Contents lists available at SciVerse ScienceDirect

Estuarine, Coastal and Shelf Science

journal homepage: www.elsevier.com/locate/ecss

Diversity of immune genes and associated gill microbes of European plaice *Pleuronectes platessa*

K. Mathias Wegner[a,b,c,*], Lisa N.S. Shama[a,b], Florian Kellnreitner[b], Moritz Pockberger[b]

[a] *Leibniz Institute for Marine Sciences (IFM-Geomar), Evolutionary Ecology of Marine Fishes, Düsternbroker Weg 20, 24105 Kiel, Germany*
[b] *Alfred Wegener Institute for Polar and Marine Research, Wadden Sea Station Sylt, Hafenstrasse 43, 25992 List/Sylt, Germany*
[c] *Institute of Integrative Biology, Experimental Ecology, ETH Zürich, Unversitätstrasse 16, CH-8047 Zürich, Switzerland*

ARTICLE INFO

ABSTRACT

Genetic variability of marine fish species is much higher than in most other vertebrates. Nevertheless, some species with large population sizes including flatfish such as European plaice *Pleuronectes platessa* show signs of population collapse and inbreeding. Taking plaice as a flagship example for fisheries-induced genetic changes also affecting the Wadden Sea, we determined the amount of genetic variability at antigen-presenting genes of the Major Histocompatibility Complex (MHC) and its potential interaction with the microbiota associated to gill tissue using a next-generation parallel tag sequencing approach. Genetic variation at MHC class IIB genes was extremely large, with 97 alleles found in 40 fish from different age cohorts. Although a strong signal of positive selection was present ($dN/dS = 4.01$) and we found significantly higher allelic diversity in 0+ fish than in older age classes, the amount of genetic variation maintained within the population may not have exceeded neutral expectations derived from mitochondrial markers. Associated microbes revealed significant spatiotemporal structure with 0+ fish displaying the highest microbial diversity as well as the highest diversity of potentially pathogenic genera. Overall the correlation between MHC genotypes and bacterial abundance was weak, and only few alleles significantly correlated with certain bacterial genera. These associations all conferred susceptibility (i.e. presence of an allele correlated to higher number of bacteria), either suggesting age-dependent selection on common alleles or weak selection on resistance against these bacterial genera. Taken together, our data suggest that selection coefficients of balancing selection maintaining immunogenetic diversity may be relatively small in large marine populations. However, if population sizes are further reduced by overharvesting, the response to increasing balancing selection coefficients will be largely unpredictable and may also negatively influence the adaptive potential of populations.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The European Wadden Sea is an important nursery ground for many fish species including commercially important flatfish species such as European plaice *Pleuronectes platessa* (van Beek et al., 1989). Nursery and spawning grounds can be considered as one of the major factors determining population structure. Use of separated nursery grounds or migration of larval cohorts from separated adult spawning grounds into the same nursery areas can result in structure within or between nursery grounds (Hovenkamp, 1991; Bolle et al., 2009; van der Veer et al., 2009). Both sources could potentially increase genetic differentiation between spawning/nursery grounds, however,

no significant genetic structure in *P. platessa* could be detected throughout the entire North Sea continental shelf using nuclear (Hoarau et al., 2002) or mitochondrial markers (Hoarau et al., 2004). Rather, continental shelf populations were highly polymorphic for both marker types suggesting that effective population sizes are very large and that random genetic drift contributes only little to population differentiation. Despite these ample amounts of genetic variation, overexploitation has led to signs of inbreeding in contemporary populations when compared to historic samples (Hoarau et al., 2005), indicating that the loss of functionally relevant genetic variation might impact population persistence.

