

# Genetic variation and balancing selection at MHC class II exon 2 in cultured stocks and wild populations of orange-spotted grouper (*Epinephelus coioides*)

Z.N. Meng\*, S. Yang\*, B. Fan, L. Wang and H.R. Lin

State Key Laboratory of Biocontrol, Institute of Aquatic Economic Animals and the Guangdong Province Key Laboratory for Aquatic Economic Animals, School of Life Sciences, Sun Yat-Sen University, Guangzhou, China

\*These authors contributed equally to this study.

Corresponding authors: L. Wang / H.R. Lin

E-mail: lewang.wang@hotmail.com / lsslhr@mail.sysu.edu.cn

Genet. Mol. Res. 11 (4): 3869-3881 (2012)

Received February 2, 2012

Accepted August 3, 2012

Published November 12, 2012

DOI <http://dx.doi.org/10.4238/2012.November.12.4>

**ABSTRACT.** Major histocompatibility complex (MHC) molecules play vital roles in triggering adaptive immune responses and are considered the most variable molecules in vertebrates. Recently, many studies have focused on the polymorphism and evolution mode of MHC in both model and non-model organisms. Here, we analyzed the MHC class II exon 2-encoding  $\beta$  chain in comparison with the mitochondrial *Cytb* gene and our previously published microsatellite data set in three cultured stocks and four wild populations of the orange-spotted grouper (*Epinephelus coioides*) in order to investigate its genetic variation and mechanism of evolution. We detected one to four alleles in one individual, suggesting that at least two loci exist in the orange-spotted grouper, as well as a particularly high level of allelic diversity at the MHC loci. Furthermore, the cultured stocks exhibited reduced allelic diversity compared to the wild counterparts. We found evidence of balancing selection at MHC class II exon 2, and codon

sites under positive selection were largely correspondent to the protein-binding region. In addition, MHC class II exon 2 revealed significant differences between population differentiation patterns from the neutral mitochondrial *Cytb* and microsatellites, which may indicate local adaptation at MHC loci in orange-spotted grouper originating from the South China Sea and Southeast Asia.

**Key words:** MHC class II; Orange-spotted grouper; Allelic diversity; *Epinephelus coioides*; Balancing selection

## INTRODUCTION

The major histocompatibility complex (MHC) is a multi-gene family central to the vertebral immune system. The two major classes, MHC class I and class II, encode cell surface receptors that present pathogen-derived peptides to T cells, thus playing pivotal roles in activating adaptive immune responses (Klein, 1986). MHC class I, expressed in almost all nucleated cells, recognizes and presents peptides from intracellular pathogens, especially viruses. Meanwhile, MHC class II, expressed by specialized antigen-presenting cells such as dendritic cells, macrophages, and B cells, picks up and presents peptides from extracellular pathogens, such as many bacteria, to CD4<sup>+</sup> receptors in T-helper cells, and triggers pathogen-specific antibody production (Dengjel et al., 2005). The functions of parasite resistance are considered intimately connected to MHC class II genes (Anmarkrud et al., 2010). The most important functional region for MHC class II molecules is considered the protein-binding region (PBR) corresponding to the  $\beta 1$  domain encoded by exon 2.

Consistent with the essential functions of MHC in recognizing various pathogens and activating specific immune responses, the MHC contains the most variable functional genes and high allelic diversity in order to confer greater pathogen resistance and survival advantages (Piertney and Oliver, 2006). It is reported that its functions in immune responses are due to not only allelic polymorphism at a single locus, but also gene duplication (Axtner and Sommer, 2007).

The extreme polymorphism at the MHC loci may be attributed to several selection processes. First, the co-evolution between parasites and hosts and strong selection on the MHC loci by pathogens are most likely the main driving force (Piertney and Oliver, 2006), as MHC genes are prone to exhibiting signatures of adaptation to the pathogen spectrum (Hill et al., 1991). Multiple lines of evidence have established the MHC type with disease resistance or susceptibility in different animals (Hill et al., 1991; Hill, 1999). Based on local adaptation, a geographical variation for MHC genes was also observed in birds recently (Ekblom et al., 2007; Alcaide et al., 2008), although such differentiation was not treated as the general pattern in non-model vertebrates (Bernatchez and Landry, 2003).

Second, the large variation at MHC genes is believed to be maintained by some form of balancing selection, although the exact nature of selection continues to be debated (Hedrick, 1999). As an example, a much higher level of excess non-synonymous over synonymous nucleotide substitutions was observed within the PBR (encoded by exon 2) compared to the non-PBR (also encoded by exon 2) in the  $\beta$  chain of an MHC class II molecule, and was considered evidence for “balancing selection” (Hughes and Nei, 1988). Such a selection pattern

in the PBR has been extensively reported in many species (Xu et al., 2010; Schad et al., 2011).

Another paradigm contributing to MHC polymorphism has been attributed to mate choice. Several studies show that individuals may use olfaction to assess MHC type during mate selection (Brown and Eklund, 1994). While MHC differentiates self from non-self, it could also allow differentiation of relatives from non-relatives, thus help to minimize inbreeding, improve gene pool variation, and enhance progeny immunological capacity (Piertney and Oliver, 2006). However, controversial study results have questioned the validity of mate selection as a major driving force of MHC diversity (Jordan and Bruford, 1998). In addition, maternal-fetal incompatibility may also play an important role in contributing to a high level of polymorphism at the MHC genes (Clarke and Kirby, 1966).

