was assessed with a co-inertia analysis (COIA; Dray *et al.* 2003) on the R package ade4 v1.5-1 (Dray *et al.* 2007). This method is very robust to autocorrelation and has three main advantages. First, complex variation is not reduced to a single number summary metric such as (F_{ST}), and variation is preserved as principle components before comparing the matrices in the COIA. Second, there is no need to designate paired comparisons as this method compares the correlation of matrices for all populations. Third, this method does not rely on any mutational and population assumptions (see Jombart *et al.* 2009). Two principal component analyses (PCA) were run on each data set constituted of alleles frequencies of the 10 populations sampled in 2007–2008. The significance of the relationships between the two matrices was assessed by comparing the COIA estimated from the real data set with the COIA distribution estimated after 10^4 bootstraps. This analysis will later refer as $COIA_{2G}$ in reference to the two genetic data sets (MHC and neutral SNPs).

Introgression

We were interested in the potential non-neutral nature of introgression of domestic MHC IIβ alleles into wild populations as predicted by the rare allele advantage hypothesis (Slade & McCallum 1992). We applied the genomic cline method (Gompert & Buerkle 2009) and the procedure of Zhu *et al.* (2005) to estimate clines for multi-allelic MHC IIβ data using genotypic frequencies as a function of genome-wide admixture and to test whether these homozygote or heterozygote clines were consistent with a null model of neutral introgression. Thus, we used a previously described procedure to differentiate two parental allele clusters into either

domestic or wild and a third cluster composed of hybrids (Lamaze *et al.* 2012). Only individuals caught in 2007–2008 in the stocked lakes were used in this analysis as the relatively high number of SNP markers ($n = 231$) should reduce clustering bias. The R package INTROGRESS v1.2.3 (Gompert & Buerkle 2010) was used to calculate the multinomial regression clines. Prior to regression fitting, the multi-allelic MHC IIβ data were reduced to two allelic classes that preserved the observed allele frequency differentials between parental species, without a loss of information or distorting the relationship between the parental populations (Zhu *et al.* 2005; Gompert & Buerkle 2009).

Parasitology

Stomach contents of fish sampled in 2008 were screened to initially survey the variation in local parasite abundance in 97 individuals within 10 populations from the two wildlife Reserves (Table S3, Supporting information). Individuals from the Portneuf wildlife reserve ($n = 53$) were also genotyped at MHC IIβ and at 231 SNPs. The complete digestive tract was screened for 90 individuals sampled in 2009. Eukaryote parasites were screened in each individual under a stereomicroscope with cold light illumination. The stomach (only for 2008), pyloric caeca and gut were inspected for the number of Nematoda, Cestoda, Trematoda and Acanthocephala parasites. Additionally, stomach and intestinal contents were screened separately for helminthes. Parasites were identified to the genus level. For a simple description of the differences in parasitic community between lakes, each taxon was analysed separately. We treated each parasite taxon as a dependent variable, and differences between lakes were determined using generalized linear models (GLM; with a quasi-Poisson error and a log-link function to compensate for overdispersion) with the statistical software R version 3.0.1 (R Core Team 2013).

MHC and pathogen covariation

We used a COIA for linking two matrices with the only constraint that rows have the same weight for each matrix (Dray *et al.* 2003). First, we did a coordinate analysis (CA) on the presence/absence of MHC IIβ alleles and a PCA on parasite abundance, using the row weights of the CA. Individuals were pooled across populations for 2009, assuming that individual allele–parasite covariation should hold across populations and to increase sample size. Relationships between the two matrices were assessed by comparing the COIA estimated from the real data set with the COIA distribution estimated after 10^4 bootstraps. This second COIA is referred below as $COIA_{GP}$ in reference to the two genetic and parasitic data sets used in this analysis. A GLM was used (quasi-Poisson error; log-link function) to assess the impact of introgression and the environment (lake) on the parasitic load in 2009. The proportion of domestic admixture of each individual was estimated with STRUCTURE ($K = 5$ as the addition of JC and LDE; see Lamaze *et al.* 2013 for justification and details). Parasitic load was standardized for the fish length (using a GLM) to obtain a comparable number of parasites per unit of length.

Gene expression analysis

Before testing the effects of the variables on gene expression, we assessed their colinearity (Zuur *et al.* 2009). For brook charr sampled in 2009, we first used a GLM to test whether any of the following seven independent variables had a significant impact on the MHC Iα and IIβ expression: admixture, minisatellite length, Fulton's condition index (see description below), parasitic load, sex, lake, capture method. The admixture level corresponded to the addition of the two domestic genetic background: JC and LDE estimated with STRUCTURE with $K = 5$ populations (Lamaze *et al.* 2013). A quadratic effect was added for the admixture as it increased the fit of the model (see Results) and also to test for nonadditive inheritance. The Fulton's condition index was used as a proxy for the general physiological state of the fish and was calculated as: $K = W/L^3 \times 100$, where W represents the mass (g) and L the fork length (mm) (Ricker 1975). We also included the length of the minisatellite present in the intron 2 of the MHC IIβ. The minisatellite length was calculated as the average length between the two allele copies within each fish assuming that they are codominantly expressed. The MHC IIβ expression data conformed to assumptions of normality and homoscedasticity, but the MHC Iα expression data did not. Therefore, the MHC Iα expression data were $\log^{10}$-transformed prior to incorporation in the linear model.

In a second step of confirmation of the gene expression and MHC/parasite co-evolution models and for a more integrative analysis, a multiple factor analysis (MFA) was conducted only for 2009, with the R package FACTOMINER (Lê *et al.* 2008). This method enables the analysis of complex interactions of multiple data sets (continuous or categorical) where variables are structured into groups. In this analysis, we included all the variables previously described in the gene expression analysis as well as the MHC IIβ alleles and parasites abundance for each taxon. This analysis was composed of seven groups of variables. The first variable group included gene expression level from the two MHC loci.

The second and third groups were the exon 2 and minisatellite length in the intron 2 genotypes that constituted the allelic information for the MHC IIβ gene. The fourth was the physiological group and contained the weight and the Fulton's condition index of fish. In the fifth group, we included the proportion of each five genetic background calculated with STRUCTURE with $K = 5$ (Lamaze *et al.* 2013), as the analysis could incorporate correlated variables. Then the individual admixture level was incorporated as a partition of each of the five genetic background variables namely: BER9, BOU, LDE, JC and MET9 (Table S1, Supporting information). In the sixth group, the number of each parasite taxa per individual was included. The seventh was a group of categorical variables constituted with the sex, lake and capture method. To represent independent variables used in the GLM, an eighth group of supplementary variable was created. This variables did not contributed to the analysis but could be visualized the correlation circle after computation (see Results). This group was composed of the domestic background at $K = 5$, the MHC IIβ minisatellite length average and the parasitic load. The continuous variables were centred to a mean of zero and a variance of one before being used in the analysis.

To analyse the MHC gene expression data collected from fish reared in the laboratory, we used an analysis of covariance (ANCOVA) to test whether the length of the minisatellite was inversely correlated with MHC IIβ gene expression and for an interaction with temperature. MHC IIβ gene expression was normalized prior to incorporation into the ANCOVA with a BoxCox transformation. The three independent variables tested were the minisatellite length, the temperature and Fulton index. Then, a least significant difference test with a Bonferroni correction ($\alpha = 0.05$) was applied as a post hoc test to compare distributions for significant variables.

Results

MHC variation

Of the 426 individuals sequenced at the MHC IIβ locus, 384 individuals (90%) were successfully genotyped. Individuals with less than 100× coverage ($n = 26$) and individuals harbouring a third allele ($n = 16$) after the genotyping pipeline were discarded. From the 27 reruns of individuals for quality control, one individual was inconsistent and therefore discarded, for a final 96% confidence in NGS sequencing and genotyping after the pipeline procedure. Additionally, among the 90 individuals (caught in 2009) for whom MHC IIβ was Sanger sequenced, homozygotes showed 100% sequence concordance with 454 sequencing. We found a 93.3% concordance value for genotypes after NGS genotyping pipeline with Sanger as 6.7% of Sanger heterozygotes were called as homozygotes with 454, indicating that our conservative approach to exclude artefacts was also resulting in missing some true alleles in a small percentage of individuals. We identified 29 unique 255 bp DNA sequences, which correspond to 29 unique amino acid sequences (Fig. S1, Supporting information). A codon insertion was found at position 58 in the amino acid sequence for three alleles, and a codon deletion was found at position 59 for five other alleles compared with the rest of sequences (Fig. S2, Supporting information). These 29 alleles were characterized by high degrees of nucleotide (83/255 sites segregating, $\pi = 0.166$) and amino acid (42/85 sites segregating, $\pi = 0.409$ average substitutions per site) polymorphism. Of the alleles identified, alleles 6 and 21 were previously described by Croisetière *et al.* (2008) as *Safo*-DAB*0401 (EU478854.1) and *Safo*-DAB*0301 (EU478853.1) with 100% coverage and e-values of 6×10^{-126} and 1×10^{-127}, respectively. The allele *Safo*-DAB*0301 was suggested as being resistant to *Aeromonas salmonicida*, whereas *Safo*-DAB*0401 was shown to increase susceptibility to infection in the experimental study of Croisetière *et al.* (2008). The *Safo*-DAB*0301 (resistant) allele was present in JC (2.4%), MET (HS) (1.5%), RIV (MS) (1.4%) with the highest frequency observed for two NS populations (20.6% and 58.3%, respectively, for SOR and MAI; Fig. S3, Supporting information). The *Safo*-DAB*0401 (susceptible) allele was present in HS and MS populations (BEL, MET, RIV, VEI) with the maximum observed in RIV (13.8%), and the lowest frequency was observed in JC (2.4%). This allele was never found in NS populations (Fig. S3, Supporting information).