On the other hand, adaptive genetic divergence shaped by natural selection often exceeds neutral divergence (McKay and Latta, 2002), even in the face of high rates of gene flow and large population sizes as commonly observed in marine species (Luttikhuizen et al., 2003). Despite this general pattern, only little is known about genetic diversity of genes under selection. The prime

example for selectively maintained genetic variation in vertebrates are antigen-presenting genes of the major histocompatibility complex (MHC) (Bernatchez and Landry, 2003), which have also been characterised in many teleost species (Wegner, 2008). One characteristic feature of nearly all species investigated is the striking polymorphism of these genes. Since alleles usually show elevated non-synonymous to synonymous substitution rates (dN/dS), balancing selection is thought to maintain the high polymorphism. MHC genes form a key component of the adaptive immune response by presenting peptide antigens derived from pathogens to T cells. Since higher individual MHC diversity means better immunological surveillance, heterozygote advantage is an obvious mechanism of balancing selection that can maintain genetic diversity (Doherty and Zinkernagel, 1975), and could explain resistance against several disease agents in many vertebrate taxa, for example, salmonids (Arkush et al., 2002; Turner et al., 2007). MHC genes are also poly-morphic in flatfish species, and can partly account for resistance against the widespread bacterial fish pathogen *Vibrio/Listonella anguillarum* in Japanese Flounder *Paralichthys olivaceus* (Zhang et al., 2006). Such bacterial pathogens may cause high mortalities not only in flat fish (Zheng et al., 2004; Zhang et al., 2006), but also in fish in general (Austin and Austin, 2007).

Opportunistic bacterial pathogens often express a temperature correlated virulence (Zheng et al., 2004), and rising ocean surface temperatures are likely to increase disease severity especially in shallow coastal systems. In light of these projected changes, the evolutionary potential of such populations in terms of their standing genetic variation in immunity will be crucial to prevent population decline, especially for species with high juvenile mortality and type III survival curves. Therefore, we investigated the correlation between diversity of pathogenic bacteria and diversity of the host's immune system in different age cohorts of Wadden Sea European plaice. We first characterised the immuno-genetic diversity at MHC class IIB genes and the associated microbes each fish was exposed to by parallel deep sequencing (Margulies et al., 2005; Binladen et al., 2007; Babik et al., 2009; Galan et al., 2010) of individually tagged MHC class IIB and bacte-rial 16S rRNA amplicons in gill samples of plaice stemming from the Sylt-Rømø bight in the Northern Wadden Sea. MHC diversity has been linked indirectly to open-water microbial diversity in Atlantic salmon *Salmo salar* (Dionne et al., 2007) but a direct connection between MHC genotype and microbial colonisation is so far lacking (Ley et al., 2008). It is nevertheless conceivable that such coevolu-tion with mutualistic and pathogenic bacteria will also harbour an MHC-dependent component. Specifically, we chose 0+ fish from a Wadden Sea nursery ground and compared those with older, open-water fish to determine whether allele frequency changes between cohorts indicate the action of selection, and whether this change can be attributed to the presence of different pathogenic bacteria associated with different life stages. Therefore, we can determine not only the bacterial and pathogenic diversity encountered at different sites during an individual's life time, but can also attempt to connect this diversity to diversity in selectively relevant genes of the immune system, and thereby identify potential causative agents of balancing selection in a genetically highly variable, yet threatened marine fish species.

2. Material & methods

2.1. Fish sampling

Our main focus was to compare juveniles (0+ fish) with plaice of higher age classes (1+ fish), therefore, we employed two different sampling techniques. A cohort of juvenile fish was caught using a beach seine in a tidal gully in Blidsel bight at low tide (Table 1). Older age classes (1+ fish) were caught by trawling using a vessel at three sites covering a depth range from ∼2 m (station 3 & 5) to ∼5 m (station 8, see Table 1 for details). We tried to balance sample sizes between juvenile and older fish, but were somewhat limited by the availability of fish from trawls; hence sample size shows some seasonal variation due to low numbers of plaice in trawls. The available fish nevertheless reflect a representative sample of the fished sites. After catching, young fish were killed in an excess of MS222, and older fish were killed by concussion and immediately frozen at −20 °C until further use.

On the day of dissection, fish were thawed, weighed to the nearest milligram, and overall length was measured before a gill arch was removed for DNA extraction. We chose to use a gill arch because gills represent a large surface contact area and have shown immunological activation by elevated MHC class IIB transcript levels in other fish species (Wegner et al., 2006). DNA was extracted with the Wizard Genomic DNA purification system (Promega, Mannheim, D) following the manufacturer's instructions. Quality and concentration of DNA was checked by photometric measure-ment using a Nanodrop 1000 (Thermo Scientific, Wilmington, USA) spectrophotometer. One aliquot of each DNA extract was diluted to a concentration of 30 ng μl^{-1} for microbial fingerprints while another aliquot was taken undiluted for MHC genotyping.