Natural selection and random genetic drift play important roles in the domestication process of fish populations, which typically lead to depletion of the levels of genetic diversity of cultured stocks (Karaïskou et al., 2010). Such reduced genetic variation will result in adverse effects on farmed populations, especially in the aspect of pathogen resistance. However, the studies on genetic changes between farmed and wild populations have mainly focused on neutral markers. Little is known about the differences of genetic variation between the protein-coding region of farmed and wild populations (Karaïskou et al., 2010). Possessing extremely high levels of genetic polymorphism, MHC genes are considered very sensitive markers for detecting genetic variation and differentiation between farmed and wild populations (Karaïskou et al., 2010). Since the process of adaptation, selection, and domestication are common in fish culture, MHC molecules that are under selective pressure could be useful tools for studying the genetic changes between cultured stocks and the wild population. During farming processes, adaptation to specific environments and the effects of bottleneck and random genetic drift stemming from a limited number of broodstocks can reduce the levels of genetic variation of cultured stocks, which can be detected by investigating MHC molecules (Karaïskou et al., 2010).

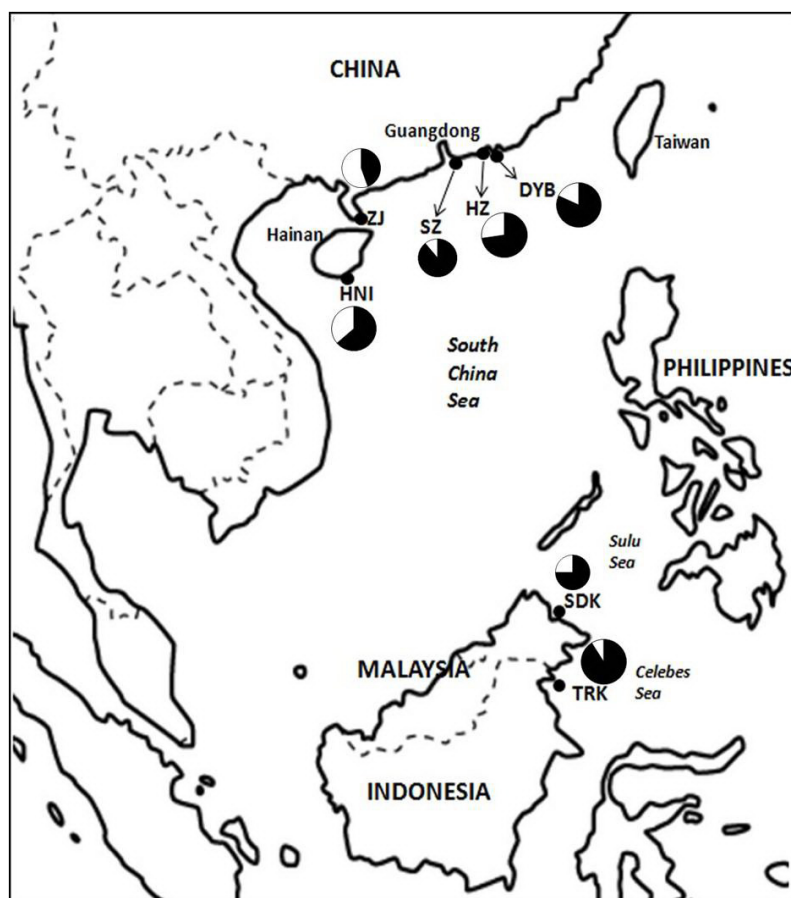
The orange-spotted grouper (*Epinephelus coioides*) is distributed across a large area, from eastern Africa southward to South Africa and eastward to the western Pacific. Especially in the Asia-Pacific region, it is treated as an economically important marine food fish species. In recent years, this species has been successfully cultured in China and the Southeast Asian countries (Wang et al., 2011). In this study, we isolated and sequenced the MHC class II exon 2 in three cultured stocks and four wild populations of orange-spotted grouper originating from the South China Sea and Southeast Asia. Furthermore, we analyzed the genetic polymorphism and evolution patterns of MHC class II exon 2 in comparison with mitochondrial DNA (mtDNA), as well as the microsatellite data set of our previous study (Wang et al., 2011). Our aims were to i) investigate the genetic variation of the MHC class II gene and genetic changes between cultured stocks and wild populations of orange-spotted grouper, ii) test for signatures of selection across codon sites to explore the mechanisms of adaptive evolution for cultured stocks and wild populations, and iii) study the population divergence patterns for different marker types to help understand the mechanism of local adaptation at the MHC class II gene.

## MATERIAL AND METHODS

### Sample collection

The orange-spotted grouper samples were from the South China Sea and Southeast

Asia, which included four wild populations and three cultured stocks as described by Wang et al. (2011) (Figure 1). The number of individuals used for MHC and mt-DNA analysis in this study is shown in Table 1.



**Figure 1.** Sampling localities of three cultured stocks and four wild populations of the orange-spotted grouper originated from the South China Sea and the Southeast Asia. Genetic variation in means of contributions from different lineages of mt-*Cytb* gene is denoted as pies, of which size is proportional to the number of haplotypes and colors represent lineages corresponding to Figure 3. Note: wild populations including DYB, HNI, SDK and TRK; cultured stocks including HZ, SZ and ZJ.