Signatures of positive selection

The *Z* test gave an overall higher ω (dn/ds) for the residues aligning with the human ABS (24 codons; $\omega = 1.84$; $P = 0.034$) when compared to non-ABS (61 codons; $\omega = 1.77$; $P = 0.122$) and was statistically different from neutrality when estimated with a JC + Γ ($\Gamma = 0.172$). Using the MEME algorithm with the F81 codon site mutation model, after accounting for the presence of recombination (nucleotide position 147, cAIC improvement = 305.202), 17 of 85 sites with a $\omega > 1$ and statistically different from neutrality were detected. Among these, eight matched with the human ABS residues of the HLA-DQ/DR β1 domain (Brown *et al.* 1993; Bondinas *et al.* 2007). Also, two sites aligned with the homodimerization patch of the HLA-DQ/DR β1 domain. The positions 49 (corresponding to position

56 in HLA-DQ/DR β1 domain) and 53 (corresponding to position 60 in HLA-DQ/DR β1 domain) correspond to conserved sites among salmonids studied to date (Landry & Bernatchez 2001; Pavey *et al.* 2013). The five remaining residues under selection were located next to human ABS. These residues were mapped on the 3D protein model simulated, which was of high quality (*Z*-score = 2.24; coverage of the threading alignment = 95%; PDB hit = 3lqzB, HLA-DP1B; Fig. S4, Supporting information).

Genetic variation at the MHC IIβ and neutral markers

The genome scan of 231 SNPs for all sampled populations in 2007–2008 identified 25 and 41 SNPs that were significantly more and less divergent than neutral expectations. After this removal, 165 SNPs, potentially evolving under neutral expectations, were retained for further analyses.

Population allelic richness (*Ar*) at MHC IIβ ranged from 5.6 alleles in CAR, a nonstocked population, to 13.4 alleles in MET, one of the heavily stocked populations (Table S1, Supporting information). Stocking had a significant effect on the MHC IIβ *Ar* when pooling data from all sampling years in the Portneuf wildlife reserve ($F = 4.937$, d.f. = 2, $P = 0.046$) with a significantly higher *Ar* in HS lakes compared with the other lake groups (estimate ± SE = −4.285 ± 1.607, $t = 2.668$, $P = 0.032$; estimate ± SE = −4.258 ± 1.607, $t = 2.643$, $P = 0.033$; for the MS and NS comparison, respectively). The mean expected heterozygosity per population ranged from 0.679 to 0.861 (Table S1, Supporting information) and tended to increase proportionally to the stocking intensity ($F = 1.47$, d.f. = 2, $P = 0.293$): HS (0.793 ± 0.029) > MS (0.747 ± 0.058) > NS (0.679 ± 0.084).

In the neutral SNPs PCA, the first two axes explained 54.9% of the variance (Fig. 1A). In the MHC PCA, the first two axes explained 41.3% of the variance (Fig. 1B). From the two PCA analyses, the first axis separated the population mainly depending on their stocking status. This was also supported by previous STRUCTURE analyses (Marie *et al.* 2010; Lamaze *et al.* 2012). The second axis of the SNPs PCA mainly partitioned the variance between three populations MAI, CAR and ARC, whereas the second axis of the MHC PCA structured the populations in three groups (MAI, SOR; AMA, ARC, CAR, RIV, VEI; BEL, JC, MET).

The $COIA_{2G}$ analysis revealed a significant positive correlation between the SNP and MHC data sets with a RV = 0.81 ($P < 0.001$). The two data sets are represented on the same factorial map, where dots represent the potentially neutrally evolving SNPs and tip of arrows represent the MHC IIβ data (Fig. 1C). Eleven MHC IIβ alleles (1, 5, 8, 10, 11, 17, 22, 23, 25, 26, 27) mainly loaded on the first axis and contributed the most to the discrimination of HS and domestic *vs.* MS and NS populations ($F1 = 48.0\%$). Five MHC IIβ alleles (9, 18, 19, 21, 29) mainly loaded on the second axis and contributed to the discrimination of two NS (MAI and SOR) population from the rest of the populations ($F2 = 19.8\%$). The MS populations clustered with one NS population (CAR) and the two HS (AMA and MET) populations tended to be intermediate between the domestic and MS populations regarding either data sets. The length of the vector is inversely proportional to the correlation between the two data sets for a given population (Fig. 1C). Thus, the MAI (NS) and RIV (MS) populations possessed neutral SNPs and MHC IIβ allele frequencies that were highly correlated, whereas AMA (HS) showed the weakest correlation between both types of markers.

Introgression

We found a significant departure from neutrality in the introgression of the MHC IIβ gene ($P = 0.002$; Fig. 2A, B). In particular, a reduced introgression rate for domestic MHC IIβ alleles into the wild populations compared with the neutral SNP markers was observed. The homozygote cline fell within the 95% confidence interval of simulated neutral genomic clines (Fig. 2A), whereas the heterozygote cline suggested a pattern of under-dominance (Fig. 2B).

Parasite screening and MHC/parasite co-evolution

We found four distinct parasitic taxa within the digestive tract of the 187 individuals screened, namely *Echinorhynchus* sp. (phylum: Acanthocephala), *Eubothrium* sp. (class: Cestoda), *Sterliadochona* sp. (phylum: Nematoda) and *Crepidostomum* sp. (class: Trematoda; Table S4, Supporting information). Among the individuals screened, 97% were infected by Trematoda, 42% by Nematoda, 24% by Cestoda and 11% by Acanthocephala. Only two fish had no parasites. The GLM indicated that the level of infection varied significantly between lakes for all taxa ($P < 0.001$), albeit only marginally for Acanthocephala ($P = 0.059$). Nematoda and Trematoda had the highest mean infection number per individual and the highest variation between lakes, from 0.0 up to 324.6 and 50.7 up to 249.9 parasites per fish, respectively. In contrast, the other two taxa had a moderate and less variable infection per fish from 0.0 up to 3.1 Acanthocephala and from 0.02 up to 2.6 Cestoda.

The first two axes of the $COIA_{GP}$ accounted for 80.0% of the variation shared between the genetic and parasitic matrices (Fig. 3A), and no significant relationship