2.2. Genotyping of MHC and microbial diversity

To measure immunogenetic diversity of MHC class IIB genes we used a 189 bp (excluding primer sequence) PCR product amplifying within the exon 2, which is the most polymorphic region of the MHC class IIB gene in teleost fish (Wegner, 2008). Primers were designed from the sequence of a 3′-RACE-PCR product previously obtained by following the SMART-RACE protocol (Clontech, Saint-Germain-en-Laye, F) using a conserved flatfish MHC class IIB forward primer (Zhang et al., 2006). This PCR product covered the region from the start of the exon 2 to the 3′-end of the transcript. Fragments of the correct expected size were excised from the gel and cloned into the TOPO-TA (Promega, Mannheim, D) vector before transformation into electro-competent cells. Inserts were amplified using M13 primers and sent out for sequencing (Microsynth, Balgen, CH). Sequences were aligned with Japanese Flounder sequences (Zhang et al., 2006) to identify putative exon–exon boundaries, which we then used to design internal exon 2 primers in conserved parts of the gene. PCR reactions contained 1 μl of undiluted DNA extract, 0.05 μl of Hot start taq (5 units μl^{-1}, Promega, Mannheim), 1 mM of each dNTP, 4 μl of reaction buffer (Promega, Mannheim), 1 μl of each forward and reverse primer (5 μM) and 11.5 μl of molecular grade water. After initial denaturation and hot-start activation for 10 min at 95 °C, we used a comparatively low number of 24 PCR cycles consisting of 30 s denaturation at 95 °C, 30 s

Table 1
Characteristics of sampling stations and sampled plaice.

Site	Coordinates (depth [m])	Date	Temp. [°C]	Salinity	Length [cm] mean ± sd	Weight [g]	Condition index[a]
Blidsel bight	54°59.38′N, 8°23.07′E (0.5 m)	18.05.09 (n = 22)	12.4	27.6	4.79 ± 0.90	1.39 ± 0.71	−0.20 ± 0.10
Station 3	55°01.59′N 8°26.81′E (2 m)	21.04.09 (n = 5)	10.8	28.0	9.53 ± 1.47	9.91 ± 4.63	−0.50 ± 0.49
Station 5	54°59.50′N 8°25.69′E (2 m)	22.04.09 (n = 4)	11.9	27.7	11.46 ± 3.12	18.92 ± 16.56	−0.04 ± 0.77
Station 8	55°04.25′N 8°29.71′E (5 m)	17.02.09 (n = 9)	1.3	29.9	8.8 ± 0.76	6.60 ± 1.99	0.54 ± 0.33

[a] Condition index was calculated as the deviation from the cubic function regressing weight on length.

annealing at 60 °C and 90 s extension at 72 °C. A final elongation of 7 min ended the PCR. Each individual fish was tagged by a unique combination of multiplex identifiers (MID, Roche, Basel, CH) integrated into forward and reverse primers (Binladen et al., 2007). We used a total of 20 tagged primers consisting of a modified 454 sequencing adaptor (GATC Biotech, Konstanz, D), the 454 sequencing key, a MID tag and the gene-specific sequence. Hence, an example of a forward primer would have the following sequence: 5′-CCATCT-CATCCCTGCGTGTCTCCGAC *TCAG* **ACGAGTGCGT** GAGCTGAACGG-CATTGAGTACA-3′, while a reverse primer would look like this: 5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG *TCAG* **ACGAGTGCGT** CTTCGT CAGAGCACTTTCTG-3′, with the 454 sequencing key in italics, the MID tag in bold and gene specific sequence underlined. We used MID tags 01 -11 (Roche, Basel). In addition to primers containing MID tags, we also used primers that did not contain MID tags to evaluate the effect of MID tags on amplification bias. To detect artefacts originating from in vitro PCR recombination, each individual was amplified in two independent reactions containing distinct MID combinations. We specifically used undiluted genomic DNA to generate a wide range of individual coverage in order to be able to quantify the effect of different MID tags on the performance of each PCR reaction and the effect of coverage on allele assignment.