### MHC and mt-DNA sequencing

The polymorphism of MHC class II exon 2 was screened by single-strand conformation polymorphism (SSCP) analysis. A 258-bp fragment of exon 2 was targeted for sequencing using the following primer set: 5'-GTTTTTGTTTCCTCAGATGG-3' and 5'-CAGGTGAGTC TGTGTTTCTG-3'. **Double-strand DNA amplification was performed in a 20- $\mu$ L volume containing 0.5  $\mu$ M of each primer, 0.2 mM dNTPs, 1.5 mM MgCl<sub>2</sub>, 1X PCR buffer, 1 U Taq DNA polymerase (Takara, China), and 20-50 ng template genomic DNA. The PCR was conducted**

for 5 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 56°C, and 30 s at 72°C, with a final elongation of 2 min at 72°C. PCR products were checked against a 50-bp DNA ladder (Takara) stained with ethidium bromide. In SSCP analysis, 5 µL PCR products was first mixed with 10 µL denatured loading dye containing 95% formamide (v/v), 10 mM NaOH, 0.05% bromophenol blue (w/v), and 0.05% xylene cyanol (w/v). The mixture was denatured at 95°C for 10 min and chilled on ice immediately, and then loaded on a 12% non-denaturing polyacrylamide gel (37.5:1) for genotyping. Electrophoresis was carried out for 16 h, with the gel consistently cooled to 4°C. The resulting bands were visualized by silver staining. A very high level of polymorphism was detected in this analysis. Thus, all SSCP bands were cut out and incubated in 100 µL TE buffer for at least 3 h at 37°C. After centrifugation, 2 µL incubated solution was used as the template for PCR using the same conditions as described above. The PCR products were again screened using SSCP analysis, and every band was sequenced in both directions with the primers above using an ABI PRISM 3730 DNA automated sequencer (Applied Biosystems, USA).

We amplified the entire cytochrome b (*Cytb*) gene in mt-DNA analysis using the following primer set: 5'-TAACCAGGACTTATGGCTTG-3' and 5'-AGAACGCCGGTCTTGTAAGC-3'. PCR amplification was performed in a 50-µL volume containing 0.5 µM of each primer, 0.2 mM dNTPs, 1.5 mM MgCl<sub>2</sub>, 1X PCR buffer, 1 U Taq DNA polymerase (Fermentas), and 20 ng template genomic DNA. The PCR program was as follows: 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, and 60 s at 72°C, with a final extension of 5 min at 72°C. The PCR products were then purified and used as templates for sequencing in both directions with the above primers. Sequencing was conducted with an ABI PRISM 3730 DNA automated sequencer.

## Data analysis

The MHC class II exon 2 and mt-*Cytb* sequences were edited and aligned with BioEdit version 7.0.5.3 (Hall, 1999) and DnaSP 4.10.4 (Rozas et al., 2003). Summary statistics of genetic diversity, including haplotype and nucleotide diversity, were also calculated using DnaSP 4.10.4 (Rozas et al., 2003).

The identified sequences at the MHC locus were checked for homology against *Xiphophorus* fishes (Summers et al., 2009). New alleles were nominated according to accepted nomenclature rules (Davies et al., 1997). The functions of specific MHC amino acid residues in antigen recognition have been extensively emphasized in many previous studies (Brown et al., 1993; Tong et al., 2006; Alcaide et al., 2008). Therefore, we assigned the PBR and non-PBR in our analysis according to Tong et al. (2006).

Evidence for selection at this locus was tested by estimating the ratio ( $\omega$ ) of nonsynonymous ( $d_N$ ) to synonymous ( $d_S$ ) substitutions. The overall  $\omega$  values were computed using the traditional method of Nei and Gojobori (1986) in MEGA version 5 (Tamura et al., 2011). This analysis was performed over all sites and for the PBR, and was tested against the null hypothesis of strict neutrality ( $d_N = d_S$ ) using a Z-test with 1000 bootstraps (Nei and Gojobori, 1986). Traditional tests of neutrality such as the Fu and Li F\*, the Fu and Li D\*, and the Tajima D were estimated using DnaSP 4.10.4 (Rozas et al., 2003), which can provide an indication of non-neutral evolution. In addition, we studied the spatial variation of positive selection on amino acid sites using maximum likelihood methods by investigating the

ratio ( $\omega$ ) of  $d_N/d_S$  substitution rates. We used the code-based likelihood program CODEML implemented in the PAML 3.14 program suite to detect positive selection among amino acid sites (Yang, 1997). Positive selection would be indicated if  $\omega (d_N/d_S) > 1$ . This analysis was conducted under M0 (one-ratio), M1a (nearly neutral), M2a (positive selection), M7 ( $\beta$ ), and M8 ( $\beta$  and  $\omega$ ) models, where M7 and M8 are considered robust against recombination (Anisimova et al., 2003). The likelihood ratio test was used to compare M1a and M2a, and M7 and M8. If M2a and M8 fit significantly better than their counterparts did, the codon sites under positive selection would be identified through the Bayes empirical Bayes (EBE) method (Yang et al., 2005).