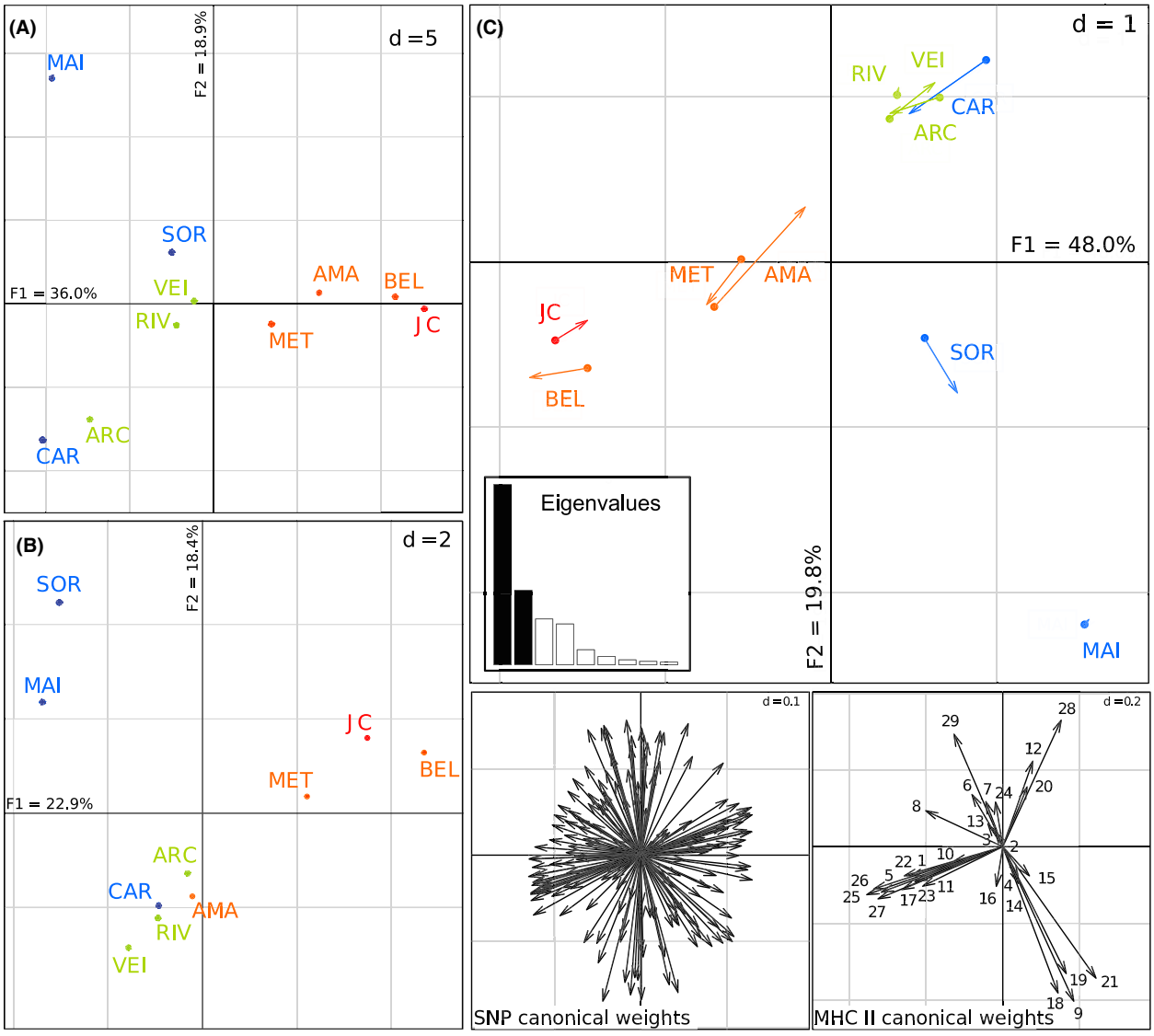

Fig. 1 Co-inertia analysis (COIA) between the frequencies of 165 potentially neutral SNPs and the major histocompatibility complex (MHC) IIβ loci for nine wild populations and one domestic population (A) Results of a principal component analyses on the SNPs and (B) on the MHC IIβ locus. Dots represent the respective relatedness of each population within the first two factorial plans. (C) The relative position of each population on the plan for the first two factors of the COIA is conditional to the MHC IIβ (the tip of the arrow) or the SNP genotypes (dot). The vector length and angle give the translational coefficient of the population position from one to the other data set. The strength of the correlation between the two data set for each population is inversely correlated with the length of the vector. The domestic population is in red (JC), the heavily stocked populations are in orange (AMA, BEL, MET), moderately stocked populations are in green (ARC, RIV, VEI), and three nonstocked populations are in blue (CAR, MAI, SOR). The COIA canonical weights for each allele for each of the two data sets are given. The length of vectors here refers to the amount of variance and covariance they explain in a particular plan direction of the COIA.

was found between them (global co-inertia RV coefficient = 0.099, simulated $P = 0.344$). Nevertheless, the $COIA_{GP}$ factor analysis indicated an association between certain parasites and the presence of specific MHC IIβ alleles (Fig. 3A). Eight alleles (1, 2, 6, 7, 13, 23, 25, 28) were negatively associated with the presence of all four parasitic taxa and therefore are potential resistant alleles, whereas five alleles (4, 10, 11, 15, 22) were strongly positively associated with the presence of at least one parasitic taxon. More specifically, allele 10 and 11 presented the strongest positive association with Cestoda, allele 15 with Trematoda, allele 22 with Nematoda and allele 4 with Acanthocephala (Fig. 3B). The MFA analysis generally confirmed the positive association between these alleles with the respective taxon (Fig. 4D). Also, allele 29 was strongly associated

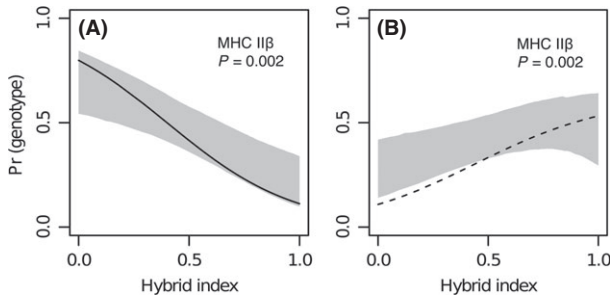

Fig. 2 Multinomial regression fit for the major histocompatibility complex IIβ gene contrasted with a genomic cline along an admixture gradient between the hatchery and wild brook charr in six stocked lakes. Plots depict the probability of domestic genotypes (1.0 = homozygous for domestic alleles; 0.5 = heterozygous; 0.0 = homozygous for wild alleles) as a function of the hybrid index, which quantifies the fraction of wild alleles across 231 SNP markers (see Lamaze *et al.* 2012 for analytical details). To facilitate interpretation, the results are split into two panels, with panel (A) displaying data for homozygotes and panel (B) displaying data for heterozygotes. The grey shading displays the 95% confidence intervals for a genomic cline given neutral introgression for the frequency of homozygotes (A) and heterozygotes (B) and the solid or dotted line illustrates the estimated cline based on the observed domestic homozygous or heterozygotes genotypes. The *P* value for the test of departure from neutrality is given.

with Acanthocephala (Fig. 4D). After standardizing for fish length, we found a positive correlation between the parasitic load and domestic strain introgression indicated by individual admixture value (coefficient ± SE: 0.579 ± 0.195, $P < 0.001$). Finally, lake had a significant effect on the parasitic load (coefficient ± SE: 0.412 ± 0.148, $P = 0.004$).

Gene expression analysis

In the wild populations, including a quadratic effect provided a better fit to the linear model (AIC = −7.614 *vs*. AIC = −4.245). Also, the level of domestic admixture had a significant negative effect on MHC IIβ gene expression (β = −0.831, $P = 0.031$; Table 1). This negative relationship was confirmed in the MFA analysis (Fig. 4A,B). However, there was no significant main effect of minisatellite length (β = 0.001, $P = 0.469$), Fulton's condition factor (β = −0.122, $P = 0.699$), parasitic load (β = 0.003, $P = 0.204$), sex (β = 0.019, $P = 0.705$), lake (β = 0.100, $P = 0.175$) or capture method [β = 0.004 (gill-net) and β = −0.148 (trap-net), $P = 0.082$] on MHC IIβ gene expression (Table 1). Our results also revealed that domestic admixture (β = −0.129, $P = 0.916$), as well as lake (β = 0.209, $P = 0.384$), capture method [β = 0.266 (gill-net) and β = 0.385 (trap-net), $P = 0.318$], or sex (β = 0.012, $P = 0.942$) had no effect on the MHC Iα expression (Table 2). However, an interaction between Fulton's condition factor and parasitic load affected the MHC Iα expression (β = 0.186, $P = 0.007$; Table 2). The Fulton's condition factor positively affected the slope of the MHC Iα gene expression for higher parasitic load. Similarly, the higher parasitic load, the greater (more positive) the effect of Fulton's condition factor on the level of MHC Iα gene expression.

The first axes of the MFA accounted for 15.9% of the variance and mainly discriminated among the different strains of brook charr and parasites taxa (Fig. 4A). The MFA provided visual support for the negative correlation between domestic admixture and MHC IIβ expression as well as MHC Iα expression (Fig. 4A,B). These three variables mainly loaded on the second axis ($F2 = 8.7\%$). The MFA provided support for a negative correlation between minisatellites length and MHC IIβ expression on the second axis (Fig. 4A,B). Also, longer minisatellite vectors mainly loaded in the same direction as the domestic backgrounds (Fig. 4A,B). Finally, a positive correlation between Cestoda prevalence and MHC Iα and IIβ genes expression was observed (Fig. 4A,C).

Similarly to the pattern observed in the wild, we confirmed in the laboratory experiment as previously reported (Croisetière *et al.* 2010) that the minisatellites length in the second intron of MHC IIβ had a positive and significant main effect on MHC IIβ expression (β = 0.313; $P < 0.001$), while controlling for other sources of variation (Table 3). Also, temperature (β = 0.344; $P < 0.001$) and Fulton's condition index (β = 0.402; $P = 0.005$) had a significant effect on MHC IIβ expression (Table 3). Finally, we found that the minisatellites length interacted with the temperature on the MHC IIβ expression (β = −0.301; $P < 0.001$; Fig. S5, Supporting information; Table 3). The temperature negatively affects the origin of the slope of the MHC IIβ gene expression for longer minisatellites length.