In contrast to MHC products, microbial diversity was measured with genomic DNA of a constant concentration (30 ng μl^{-1}) to allow for between sample comparisons of bacterial abundances. We used an informative region of the 16S rRNA gene spanning position 535-904/912 in the *Escherichia coli* genome covering the variable regions V3 and V4 for ribotyping. Using a PCR product of this length increases the precision of taxonomic assignment (Huber et al., 2009) and should provide high resolution due to the inclusion of two variable regions. To avoid cross-amplification of host 18S rRNA, we performed a nested PCR starting with primers specific for Eubacteria (27f: AGA GTT TGA TCA TGG CTC AG and 1492r: GGT TAC CTT GTT ACG ACT T (Suzuki and Giovannoni, 1996)). The first nested PCR consisted of 30 ng of genomic DNA, 0.05 μl of Hot start taq (5unit/μl, Promega), 1 mM of each dNTP, 4 μl of reaction buffer (Promega), 1 μl of each forward and reverse primer (5 μM) and 11.5 μl of molecular grade water. Cycling started with an initial denaturation and hot start activation of 10 min at 95 °C followed by a low number of 16 cycles of 30 s denaturation at 95 °C, 30 s at 50 °C and 90 s at 72 °C and a final extension of 10 min at 72 °C. One μl of each PCR product was then diluted in 99 μl of molecular grade water before the internal stretch was amplified using 454 tagged primer combinations analogous to the MHC amplicon approach above, the only difference being that gene specific primer sequences were 16S_intF 5′-CCAC-GAGCCGCGGTAAT-3′ and 16S_intR 5′-CCGTCAATTCMTTTAAGTTT-3′.

After amplification, single PCR reactions were purified using the MinElute 96 kit (Qiagen, Hilden) before 2 μl of each elution was used for pooling. To eliminate remaining primer-dimer, both pools were purified again using the Wizard PCR clean-up system (Promega, Mannheim) following the manufacturer's instructions. After confirming the sole presence of the desired PCR product without any traces of primer by gel electrophoresis, the concentration of both pools was measured. To increase relative coverage of the 16S rRNA product, we combined both pools at a 1:7 (MHC:16S) ratio and adjusted the amount and concentration to the desired 20 ng/μl using a speed vac. Both microbial and MHC class IIB genotyping was done by deep amplicon sequencing of individually barcoded PCR reactions on the 454 FLX genome sequencer (Roche, Basel, CH) using Titanium chemistry. Sequencing was performed by GATC biotech (Konstanz, D).

2.3. Data analysis

Assignment of reads to individual PCRs was done using in-house perl scripts. In short, within each raw read we first looked for either gene specific primer and then identified MID tags. In cases where MID could not be identified due to potential sequencing errors, one base at a time within putative MID regions was mutated until a correct MID was found. This strategy makes use of the fact that any pairwise difference between two MID tags used here was at least six point mutations, and enabled us to use 5% (=7076 reads) more reads for the subsequent analysis. After correct assignment of single reads to individuals, primer sequences were trimmed and the remaining sequence was added to a single fasta file representing MHC or 16S rRNA reads for each individual respectively.