In terms of genetic divergence, we calculated the overall population differentiation ( $F_{ST}$ ) for MHC class II exon 2, *Cytb*, and our previously published microsatellite data using FSTAT version 2.9.3.2 (Goudet, 2002) and Arlequin version 3.11 (Excoffier et al., 2005). For the microsatellite data, we generated the 95% confidence interval (95%CI) of the  $F_{ST}$  values by bootstrapping over loci. We also estimated the nucleotide sequence-based pairwise population genetic differentiation for MHC and mt-*Cytb* using Arlequin version 3.11 (Excoffier et al., 2005) in order to compare with the microsatellite data. The relationship of population differentiation patterns among MHC class II exon 2, *Cytb*, and the microsatellites was tested using the Mantel test. Furthermore, phylogenetic and molecular evolutionary analyses of both types of sequences were conducted by maximum likelihood using MEGA version 5 with 1000 bootstraps (Tamura et al., 2011). Before this analysis, the most appropriate evolutionary model (TN92 + G + I and HKY + G for MHC and *Cytb*, respectively) was estimated with the Bayesian Information Criterion in MEGA version 5.

## RESULTS

### Genetic variation at MHC exon 2 and mt-*Cytb*

We sequenced exon 2 of the MHC class II  $\beta$  gene in 87 individuals and obtained a final alignment of 258 bp after correcting polymerase errors based on sequencing in both directions and by excluding irreproducible single-nucleotide polymorphisms. These 87 individuals produced 222 sequences, among which 184 were unique (Table 1). These alleles have been designated *Epc-DAB\*0101* to *Epc-DAB\*8601* (GenBank accession Nos. JN980404-JN980625). No insertions, deletions, or stop codons were detected for these sequences, which suggest that all sequences are functional molecules in the genome. Analysis of the 258-bp sequence showed that 97 nucleotides (37.597%) and 51 predicted amino acid sites (59.302%) were variable. The nucleotide diversity for each sample ranged from 0.098 (SZ) to 0.133 (HNI) (Table 1). These sequences exhibited an unequally distributed pattern among individuals because one to four alleles were identified in each individual, which suggests that at least two loci of the MHC class II  $\beta$  gene exist in the genome of the orange-spotted grouper. For this reason, it was impossible to assess the genetic changes between cultured and wild samples by estimating genetic diversity in the form of heterozygosity. Therefore, we used allelic diversity, calculated in the form of haplotype diversity for each sample, and the allelic diversity ranged from 0.983 (HZ) to 1.000 (SDK). We also determined that the allelic diversity of cultured samples was significantly reduced compared to their wild counterparts (two-tailed *t*-test,  $P = 0.001$ ; Table 1).

**Table 1.** Number of individuals analyzed and genetic diversity for major histocompatibility complex (MHC) class II exon 2 and mt-*Cytb* gene in each sample.

Samples	Origin	MHC				Mitochondrial			
		n	N	Ad	$\pi$	n	H	Hd	$\pi$
HZ	Cultured	11	25	0.983	0.108	11	5	0.818	0.007
SZ	Cultured	15	35	0.985	0.098	9	4	0.694	0.003
ZJ	Cultured	11	30	0.986	0.118	9	4	0.750	0.007
DYB	Wild	10	21	0.995	0.128	12	5	0.618	0.005
HNI	Wild	12	33	0.998	0.133	11	5	0.667	0.006
TRK	Wild	11	33	0.994	0.119	11	4	0.600	0.003
SDK	Wild	17	45	1.000	0.112	8	3	0.607	0.006
Total	Wild	87	222	0.998	0.118	71	11	0.705	0.005

n = number of individuals analyzed; N = number of sequences isolated; Ad = allelic diversity; Hd = haplotype (gene) diversity;  $\pi$  = nucleotide diversity; H = haplotypes. The difference of Ad for the MHC gene between cultured stocks and wild populations is statistically significant ( $P = 0.001$ ).

Mitochondrial *Cytb* sequence alignment comprised 1141 bp and showed no evidence of stop codons, insertions, or deletions (GenBank accession Nos. JN980626-JN980696). Of the entire 1141-bp sequence, 20 variable sites with 19 parsimony informative sites and 11 haplotypes were detected among the 71 individuals analyzed (Table 1). The haplotype and nucleotide diversity for each sample ranged from 0.600 (TRK) to 0.818 (HZ) and from 0.003 (SZ and TRK) to 0.007 (HZ and ZJ) (Table 1). No significant reduction of genetic diversity was observed in cultured samples as compared to their wild counterparts ( $P > 0.05$ ; Table 1).

### Patterns of selection at MHC exon 2

Using the aforementioned methods in our analyses, we did not detect significant differences in terms of selection between cultured and wild samples. Considering the fact that the cultured stocks were not significantly different from the wild populations in terms of nucleotide diversity and that all samples exhibited no significant pairwise differentiation (see below), we pooled the samples to investigate the patterns of selection. All traditional testing methods yielded positive values (Fu and Li  $D^* = 1.462$ ,  $P > 0.05$ ; Fu and Li  $F^* = 1.017$ ,  $P > 0.10$ , and Tajima  $D = 0.174$ ,  $P > 0.10$ ), which revealed signals of balancing selection at MHC class II exon 2. However, the results of the Z-tests indicated that positive selection might exist both in cultured and wild samples in the PBR only (Table 2). The maximum likelihood ratio tests between Model M1a ( $\ln L = -19795.429$ ) and M2a ( $\ln L = -8707.819$ ) and between M7 ( $\ln L = -15087.362$ ) and M8 ( $\ln L = -8632.992$ ) revealed highly significant differences ( $P < 0.001$ ; Table 3), which also indicated positive selection at sites in the MHC class II  $\beta$  gene. Further analyses based on EBE showed that M2a and M8 identified 10 and 9 codons, respectively, under positive selection with a posterior probability of more than 95% (Table 3 and Figure 2). In addition, the codon sites under positive selection largely corresponded to the amino acid sites of the PBR (Figure 2).