Discussion

This study improves our understanding of the functional impacts of stocking on immunity by using a multifaceted approach with stocked and un-stocked populations of brook charr. We reported that the MHC IIβ allelic variation observed in populations is a combination of a complex interplay between introgressive, neutral and selective processes. Also, we showed that domestic alleles have a lower introgression rate, hypothetically as a result of negative selection. Furthermore, not only individuals with domestic background were more infected, but we also found an association between individual MHC IIβ alleles and parasite species, suggesting a possible local adaptation of the allelic pool to specific parasites, which was further con-

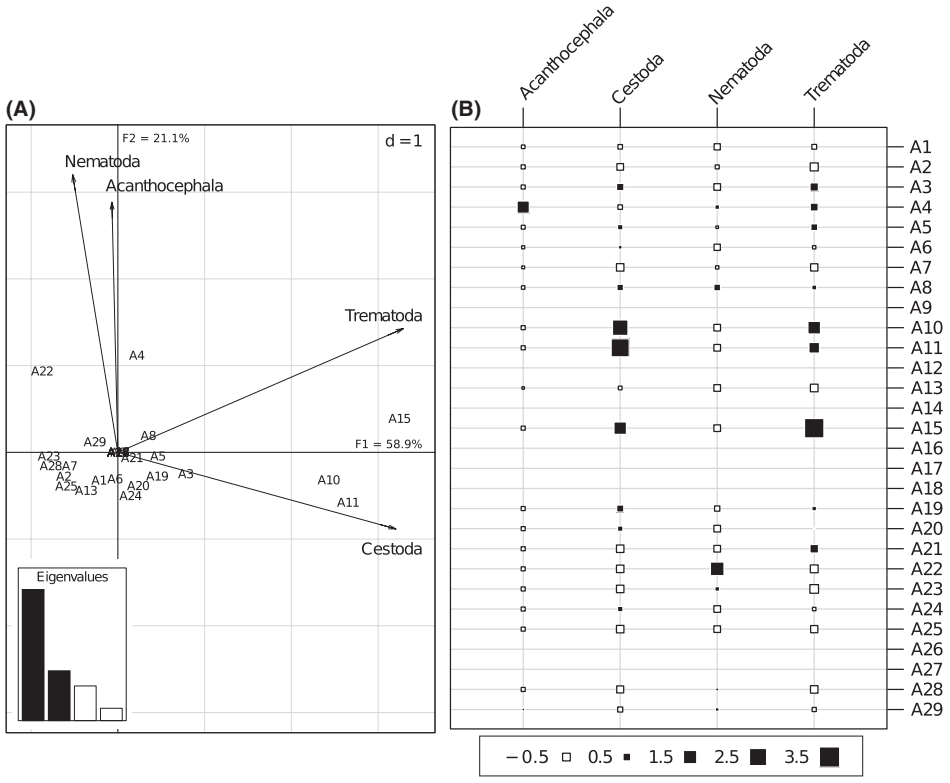

Fig. 3 Projection of a co-inertia analysis (COIA) between parasitological infection intensity and genetic variation at major histocompatibility complex (MHC) IIβ. A) Parasitological and genetic COIA analysis projection on the first two factor plans. Vectors represent the parasite data, whereas MHC IIβ alleles (A−) are represented by a location on the plan. Variables (parasite content and MHC IIβ alleles) in the same direction or close together are positively associated, whereas those located in the opposite direction are considered as negatively associated. MHC IIβ alleles located in the centre do not structure the data. (B) Indicated covariance of factors of the COIA analyses illustrated by square sizes. The size of the squares is proportional to the sum of square covariance between the gastrointestinal parasite count and MHC IIβ allele variation. Black squares indicate positive and white squares negative associations.

firmed by the allele segregation on the $COIA_{2G}$ map. Finally, we found evidences for a functional link between artificial selection for longer *cis*-regulatory minisatellite and the down-regulation of MHC Iα and IIβ genes in domestic individuals, suggesting lower performance against pathogens in nature.

The domestic strain had a higher allelic richness at the MHC IIβ gene compared with the nonstocked populations despite the expectation that drift, founder effects and artificial selection during domestication should have lowered the genetic diversity (e.g. van Oosterhout *et al*. 2007). This somewhat counterintuitive result could be partly explained by the fact that domestic brook charr in Québec represent a meta-population where brood stocks management routinely involves fish exchange between different hatcheries. Thus, a high genetic diversity may be maintained by generating artificial gene flow between different hatcheries, which reduces genetic drift (Marie *et al.* 2010; Lamaze *et al*. 2012). The increased allelic richness and reduced genetic divergence between populations at the MHC IIβ gene, as a function of stocking intensity, is concordant with previous results obtained using microsatellites and SNPs analyses (Marie *et al.* 2010; Lamaze *et al.* 2012). Thus, stocking has three different effects on MHC IIβ diversity: (i) homogenizing the genetic structure of MHC IIβ among wild populations, (ii) increasing the intrapopulation genetic diversity because of (iii) the introduction of new potentially maladaptive domestic MHC IIβ alleles into wild populations, including the *A. salmonicida* susceptible allele 6 (*safo*-DAB*0401), which is not otherwise found in wild brook charr.

The new $COIA_{2G}$ method we used gives insights into how selection or drift operates within population, which is an improvement over previous analyses (e.g. Bernatchez & Landry 2003). First, the strong correlation observed between neutrally evolving SNPs and MHC IIβ in MAI (NS) and RIV (MS) [i.e. the two types of markers cosegregate on the same location on the graph (short vector length)], suggests that neutral processes likely shaped MHC IIβ allelic frequency in these populations. One possible explanation is that MAI and RIV populations are small and isolated (Marie *et al.* 2010), meaning that the effect of genetic drift on MHC could

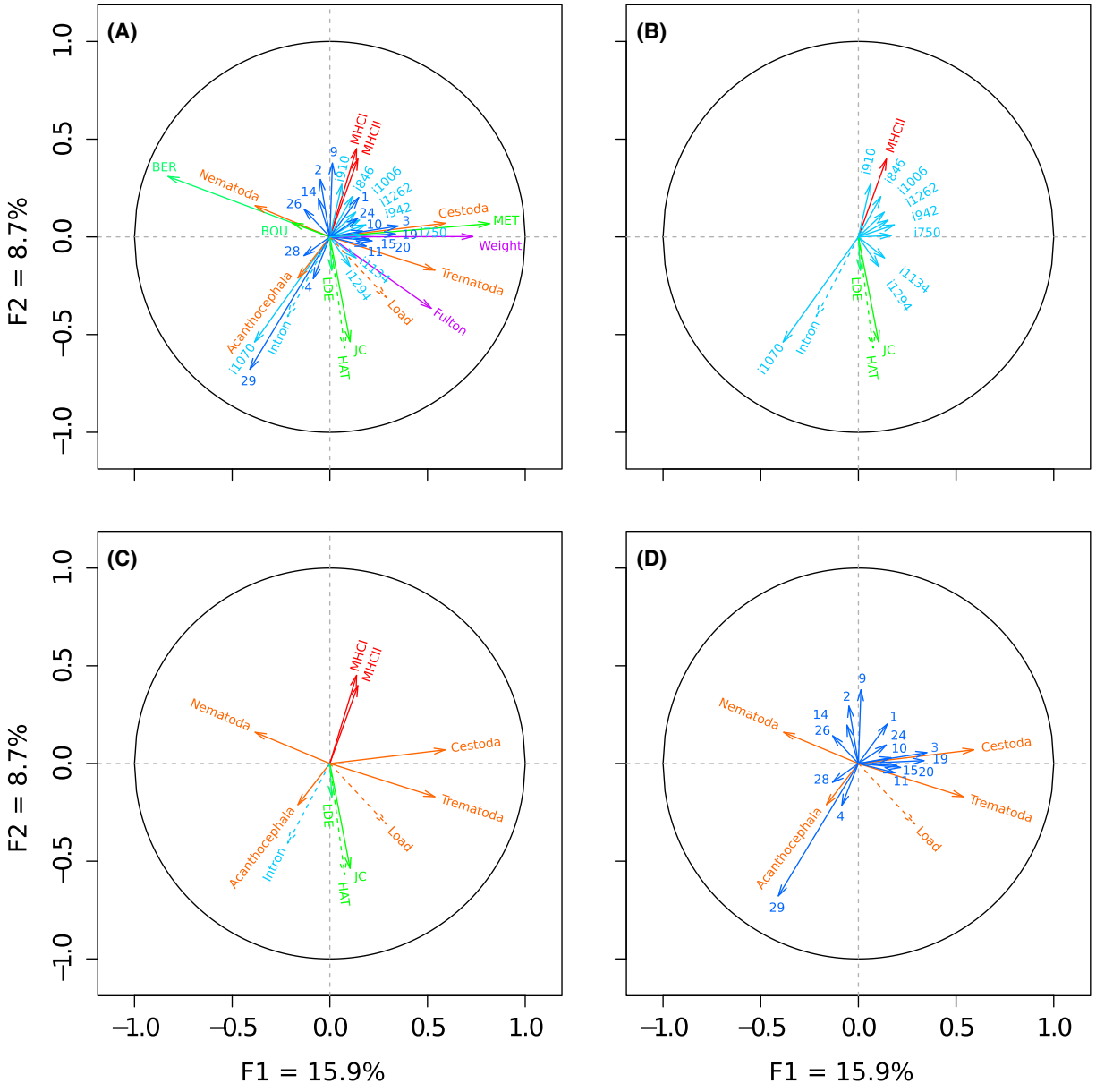

Fig. 4 Multiple factor analysis correlation circle of continuous variables. Colours are used for visual clarity as follow: green = genetic background variables, orange = parasites variables, light blue = minisatellite length in intron two of major histocompatibility complex (MHC) IIβ, dark blue = the alleles of MH IIβ, red = the expression of MHC IIβ and MHC Iα, purple = physiological variables: weight and Fulton's condition index. Categorical variables are not shown (lakes; capture; sex). Dashed vectors representing supplementary variables that are not used in the computation of the MFA. 'Parasitic load' is the total number of parasite per individual irrespective of taxa and corrected for fish body length. 'Domestic admixture' is the addition of hatcheries genetic background (LDE and JC) calculated with STRUCTURE for $K = 5$. Variables located close to the centre do not structure the data, and vectors having a length shorter than the square cosine of 0.02 are not drawn. (A) Represents all variables, (B) represents only the minisatellite length, the domestic background and MHC genes expression, (C) excludes all MHC alleles and minisatellite length, and (D) includes MHC IIβ alleles and parasites taxa and load.