**Table 2.**  $d_N/d_S$  (omega) values and P values of Z-tests (Z) for partitioned regions of MHC class II exon 2.

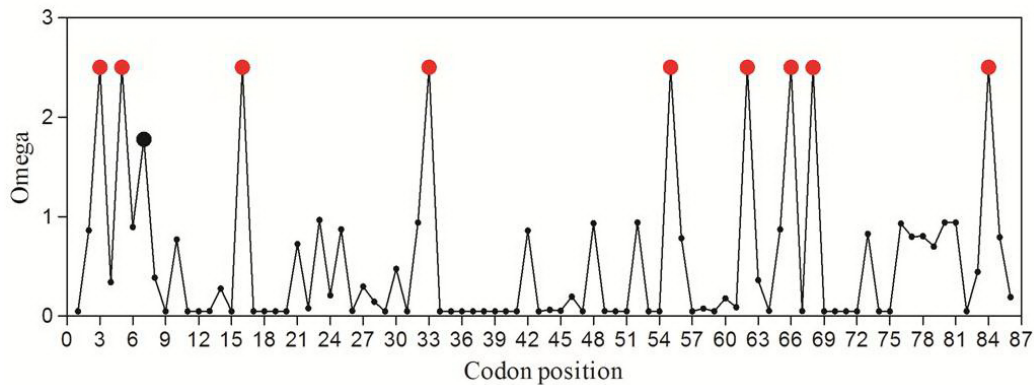
Sample	All domain		PBR		Non-PBR	
	Omega	P (Z)	Omega	P (Z)	Omega	P (Z)
All	3.281	0.001	2.850	0.005	1.375	0.172
Wild	3.170	0.002	2.932	0.004	1.420	0.158
Cultured	3.040	0.003	3.078	0.003	1.157	0.250

PBR = protein-binding region.

**Table 3.** Results of maximum likelihood models of the MHC class II exon 2 inferred from CODEML implemented in the PAML.

Model	Ln likelihood	Parameter estimates	Positively selected sites
M0 (one ratio)	-25499.727	$\omega = 643.568$	None
M1a (neutral)	-19795.429	$p_0 = 0.057, p_1 = 0.943, \omega_0 = 0.014, \omega_1 = 1.000$	Not allowed
M2a (selection)	-8707.819	$p_0 = 0.541, p_1 = 0.412, p_2 = 0.047, \omega_0 = 0.029, \omega_1 = 1.000, \omega_2 = 999.000$	<b>3, 5, 7, 16, 23, 33, 55, 62, 66, 68, 84</b>
M7 ( $\beta$ )	-15087.362	$p = 0.092, q = 0.024$	Not allowed
M8 ( $\beta$ & $\omega$ )	-8632.992	$p_0 = 0.959, p_1 = 0.041, p = 0.121, q = 0.173, ws = 999.000$	<b>3, 5, 7, 16, 33, 55, 62, 66, 68, 84</b>

Positively selected amino acid sites were calculated with Bayes empirical Bayes (Yang et al., 2005). Codon sites under positive selection at the 95% level are denoted in bold.



**Figure 2.** Spatial variation of the selection parameter  $\omega$  across MHC class II exon 2. Parameter estimates were carried out under the model M2a and M8 in the program CODEML implemented in the PAML. Sites under positive selection with model M2a (red and black) and M8 (red) were denoted.

### Patterns of population differentiation

The overall  $F_{ST}$  values for MHC class II exon 2 and *Cytb* were 0.019 and 0.011, respectively, which were significantly lower than that of the microsatellite data, 0.036 (95%CI = 0.027-0.047). Pairwise estimates of population differentiation showed that almost all the divergences between samples were not significant for MHC class II exon 2 and *Cytb* after Bonferroni's correction for multiple comparisons ( $P = 0.0024$ ) except for one, which was between HZ and SDK at the MHC locus ( $F_{ST} = 0.029$ ;  $P = 0.001$ ) (Table 4). These results were obviously different from those of the microsatellite data, which demonstrated significant pairwise population divergence except for two pairs (35). We did not detect any correlation of population differentiation patterns among MHC class II exon 2, *Cytb*, and the microsatellite loci in the Mantel tests, all of whose P values exceeded 0.05 (35). From comparison between the cultured stock and wild populations, we found some evidence of genetic divergence between and among these two groups at the MHC locus. Pairwise  $F_{ST}$  analysis showed that 14 of 21 comparisons were significant at 0.05, among which 12 comparisons were between cultured and wild samples, and among cultured stocks (Table 4). Phylogenetic analysis based on *Cytb* revealed two genetic lineages among all samples, and the haplotype distribution exhibited a

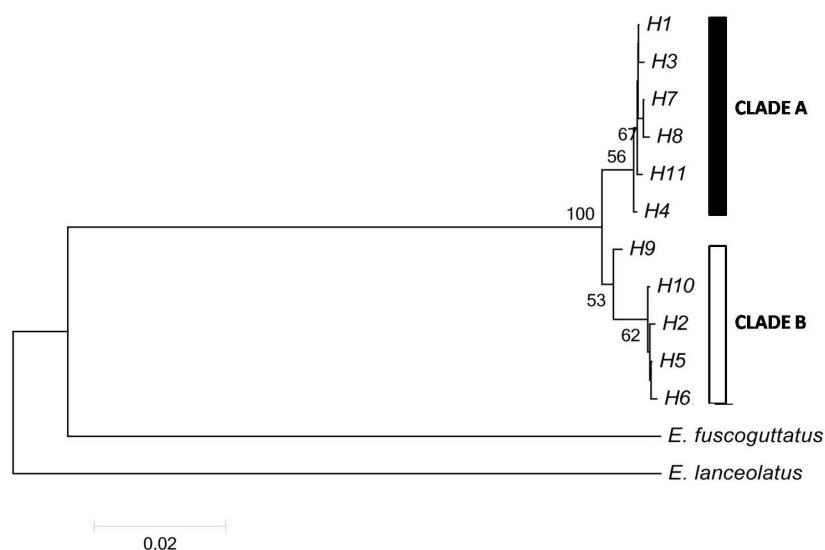


different pattern compared to that of the MHC class II  $\beta$  gene, which displayed only one allelic lineage (Figures 1 and 3). Moreover, the phylogenetic analyses of these two marker types revealed no geographical structuring among the wild samples (Figure 1 and Table 4).

**Table 4.** Matrix of pairwise  $F_{ST}$  values based on MHC class II exon 2 (below diagonal) and *Cytb* gene (above diagonal) among four wild populations and three cultured stocks of the orange-spotted grouper.

Pop	HZ	SZ	ZJ	DYB	HNI	TRK	SDK
HZ	0	-0.012	0.068	-0.033	-0.047	0.041	-0.083
SZ	0.045*	0	0.261	-0.086	-0.004	-0.102	-0.063
ZJ	0.021*	0.033*	0	0.165	0.038	0.316	0.056
DYB	0.021*	0.036*	0.015	0	-0.059	-0.066	-0.101
HNI	0.028*	0.028*	0.012	-0.020	0	0.030	-0.099
TRK	0.028*	0.035*	0.020*	0.006	-0.003	0	-0.027
SDK	0.029**	0.029*	0.009	0.007	0.014*	0.025*	0

\*Significant at 0.05 level; \*\*significant after sequential Bonferroni's correction; corrected P = 0.0024.



**Figure 3.** Maximum likelihood tree based on 11 haplotypes of the mt-*Cytb* gene of the orange-spotted grouper under the mutation model of HKY+G. Bootstrap replicates of more than 50% are presented. *Epinephelus fuscoguttatus* (HQ174891) and *E. lanceolatus* (HQ174847) are used as outgroup.

## DISCUSSION

### Genetic polymorphism within and between cultured stocks and wild populations at MHC locus

Until now, little is known about the polymorphism and genetic structure of the MHC in marine fishes, due to either the high number of gene copies in the whole genome or to complicated patterns of evolution. In our study, we detected up to four different sequences in a single individual, suggesting that at least two loci of the MHC class II  $\beta$  gene exist at the

genome level in the orange-spotted grouper. Such a high level of gene duplication has also been reported in other fish species (Xu et al., 2010), amphibians (Hauswaldt et al., 2007), birds (Westerdahl et al., 2004), and mammals (Schad et al., 2011). The presence of these multiple loci has been argued as being one important reason for the high level of polymorphism at MHC genes (Grimholt et al., 2003), although its mechanism is still far from understood. Moreover, we detected a much higher level of sequence variation, including nucleotide diversity (ranging 0.098-0.113) and allelic diversity (ranging 0.983-1.000), compared to that of *Cytb* (ranging 0.003-0.007 and 0.600-0.818, respectively) (Table 1) that also indicates much higher genetic polymorphism of the MHC class II  $\beta$  gene. The sequence variation of the MHC class II  $\beta$  gene possibly associates with the properties of binding antigen peptides and directly reflects the level of pathogen diversity (Jones et al., 2006). Although the patterns of gene duplication limited our study to allelic heterozygosities, we found a significant reduction in allelic diversity for the cultured stocks as compared to the wild populations (Table 1) that corresponds to the genetic diversity patterns revealed by our previous microsatellite analysis (Wang et al., 2011). This significantly reduced genetic diversity was most likely due to founder effects and random genetic drift in the cultured stocks, which were produced by broodstocks with a small number of individuals as described by Wang et al. (2011). At the same time, such a high level of genetic variation at the MHC locus may suggest that a bottleneck has not occurred in the cultured stocks, which is also consistent with the results of previous microsatellite analysis (Wang et al., 2011). As for mt-*Cytb*, we did not observe a reduced pattern of genetic variation in the cultured stocks as compared to their wild counterparts (Table 1), an indication that mt-DNA is not sensitive enough to reflect recent gene flow (Rokas et al., 2003). Regarding genetic divergence, we also detected some evidence of differentiation between cultured and wild samples and among cultured samples. This may indicate that aquaculture practices have influenced cultured stocks by selection, founder effects, and random genetic drift (Wang et al., 2011). However, this should be considered with caution because such differentiation was no longer significant after correction for multiple comparisons. The low genetic differentiation at the MHC locus may be reasonable since the studied cultured samples were the first generation of captive broodstocks and had not been artificially selected (Wang et al., 2011).