be very strong and reduce the efficiency of selection (Landry & Bernatchez 2001; Evans *et al.* 2010a; Sutton *et al.* 2011). Second, the $COIA_{2G}$ suggested that MHC IIβ has also evolved under strong directional selection in AMA (HS) and CAR (NS), which showed a low correlation (long vector length) between MHC IIβ and neutral SNPs. The vector direction away from the domestic MHC IIβ alleles and the presence of the second greatest Nematoda infection in AMA in 2008 suggested that domestic migrant MHC IIβ alleles were

Table 1 Results for the major histocompatibility complex IIβ gene expression in natural environment obtained from the linear model

Variables	β ± SE	*F*	*P* values
Introgression	0.543 ± 0.315	2.421	0.117
Introgression^2	−0.831 ± 0.377	4.867	0.031
Minisatellite length	0.001 ± 0.002	0.531	0.469
Parasitic load	0.003 ± 0.002	1.642	0.204
Fulton	−0.122 ± 0.315	0.150	0.699
Lake	0.100 ± 0.073	1.873	0.175
Capture		2.593	0.082
Gill-nets	0.004 ± 0.062		
Trap-nets	−0.148 ± 0.082		
Sex	0.020 ± 0.052	0.145	0.705

The introgression parameter was estimated with $K = 5$ from STRUCTURE. β ± SE stands for the coefficient ± standard error. Significance (*P* values) for the fixed effects was obtained through a type III test (*F* statistic).

Table 2 Results for the major histocompatibility complex Iα gene expression in natural environment obtained from the linear model

Variables	β ± SE	*F*	*P* values
Introgression	0.018 ± 1.163	<0.001	0.988
Introgression^2	−0.129 ± 1.217	0.011	0.916
Parasitic load	−0.200 ± 0.073	7.615	0.007
Fulton	−0.846 ± 1.039	0.663	0.418
Lake	0.209 ± 0.239	0.767	0.384
Capture		1.162	0.318
Gill-nets	0.266 ± 0.204		
Trap-nets	0.385 ± 0.275		
Sex	0.012 ± 0.168	0.005	0.942
Load × Fulton	0.186 ± 0.068	7.539	0.007

The introgression parameter was estimated with $K = 5$ from STRUCTURE. β ± SE stands for the coefficient ± standard error. Significance (*P* values) for the fixed effects was obtained through a type III test (*F* statistic).

particularly maladapted in this population. The four populations highlighted above represent the two extremes of the spectrum in the comparison to neutral expectations. The remaining populations with vectors of intermediate length are intermediate between the two extremes.

We found a positive correlation between parasitic load and the extent of domestic admixture. A previous study, using experimental stocking, documented that domestic guppies (*Poecilia reticulata*) were more susceptible to gyrodactylid parasites (58% survival rate) than their wild counterparts (96% survival; van Oosterhout *et al.* 2007). In addition, introgression of domestic rainbow trout increased the susceptibility to ceratomyxosis infection (by 1.96 up to 4.25 times) in a wild population (Currens *et al.* 1997). Although we do not have any information about the parasite community in hatchery, literature abounds in protocols aiming at controlling epizootic events with medication (e.g. Ramstad *et al.* 2007). Moreover, domestic brook charr is known for disease susceptibility, for example furunculosis f(Cipriano *et al.* 2002; Bougas *et al.* 2010). Thus, any alteration in immune genes variation or recognition capability of domestic MHC alleles and lack of exposure to natural parasite communities may have reduced the fitness of captive-bred individuals in the natural environment (van Oosterhout *et al.* 2007).

Table 3 Results for major histocompatibility complex IIβ gene expression regulation in experimental condition obtained from an ANCOVA

Variables	β ± SE	*F* values	*P* values
Minisatellite length	0.313 ± 0.066	22.257	<0.001
Temperature	0.344 ± 0.063	29.661	<0.001
Fulton	0.402 ± 0.141	8.161	0.005
Minisatellite length × temperature	−0.301 ± 0.090	11.220	<0.001

Significance (*P* values) for the fixed effects was obtained through a type III test (*F* statistic). β ± SE stands for the coefficient ± standard error. Reference βs (or intercepts) for minisatellite length = short repeat, for temperature = 21 °C, as given by contrasts.

We described the same species of helminthes as previously found in other lakes close to our sampling sites (Bertrand *et al.* 2008). We found that this community varied significantly in abundance between populations, and this observation is in accordance with many fish studies including in salmonids (Dionne *et al.* 2007; Evans & Neff 2009; Eizaguirre *et al.* 2012). Also, canonical weights associated with MHC IIβ alleles revealed a clear segregation of alleles that contributed to the population structure observed on the $COIA_{2G}$ map. This suggests that divergent host–parasite interactions in different environments are predicted to result in locally adapted MHC allele pools as well as local adaptation of parasite infectivity and virulence (Kaltz & Shykoff 1998; Hedrick 2002; Kawecki & Ebert 2004; Eizaguirre *et al.* 2012). For example, in stickleback (*Gasterosteus aculeatus*), the lacustrine form harboured a higher number of alleles that were more adapted to a highly diverse parasite environment compared with the riverine form, which harboured fewer but more specialized alleles to the local parasitic community (Eizaguirre *et al.* 2012). In Chinook salmon (*Oncorhynchus tshawytscha*), Evans *et al.* (2010b) found evidence of increased fitness associated with lower MHC II allelic diversity in one population, whereas translocated families did not display this pat-

tern, which the authors interpreted as evidence for some degree of local adaptation at the MHC loci.

Quantitative positive and negative covariation between MHC IIβ alleles and helminthes was observed with the $COIA_{GP}$, as theoretically predicted and observed (Gandon & Michalakis 2000; Westerdahl *et al.* 2012), suggesting rare allele advantage or fluctuating selection promoting local adaptation (Spurgin & Richardson 2010). Similar to other studies, we found 'susceptibility' alleles, that is, alleles associated with specific pathogens (Croisetière *et al.* 2008; Dionne *et al.* 2009; Evans & Neff 2009; Pavey *et al.* 2013). Also, susceptibility alleles may represent quantitative resistance from severe infections (see, Westerdahl *et al.* 2012), as individuals with very poor alleles may die from an infection before sampling occurs. The quantitative resistance alleles may allow an infected individual to survive with a parasite by keeping it in check, allowing the individual to be sampled and the allele to be associated with the pathogen. All studies in natural systems might face this problem (see Discussion in Pavey *et al.* 2013). Interestingly, all the strong 'susceptibility' alleles (4, 10, 11, 15, 22), while positively associated with one or more parasite species, were also negatively associated with other pathogens species. This observation suggests antagonistic effects of MHC IIβ alleles in pathogen resistance, as previously observed in Loiseau *et al.* (2008), potentially promoting the maintenance of 'susceptibility' alleles in the population at least in certain conditions. A group of alleles (1, 2, 6, 7, 13, 23, 25, 28), loaded in the opposite direction of all helminthes, would suggest that they represent universally 'resistant' alleles. These alleles could also be associated with other, unmeasured pathogens, for example allele 6, while susceptible to *A. salmonicida* seems 'resistant' to helminthes.

The cline analysis indicated that domestic MHC IIβ alleles introgressed less than predicted, based on the neutral expectation, within wild populations. Namely, a reduction in heterozygote in hybrids may suggest that hybrids were less fit, then contributing to reduce the rate of introgression. Strengthening the fitness reduction of immune function associated with introgression, we recently documented for the same individuals a lower introgression rate of PSMB9 (proteasome subunit beta 9), a gene involved in immunity and located in the MHC I 'core' region (Lukacs *et al.* 2007; Lamaze *et al.* 2012). These results further suggested that local adaptation may have lowered the level of admixture of immune genes in these populations. Under the hypothesis that protein-coding and *cis*-regulatory polymorphisms in MHC IIβ gene are adaptive, reproduction with domesticated fish could lead to outbreeding depression of wild, stocked populations by reducing fitness of hybrid offspring via changes in MHC protein quality and/or quantity.