### Balancing selection at MHC locus

In our study, we detected evidence of balancing selection at MHC class II exon 2. First, all the traditional tests, including the Fu and Li D\*, Fu and Li F\*, and the Tajima D yielded positive values, which revealed signals of balancing selection (Rozas et al., 2003). Second, our data sets provided strong evidence of balancing selection on the amino acids in MHC class II exon 2, which revealed a very high level of allelic and nucleotide diversity as compared to a neutral mt-DNA sequence (Table 1). This high level of genetic variation is considered an indicator of balancing selection at MHC class II exon 2 (Strand et al., 2007). Nucleotide diversity in particular is expected to be high for loci undergoing balancing selection (Strand et al., 2007). Lastly, the selection patterns revealed by the Z-tests, in that the PBR exhibited higher  $\omega$  values compared to the non-PBR are considered consistent with the model under balancing selection (Hughes and Nei, 1988; Hughes and Yeager, 1998). In addition, high  $d_N/d_S$  values and the high posterior probability of selection for codon sites were, to a large degree, confined to the PBR (Tables 2 and 3 and Figure 2), also suggesting balancing

selection at the loci (Anmarkrud et al., 2010). The same patterns of balancing selection have already been reported in many studies (Bernatchez and Landry, 2003; Anmarkrud et al., 2010; Xu et al., 2010). However, it is still uncertain whether such a selection pattern is due to the overdominance of selection pressure on the PBR (Anmarkrud et al., 2010). Nevertheless, our study demonstrated a high level of genetic variation at the MHC locus in the orange-spotted grouper that is largely maintained by balancing selection. This is particularly important for cultured stocks of orange-spotted grouper, as balancing selection can reduce the adverse effects of farming practices (Karaïskou et al., 2010).

### Patterns of genetic divergence and evolution mechanism at MHC locus

We observed different patterns of genetic divergence for mt-DNA and MHC class II exon 2 based on the phylogenetic analysis, which revealed two genetic lineages for *Cytb* but only one genetic lineage for the MHC locus (Figure 3). In terms of population differentiation, the MHC locus, *Cytb*, and microsatellites also exhibited differing patterns (Mantel test,  $P > 0.05$ ). The discrepancy of the divergence patterns among these three marker types may reflect their different evolutionary mechanism not only in mutation rates but also in selective regimes (Alcaide et al., 2008). For instance, due to slow mutation rates, sex-biased inheritance, and being non-recombinant, mt-DNA is more sensitive for detecting phylogeographical structure and the historical evolutionary process, but is much less useful in reflecting recent gene flow (Rokas et al., 2003). On the contrary, hypervariable microsatellites under the stepwise mutation model are considered much more suitable for detecting recent gene flow among populations (Rokas et al., 2003). Different from the two neutral markers above, the mutation model of the MHC gene is rather complicated and far from understood. Nevertheless, the MHC gene is considered to be under pathogen-driven selection pressure because of its functions in pathogen recognition and immune response. Furthermore, considering the diverse pathogen environments in aquaculture, where cultured stocks always present much higher mortality rates compared to wild populations, there should be a much higher level of genetic differentiation at the MHC locus between cultured stocks and wild populations. However, we did not detect significant population differentiation at the MHC locus between the cultured and wild groups of orange-spotted grouper. This may be caused by the short history of aquaculture practices, where the accumulation of genetic divergence under selection pressure has been high enough to be detected (Hung et al., 2009). On the other hand, the effect of balancing selection on MHC class II exon 2 as revealed in our study could be an important reason for such non-significant differentiation, as balancing selection is considered to reduce the level of differentiation (Beaumont and Balding, 2004).

### CONCLUSIONS

In conclusion, we investigated allelic polymorphisms and tested the divergence patterns of MHC class II exon 2 within and between cultured stocks and wild populations of orange-spotted grouper from the South China Sea and Southeast Asia, which provided new insights into the genetic diversity and pathogen-driven selection of this species in this geographical region. This is also the first report on MHC polymorphisms and selection patterns in grouper species. Our results suggest that the MHC class II  $\beta$  gene can be used to monitor

genetic variation between cultured stocks and wild populations of grouper species. Importantly, further information about the status of wild populations and their susceptibility to pathogens also can be acquired by continuous investigation of the genetic changes of MHC gene diversity.

## ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation (#31001112, #30871914), the National Key Technology R&D Program of China (#2007BAD29B03), Fundamental Research Funds for the Central Universities (#111gpy66), and the National High Technology R&D Program 863 of China (#2009AA09Z401, #2012AA10A414).