Our extensive sequencing coverage revealed only one to two MHC IIβ alleles per individual, confirming the hypothesis that MHC IIβ is not duplicated in brook charr, as in many salmonids except for *Coregonus* which carries two loci (Croisetière *et al.* 2008). The presence of a 3-bp indel trans-specific polymorphism in exon 2 of the MHC IIβ gene in multiple salmonid species could be the result of balancing selection (see Discussion in Pavey *et al.* 2013). This indel could play a role in presenting antigens as an MHC IIβ allele from brook charr (*safo*DAB*0401) was characterized by a 6-bp deletion and is one of the most susceptible to *A. salmonicida* (Croisetière *et al.* 2008). We found 17 nucleotides in the MHC IIβ exon 2 under positive selection, and about half (eight) were congruent with human ABS, as reported in other salmonid studies (Brown *et al.* 1993; Bondinas *et al.* 2007). However, the structural function of the indel and the homodimerization patch in the final protein is not yet known and needs further investigation.

Introgression due to stocking affected the regulation of MHC Iα and IIβ genes as suggested by the negative correlation of expression with the domestic admixture level, a similar pattern to those recently reported for growth and energy metabolism genes (Lamaze *et al.* 2013). Moreover, this result is concordant with the under-expression of MHC genes in domestic relative to wild brook charr individuals (Bougas *et al.* 2010). Furthermore, a *cis*-regulatory minisatellite located in the 2nd intron of the MHC IIβ in brook charr revealed a genetic element acting on the regulation of MHC IIβ expression (Croisetière *et al.* 2010). Supporting the fact that a genetic regulatory component may have been selected during artificial selection, the MFA suggested that the longest minisatellites were generally associated with a domestic genetic background. Additionally, as reported by Croisetière *et al.* (2010), we found both in the wild and in controlled condition that the number of this minisatellite repeat is inversely correlated with the level of expression of the MHC IIβ. Also, we found a quadratic effect, underlying dominance rather than additive effect on pattern of gene expression. Admittedly, transcriptional deregulation is complex and unpredictable (Bougas *et al.* 2010). Nevertheless, our results tend to support a predominant role of negative selection acting in *F*1 and post-*F*1 individuals against *cis* longer regulatory minisatellite selected during domestication. Also, additional sources of environmental variation such as temperature may modulate the regulation of gene expression as suggested by Croisetière *et al.* (2010) and our results in controlled conditions. This was also previously suggested for five of

seven candidate genes involved in growth, energy, protein folding and detoxification (Lamaze *et al.* 2013).

The predominance of alleles bearing long minisatellites and the positive association between longer minisatellites and reduced transcript abundance for both MHC genes could be the result of strong artificial selection for fast growth. Such selection process often comes at the cost of inadvertently down-regulating the immune system, as predicted by the resource-dependent cost of the evolution of immune resistance and in agreement with observations in livestock (Rauw *et al.* 1998; Schulenburg *et al.* 2009; Boots 2011). In addition, a trade-off between growth and immune function has been suggested in other salmonids by comparing different types of fast-growing strains and their wild counterparts (Tymchuk *et al.* 2009). Moreover, the disease susceptibility was argued to be partly the result of a down-regulation of MHC genes (Cipriano *et al.* 2002; Bougas *et al.* 2010). Also, higher expression of MHC IIβ comes at an energetic cost as measured through a lower hepatosomatic index and increased oxidative damage (Kurtz *et al.* 2006). Indeed, we recently found a positive correlation between condition index and domestic admixture in a natural context but a negative correlation between these two variables and MHC expression (see Lamaze *et al.* 2013). Altogether, these observations support the hypothesis that domestic brook charr are reallocating their energy towards growth and fat accumulation (Crespel *et al.* 2013) at the expense of other biological functions, including immunity.

In conclusion, this study used a combination of multivariate and gene expression analyses to characterize the functional immunological impacts of introgressive hybridization following the introduction of domestic individuals into wild populations. We have shown with the new $COIA_{2G}$ analysis that variation in allele frequencies is the result of combination of neutral and selective processes that vary across different environmental context (Fig. 5a). Parasite community variation between lakes resulted in alleles and pathogens covariation, underlying local adaptation in some populations (Fig. 5b). Also, our $COIA_{GP}$ suggested that the maintenance of MHC polymorphism resulted in the coexistence of specialist and generalist MHC alleles in terms of parasite detection (Fig. 5b,c). This translated into lower domestic alleles introgression at immunogenic genes than expected under neutrality in certain circum-

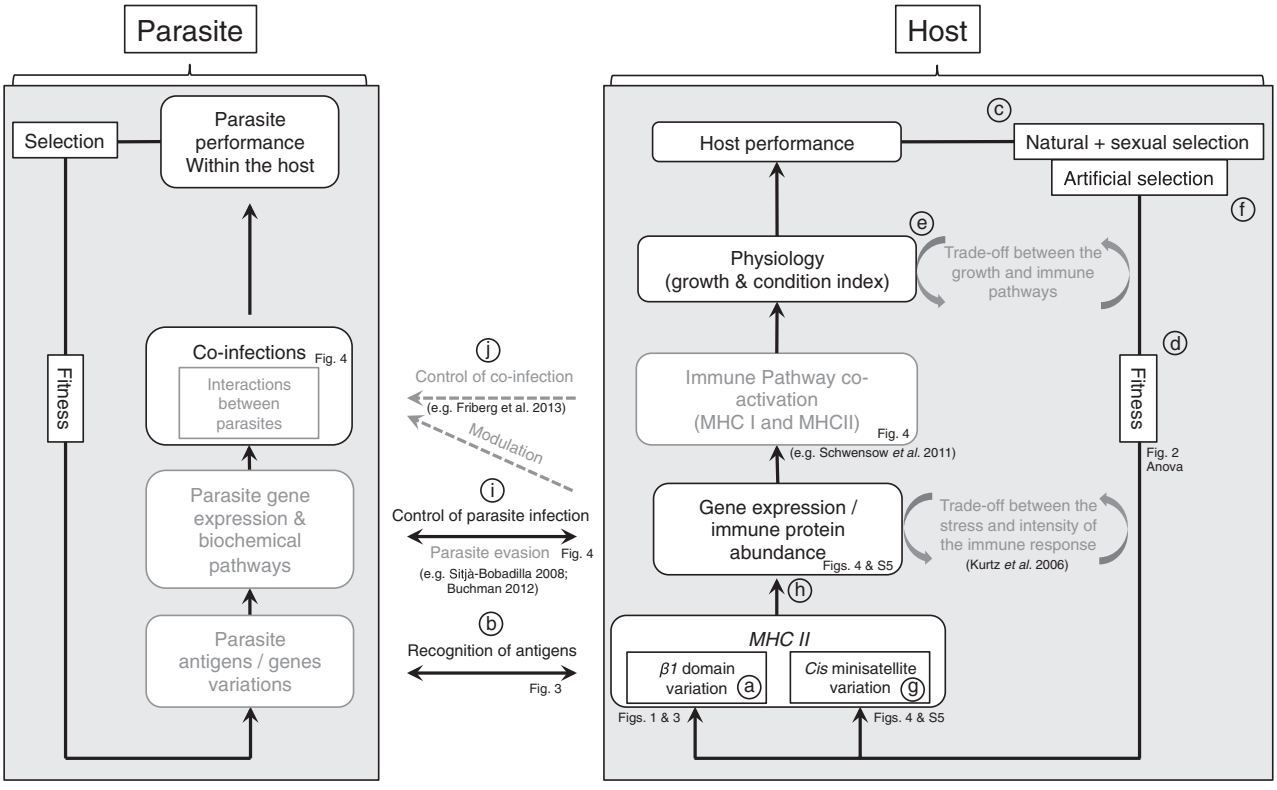

Fig. 5 Graph linking hypothetical sequence and expression variation and fitness by taking into account the dynamic host–parasite interactions. The arrows connect levels of biological organization. The grey colour represents links that have not been documented in natural conditions or are not supported by strong evidences in our study and need further investigations. The black colour represents links that were supported in our study.

stances, likely due to a lower fitness associated with an increased parasitic load in admixed individuals (Fig. 5d). We proposed a mechanistic scenario for the observed reduced fitness associated with the domestic alleles in the wild, whereby a trade-off between fast growth and immunity (Fig. 5e) promoted alleles with reduced expression during artificial selection (Fig. 5f). Indeed, we identified that *cis*-regulatory minisatellite length and admixture were positively correlated and both were associated with a down-regulation of MHC genes (Fig. 5g,h). Moreover, we have confirmed in a controlled environment that *cis*-regulatory minisatellite length down-regulated the expression of the MHC IIβ (Fig. 5g,h). A complex interplay among parasites and between host and parasite seems to emerge (Fig. 5), which suggests the importance to consider not only the sequence polymorphism (Fig. 5b) but also the regulation of expression (Fig. 5i,j; Loiseau *et al.* 2008; Gouy de Bellocq *et al.* 2008; Froeschke & Sommer 2012; Axtner & Sommer 2012). From a management perspective, these results further emphasize the need to restrict stocking activities using domestic strains to already extensively introgressed populations in order to mitigate the potential consequences on the genetic integrity and immune resiliency of locally adapted pristine populations.