## REFERENCES

- Alcaide M, Edwards SV, Negro JJ, Serrano D, et al. (2008). Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Mol. Ecol.* 17: 2652-2665.
- Anisimova M, Nielsen R and Yang Z (2003). Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 164: 1229-1236.
- Anmarkrud JA, Johnsen A, Bachmann L and Lifjeld JT (2010). Ancestral polymorphism in exon 2 of bluethroat (*Luscinia svecica*) MHC class II B genes. *J. Evol. Biol.* 23: 1206-1217.
- Axtner J and Sommer S (2007). Gene duplication, allelic diversity, selection processes and adaptive value of MHC class II DRB genes of the bank vole, *Clethrionomys glareolus*. *Immunogenetics* 59: 417-426.
- Beaumont MA and Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13: 969-980.
- Bernatchez L and Landry C (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* 16: 363-377.
- Brown JL and Eklund A (1994). Kin recognition and the major histocompatibility complex: an integrative review. *Am. Nat.* 143: 435-461.
- Brown JH, Jardetzky TS, Gorga JC, Stern LJ, et al. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364: 33-39.
- Clarke B and Kirby DR (1966). Maintenance of histocompatibility polymorphisms. *Nature* 211: 999-1000.
- Davies CJ, Andersson L, Mikko S, Ellis SA, et al. (1997). Nomenclature for factors of the BoLA system, 1996: report of the ISAG BoLA Nomenclature Committee. *Anim. Genet.* 28: 159-168.
- Dengjel J, Schoor O, Fischer R, Reich M, et al. (2005). Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. *Proc. Natl. Acad. Sci. U. S. A.* 102: 7922-7927.
- Eklblom R, Saether SA, Jacobsson P, Fiske P, et al. (2007). Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Mol. Ecol.* 16: 1439-1451.
- Excoffier L, Laval G and Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.
- Goudet J (2002). Fstat, A Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3.2). Available at [<http://www2.unil.ch/popgen/softwares/fstat.htm>]. Accessed February 8, 2012.
- Grimholt U, Larsen S, Nordmo R, Midtlyng P, 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.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.
- Hauswaldt JS, Stuckas H, Pfautsch S and Tiedemann R (2007). Molecular characterization of MHC class II in a nonmodel anuran species, the fire-bellied toad *Bombina orientalis*. *Immunogenetics* 59: 479-491.
- Hedrick PW (1999). Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313-318.
- Hill AV (1999). The immunogenetics of resistance to malaria. *Proc. Assoc. Am. Physicians* 111: 272-277.
- Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, et al. (1991). Common west African HLA antigens are associated with protection from severe malaria. *Nature* 352: 595-600.

- Hughes AL and Nei M (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335: 167-170.
- Hughes AL and Yeager M (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annu. Rev. Genet.* 32: 415-435.
- Hung PH, Thuy TTN, Supawadee P and Uthairat N (2009). Microsatellites revealed no genetic differentiation between hatchery and contemporary wild populations of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878) in Vietnam. *Aquaculture* 291: 154-160.
- Jones EY, Fugger L, Strominger JL and Siebold C (2006). MHC class II proteins and disease: a structural perspective. *Nat. Rev. Immunol.* 6: 271-282.
- Jordan WC and Bruford MW (1998). New perspectives on mate choice and the MHC. *Heredity* 81: 239-245.
- Karaïskou N, Moran P, Georgitsakis G, Abatzopoulos TJ, et al. (2010). High allelic variation of MHC class II alpha antigen and the role of selection in wild and cultured *Sparus aurata* populations. *Hydrobiologia* 638: 11-20.
- Klein J (1986). Natural History of the Major Histocompatibility Complex. Wiley, New York.
- Nei M and Gojobori T (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3: 418-426.
- Piertney SB and Oliver MK (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity* 96: 7-21.
- Rokas A, Atkinson RJ, Webster L, Csoka G, et al. (2003). Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp *Andricus quercustozae*. *Mol. Ecol.* 12: 2153-2174.
- Rozas J, Sanchez-Del Barrio JC, Messeguer X and Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.
- Schad J, Dechmann DK, Voigt CC and Sommer S (2011). MHC class II DRB diversity, selection pattern and population structure in a neotropical bat species, *Noctilio albiventris*. *Heredity* 107: 115-126.
- Strand T, Westerdahl H, Høglund J, Alatalo V, et al. (2007). The Mhc class II of the Black grouse (*Tetrao tetrix*) consists of low numbers of B and Y genes with variable diversity and expression. *Immunogenetics* 59: 725-734.
- Summers K, Roney KE, da Silva J, Capraro G, et al. (2009). Divergent patterns of selection on the DAB and DXB MHC class II loci in *Xiphophorus* fishes. *Genetica* 135: 379-390.
- Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.
- Tong JC, Bramson J, Kanduc D, Chow S, et al. (2006). Modeling the bound conformation of *Pemphigus vulgaris*-associated peptides to MHC Class II DR and DQ alleles. *Immunome Res.* 2: 1.
- Wang L, Meng Z, Liu X, Zhang Y, et al. (2011). genetic diversity and differentiation of the orange-spotted grouper (*Epinephelus coioides*) between and within cultured stocks and wild populations inferred from microsatellite DNA Analysis. *Int. J. Mol. Sci.* 12: 4378-4394.
- Westerdahl H, Wittzell H, von Schantz T and Bensch S (2004). MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity* 92: 534-542.
- Xu TJ, Sun YN and Chen SL (2010). Allelic variation, balancing selection and positive selected sites detected from MHC class Ia gene of olive flounder. *Genetica* 138: 1251-1259.
- Yang Z (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13: 555-556.
- Yang Z, Wong WS and Nielsen R (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22: 1107-1118.