Acknowledgements

We are grateful to P-A. Gagnaire for R script development for MHC genotyping. We thank M. Bachand, S. Boutin and G. Yannic for statistical comments and A. Dalziel for manuscript comments. We thank G. Côté, M-H Perreault, L. Papillon for laboratory assistance, and staff at the LARSA for fish husbandry and care. We also thank V. Bourret, S. Gerardi, B. Mayot, A. Marie, for their help with field sampling. We also want to thank B. Boyle from the sequencing platform at the IBIS institute (Université Laval, Québec, Canada) for his assistance in 454 library and genotyping. We are grateful to Associate editor J. Slate and three anonymous referees for their constructive comments and suggestions. This research was financially supported primarily by a strategic project grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to LB and DG, the Canadian Research Chair in Genomics and Conservation of Aquatic Resources to LB and The Collaborative Research and Training Experience (CREATE) programme to FL. We also acknowledge the important contributions of the Ministère des Ressources Naturelles et de la Faune du Québec (MRNF), the Société des Établissement de Plein-Air du Québec (SÉPAQ) for supporting this project. This study is a contribution to the research programme of Ressources Aquatiques Québec (RAQ).

References

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, **25**, 4692–4693.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, **16**, 613–622.

Axtner J, Sommer S (2012) The functional importance of sequence versus expression variability of MHC alleles in parasite resistance. *Genetica*, **140**, 407–420.

Beaumont MA, Nichols R (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **263**, 1619–1626.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.

Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, **16**, 363–377.

Bertrand M, Marcogliese D, Magnan P (2008) Trophic polymorphism in brook charr revealed by diet, parasites and morphometrics. *Journal of Fish Biology*, **72**, 555–572.

Bondinas GP, Moustakas AK, Papadopoulos GK (2007) The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. *Immunogenetics*, **59**, 539–553.

Boots M (2011) The evolution of resistance to a parasite is determined by resources. *The American Naturalist*, **178**, 214–220.

Bougas B, Granier S, Audet C, Bernatchez L (2010) The transcriptional landscape of cross-specific hybrids and its possible link with growth in brook charr (*Salvelinus fontinalis* Mitchill). *Genetics*, **186**, 97–107.

Bowden TJ (2008) Modulation of the immune system of fish by their environment. *Fish and Shellfish Immunology*, **25**, 373–383.

Brown JH, Jardetzky TS, Gorga JC, *et al.* (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–39.

Buchmann K (2012) Fish immune responses against endoparasitic nematodes—experimental models. *Journal of Fish Diseases*, **35**, 623–635.

Carroll SP (2011) Conciliation biology: the eco-evolutionary management of permanently invaded biotic systems. *Evolutionary Applications*, **4**, 184–199.

Cipriano RC, Marchant D, Jones TE, Schachte JH (2002) Practical application of disease resistance: a brook trout fishery selected for resistance to furunculosis. *Aquaculture*, **206**, 1–17.

Conejeros P, Phan A, Power M, *et al.* (2008) MH class IIalpha polymorphism in local and global adaptation of Arctic charr (*Salvelinus alpinus* L.). *Immunogenetics*, **60**, 325–337.

Coughlan J, Mcginnity P, O'Farrell B, *et al.* (2006) Temporal variation in an immune response gene (MHC I) in anadromous *Salmo trutta* in an Irish river before and during aquaculture activities. *Journal of Marine Science*, **63**, 1248–1255.

Crespel A, Bernatchez L, Garant D, Audet C (2013) Genetically based population divergence in overwintering energy mobilization in brook charr (*Salvelinus fontinalis*). *Genetica*, **141**, 51–64.

Croisetière S, Tarte P, Bernatchez L, Belhumeur P (2008) Identification of MHC class IIβ resistance/susceptibility alleles to *Aeromonas salmonicida* in brook charr (*Salvelinus fontinalis*). *Molecular Immunology*, **45**, 3107–3116.

Croisetière S, Bernatchez L, Belhumeur P (2010) Temperature and length-dependent modulation of the MH class II beta

gene expression in brook charr (*Salvelinus fontinalis*) by a cis-acting minisatellite. *Molecular Immunology*, **47**, 1817–1829.

Currens KP, Hemmingsen AR, French RA, *et al.* (1997) Introgression and susceptibility to disease in a wild population of rainbow trout. *North American Journal of Fisheries Management*, **17**, 1065–1078.

Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L (2007) Clinal variation in MHC diversity with temperature: evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. *Evolution*, **61**, 2154–2164.

Dionne M, Miller KM, Dodson JJ, Bernatchez L (2009) MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1555–1565.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, **256**, 50–52.

Dray S, Chessel D, Thioulouse J (2003) Co-inertia analysis and the linking of ecological data tables. *Ecology*, **84**, 3078–3089.

Dray S, Dufour A-B, Chessel D (2007) The ade4 package II: two-table and K-table methods. *R news*, **7**, 47–52.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Edmands S (2007) Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, **16**, 463–475.

Eizaguirre C, Lenz TL (2010) Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of Fish Biology*, **77**, 2023–2047.

Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012) Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, **15**, 723–731.

Evans ML, Neff BD (2009) Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Molecular Ecology*, **18**, 4716–4729.

Evans ML, Neff BD, Heath DD (2010a) MHC genetic structure and divergence across populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Heredity*, **104**, 449–459.

Evans ML, Neff BD, Heath DD (2010b) MHC-mediated local adaptation in reciprocally translocated Chinook salmon. *Conservation Genetics*, **11**, 2333–2342.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

de Eyto E, Mcginnity P, Huisman J, *et al.* (2011) Varying disease-mediated selection at different life-history stages of Atlantic salmon in fresh water. *Evolutionary Applications*, **4**, 749–762.

Fraser DJ, Cook AM, Eddington JD, Bentzen P, Hutchings JA (2008) Mixed evidence for reduced local adaptation in wild salmon resulting from interbreeding with escaped farmed salmon: complexities in hybrid fitness. *Evolutionary Applications*, **1**, 501–512.

Friberg IM, Little S, Ralli C, *et al.* (2013) Macroparasites at peripheral sites of infection are major and dynamic modifiers of systemic antimicrobial pattern recognition responses. *Molecular Ecology*, **22**, 2810–2826.

Froeschke G, Sommer S (2012) Insights into the complex associations between MHC class II DRB polymorphism and multiple gastrointestinal parasite infestations in the striped mouse. *PLoS ONE*, **7**, e31820.

Gandon S, Michalakis Y (2000) Evolution of parasite virulence against qualitative or quantitative host resistance. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **267**, 985–990.

Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, **18**, 1207–1224.

Gompert Z, Buerkle CA (2010) Introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, **10**, 378–384.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Gouy de Bellocq J, Charbonnel N, Morand S (2008) Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *Journal of Evolutionary Biology*, **21**, 1144–1150.

Grimholt U, Getahun A, Hermsen T, Stet RJM (2000) The major histocompatibility class II alpha chain in salmonid fishes. *Developmental and Comparative Immunology*, **24**, 751–763.

Grimholt U, Drabløs F, Jørgensen SM, Høyheim B, Stet RJM (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, Larsen S, Nordmo R, *et al.* (2003) MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics*, **55**, 210–219.

Handunnetthi L, Ramagopalan SV, Ebers GC, Knight JC (2010) Regulation of major histocompatibility complex class II gene expression, genetic variation and disease. *Genes and Immunity*, **11**, 99–112.

Hansen JD, Strassburger P, Thorgaard GH, Young WP, Du Pasquier L (1999) Expression, linkage, and polymorphism of MHC-related genes in rainbow trout, *Oncorhynchus mykiss*. *Journal of Immunology*, **163**, 774–786.

Harstad H, Lukacs MF, Bakke HG, Grimholt U (2008) Multiple expressed MHC class II loci in salmonids; details of one non-classical region in Atlantic salmon (*Salmo salar*). *BMC Genomics*, **9**, 193.

Harvell CD (1999) Emerging marine diseases-climate links and anthropogenic factors. *Science*, **285**, 1505–1510.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution*, **56**, 1902–1908.

Hendry AP, Stearns SC (eds) (2003) *Evolution Illuminated: Salmon and Their Relatives*. Oxford University Press, New York.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Janeway CA, Travers P, Walport M, Shlomchik M (2005) *Immunobiology: The Immune System in Health and Disease*. Garland Science, New York, New York.

Jombart T, Pontier D, Dufour A-B (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**, 330–341.

Kaltz O, Shykoff JA (1998) Local adaptation in host–parasite systems. *Heredity*, **81**, 361–370.

Kawecki T, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Klein J (1986) *Natural History of the Major Histocompatibility Complex*. Wiley, New York, New York.

Kosakovsky Pond SL (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, **22**, 1208–1222.

Krkošek M, Ford J, Morton A, *et al.* (2007) Declining wild salmon populations in relation to parasites from farm salmon. *Science*, **318**, 1772.

Kurtz J, Wegner KM, Kalbe M, *et al.* (2006) MHC genes and oxidative stress in sticklebacks: an immuno-ecological approach. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **273**, 1407–1414.

Laikre L, Schwartz MK, Waples RS, Ryman N; The GeM Working Group (2010) Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends in Ecology & Evolution*, **25**, 520–529.

Lamaze FC, Sauvage C, Marie A, Garant D, Bernatchez L (2012) Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, **21**, 2877–2895.

Lamaze FC, Garant D, Bernatchez L (2013) Stocking impacts the expression of candidate genes and physiological condition in introgressed brook charr (*Salvelinus fontinalis*) populations. *Evolutionary Applications*, **6**, 393–407.

Landry C, Bernatchez L (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **10**, 2525–2539.

Langefors A, Lohm J, Grahn M, Andersen O, von Schantz T (2001) Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **268**, 479–485.

Lê S, Josse J, Husson F (2008) FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, **25**, 1–18.

Leeds TD, Silverstein JT, Weber GM, *et al.* (2010) Response to selection for bacterial cold water disease resistance in rainbow trout. *Journal of Animal Science*, **88**, 1936–1946.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative pcr and the $2^{-\Delta\Delta C}$T method. *Methods*, **25**, 402–408.

Lohm J, Grahn M, Langefors A, *et al.* (2002) Experimental evidence for major histocompatibility complex-allele-specific resistance to a bacterial infection. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **269**, 2029–2033.

Loiseau C, Zoorob R, Garnier SP, *et al.* (2008) Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecology Letters*, **11**, 258–265.

Lukacs MF, Harstad H, Grimholt U, *et al.* (2007) Genomic organization of duplicated major histocompatibility complex class I regions in Atlantic salmon (*Salmo salar*). *BMC Genomics*, **8**, 251–251.

Marie AD, Bernatchez L, Garant D (2010) Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, **19**, 2025–2037.

McClelland E, Naish K (2007) What is the fitness outcome of crossing unrelated fish populations? A meta-analysis and an evaluation of future research directions. *Conservation Genetics*, **8**, 397–416.

Ministère des Ressources Naturelles et de la Faune du Québec (2008) *Lignes directrices sur les ensemencements de poissons. Secteur Faune Québec*, pp. 41. Direction de l'expertise sur la faune et ses habitats, Québec.

van der Most PJ, de Jong B, Parmentier HK, Verhulst S (2011) Trade-off between growth and immune function: a meta-analysis of selection experiments. *Functional Ecology*, **25**, 74–80.

Murrell B, Wertheim JO, Moola S, *et al.* (2012) Detecting individual sites subject to episodic diversifying selection (HS Malik, Ed.). *PLoS Genetics*, **8**, e1002764.

Neff B, Garner S, Heath J, Heath D (2008) The MHC and nonrandom mating in a captive population of Chinook salmon. *Heredity*, **101**, 175–185.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, **3**, 418–426.

van Oosterhout C, Smith AM, Hänfling B, *et al.* (2007) The guppy as a conservation model: implications of parasitism and inbreeding for reintroduction success. *Conservation Biology*, **21**, 1573–1583.

Pavey SA, Lamaze FC, Garant D, Bernatchez L (2011) Full length MHC II beta exon 2 primers for salmonids: a new resource for next generation sequencing. *Conservation Genetics Resources*, **3**, 665–667.

Pavey SA, Sevellec M, Adam W, *et al.* (2013) Nonparallelism in MHCIIβ diversity accompanies nonparallelism in pathogen infection of lake whitefish (*Coregonus clupeaformis*) species pairs as revealed by next-generation sequencing. *Molecular Ecology*, **22**, 3833–3849.

Plumb JA, Hanson LA (2011) *Health Maintenance and Principal Microbial Diseases of Cultured Fishes*. Wiley-Blackwell, Ames, Iowa.

Poulin R (2006) Global warming and temperature-mediated increases in cercarial emergence in trematode parasites. *Parasitology*, **132**, 143–151.

Pulkkinen K, Suomalainen LR, Read AF, *et al.* (2010) Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **277**, 593–600.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Ramstad A, Romstad AB, Knappskog DH, Midtlyng PJ (2007) Field validation of experimental challenge models for IPN vaccines. *Journal of Fish Diseases*, **30**, 723–731.

Rauw WM, Kanis E, Noordhuizen-Stassen EN, Grommers FJ (1998) Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science*, **56**, 15–33.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rhymer J, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology, Evolution, and Systematics*, **27**, 83–109.

Ricker WE (1975) Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada*, **191**, 1–382.

Roberge C, Einum S, Guderley H, Bernatchez L (2006) Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Molecular Ecology*, **15**, 9–20.

Roberge C, Páez DJ, Rossignol O, *et al.* (2007) Genome-wide survey of the gene expression response to saprolegniasis in Atlantic salmon. *Molecular Immunology*, **44**, 1374–1383.
Robinson J, Waller MJ, Fail SC, *et al.* (2009) The IMGT/HLA database. *Nucleic Acids Research*, **37**, D1013–D1017.
Rose AB (2008) Intron-mediated regulation of gene expression. *Current topics in microbiology and immunology*, **326**, 277–290.
Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, **5**, 725–738.
Schulenburg H, Kurtz J, Moret Y, Siva-Jothy MT (2009) Introduction. Ecological immunology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 3–14.
Schwensow N, Axtner J, Sommer S (2011) Are associations of immune gene expression, body condition and parasite burden detectable in nature? A case study in an endemic rodent from the Brazilian Atlantic Forest. *Infection, Genetics and Evolution*, **11**, 23–30.
Sitjà-Bobadilla A (2008) Living off a fish: a trade-off between parasites and the immune system. *Fish and Shellfish Immunology*, **25**, 358–372.
Slade RWR, McCallum HIH (1992) Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*, **132**, 861–864.
Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **277**, 979–988.
Stet RJM, de Vries B, Mudde K, *et al.* (2002) Unique haplotypes of co-segregating major histocompatibility class II A and class II B alleles in Atlantic salmon (*Salmo salar*) give rise to diverse class II genotypes. *Immunogenetics*, **54**, 320–331.
Sutton JT, Nakagawa S, Robertson BC, Jamieson IG (2011) Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology*, **20**, 4408–4420.
Tamura K, Peterson D, Peterson N, *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.
Tymchuk W, Sakhrani D, Devlin R (2009) Domestication causes large-scale effects on gene expression in rainbow trout: analysis of muscle, liver and brain transcriptomes. *General and Comparative Endocrinology*, **164**, 175–183.
Vandersteen Tymchuk W, O'Reilly P, Bittman J, Macdonald D, Schulte P (2010) Conservation genomics of Atlantic salmon: variation in gene expression between and within regions of the Bay of Fundy. *Molecular Ecology*, **19**, 1842–1859.
Villanueva B, Fernandez J, Garcia-Cortes LA, *et al.* (2011) Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs. *Journal of Animal Science*, **89**, 3433–3442.
Weir B, Cockerham CC (1984) Estimating f-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
Westerdahl H, Asghar M, Hasselquist D, Bensch S (2012) Quantitative disease resistance: to better understand parasite-mediated selection on major histocompatibility complex. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **279**, 577–584.site.
Wynne J, Cook M, Holmes B, Elliott N (2007) Allelic and haplotypic diversity at the major histocompatibility class II within domesticated Australian Atlantic salmon (*Salmo salar* L.). *Journal of Fish Biology*, **70**, 45–59.
Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
Zhu X, Luke A, Cooper RS, *et al.* (2005) Admixture mapping for hypertension loci with genome-scan markers. *Nature Genetics*, **37**, 177–181.
Zuur AF, Ieno EN, Elphick CS (2009) A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, **1**, 3–14.

The genetic and gene expression experiments in natural and laboratory conditions were conducted by F.C.L. and G.R., respectively. The parasite screening was carried out by F.C.L. The MHC genotyping procedure was performed by F.C.L., S.A.P. and E.N. The data analysis and protein modelling was performed by F.C.L. The manuscript was written by F.C.L., S.A.P., D.G. and L.B. The study was conceived by F.C.L., D.G. and L.B.

Data accessibility

Raw MHCIIb 454 reads, MHC, pathogen and gene expression matrices and databases, and MHC allele sequences are deposited in the Dryad repository (doi:10.5061/dryad.12ck0). The pipeline used to determine and assign alleles to the individuals is available at this link: github.com/ enormandeau/ngs_genotyping.

Supporting information

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

Table S1 Descriptive genetic statistics for the MHC IIβ locus.

Table S2 Sequencing Primers, Real-Time PCR Primers, and Taqman MGB Probe for Each Candidate Gene.

Table S3 Parasites abundance for 2008 and 2009.

Fig. S1 NeighbourNet of the 29 MHC II beta 1 domains observed in *Salvelinus fontinalis*.

Fig. S2 Amino acid alignment of the 29 MHC IIβ alleles of *Salvelinus fontinalis*.

Fig. S3 Allelic frequency distribution at the MHC IIβ gene.

Fig. S4 The simulated three-dimensional structure model of the beta 1 domain of *Salvelinus fontinalis* MHC class II.

Fig. S5 Box plot presenting MHC IIβ expression in brook charr head kidney as function of minisatellite repeat number in MHC IIβ intron 2 and temperature, (8 or 20 °C).

Appendix S6 Supplementary material and methods.