2.3.1. MHC genotyping

Individual MHC genotypes were determined by sorting the observed sequences of the first PCR according to their frequency found within that PCR. We filtered artefacts by an absolute (<3 reads per sequence) and a relative cut-off (<5% of total coverage of that PCR reaction), representing previously used values in systems with similar genomic complexity (Kloch et al., 2010). This was repeated for the replicate PCR of the same individual. Only common sequences occurring in both replicate PCRs were used for further analysis. This approach combines previous artefact detection methods which mainly applied frequency/divergence statistics to single out sequencing errors (Babik et al., 2009; Kloch et al., 2010) with the more conservative approach of allele calling from independent PCR reactions for each individual thus representing the "gold standard" for MHC genotyping (Lenz and Becker, 2008). The set of confirmed alleles was aligned using ClustalX and the signature of selection (dN/dS) was detected using MEGA 4.1 (Tamura et al., 2007) for the total sequence using the Nei-Gojobori method and Jukes-Cantor correction. For codon based estimates, we used omegaMap (Wilson and McVean, 2006) which applies Bayesian Markov Chain Monte Carlo procedures to estimate the dN/dS ratio. We used improper inverse priors and ran two independent chains of 100 000 iterations using a window size of three codons from which we sampled every 50th estimate after removing a burnin of 20 000 iterations. To look for small scale local structure in MHC genotypes, we ordinated individual genotypes by correspondence analysis implemented in the ADE 4 package of the *R* statistical environment (R Development Core Team, 2010), only considering common alleles of a frequency >5%. Significance of the observed spatial structure was tested by 999 randomisations implemented in class-specific methods of *randtest* (ADE 4). To get an idea of how well we covered the total allelic diversity present in the population, we performed rarefaction analysis on the number of alleles found given a set of 100 random individual resamples stemming from the whole data set or each cohort separately. We fitted an asymptotic function of the shape $N * (1 - \exp^{x/a})$, where we estimated the shape parameter a and the asymptotic value N corresponding to expected total allelic diversity.

2.3.2. Microbial diversity

We used mothur (Schloss et al., 2009) to align all 16S rRNA reads found for each individual plaice and to identify and remove chimeric sequences using the *chimera slayer* routine. Assignment of cleaned 16S rRNA reads to taxonomic units was done using a stand-alone version of RDB Classifier 2.1 (Wang et al., 2007), which gives a more robust assignment than BLAST analysis against GenBank (Liu et al., 2008). Using an assignment certainty threshold of 50% for each taxonomic level, we generated abundance tables for all taxonomic levels ranging from phylum to genus, which were used for further analysis. We specifically did not evaluate the number of OTUs in the form of unique reads, as assignment to taxonomic levels higher than genus is often ambiguous. MEGAN (Huson et al., 2007) was used to visualise taxonomic distributions for different trawl sites. Since we were mainly interested in the abundance of potentially pathogenic bacterial genera, we generated an additional

abundance table containing only genera previously described to contain pathogenic species (Austin and Austin, 2007). Similar to the MHC ordination above, we also ordinated individual microbial fingerprints by principal component analysis to detect spatial structure in associated bacteria using log-transformed read counts for each individual fish. Log-transformed read counts were also used for calculating Shannon's diversity index H, as well as for calculating total and pathogenic load. The influence of body condition and sampling locality on these summary statistics were calculated by linear models. The direct link between MHC genotype and bacterial abundance matrices was tested by co-inertia analysis (Dray et al., 2003), which has previously been used to link large MHC and parasite matrices (Deter et al., 2008; Tollenaere et al., 2008), and offers the flexibility to correlate a continuous matrix of bacterial abundance to a factorial matrix of MHC allele presence.

3. Results

3.1. Immunogenetic diversity

We could assign 60 468 sequence reads to either one of the replicated PCRs of the 40 individuals used. The average coverage per PCR was 703.01 ± 43.67 reads. Average coverage of a called allele was 378.82 ± 44.24, while discarded alleles had an average coverage of 3.90 ± 0.26, showing that we obtained deep enough coverage for reliable allele calls. The correlation between replicate PCRs was strong on the allele level ($R^2 = 0.93$, slope $= 0.98$) indicating that bias introduced by different tag combinations was small and both PCRs independently led to very similar results. We identified 97 MHC class IIB alleles among the 40 individuals of European plaice. On average, 3.53 ± 0.13 sequence variants could be recovered per individual with most fish (58%) having four different alleles, suggesting that we typed two MHC class IIB loci that were mostly heterozygous due to a high degree of polymorphism. Most alleles were singletons ($n = 74$), while the most frequent allele reached a frequency of 17%. Both loci seemed to be functional as none of the sequences showed stop codons or other signs typical for pseudogenes. Due to low statistical support for distinct groups of sequences, assignment to putative loci was not possible (data not shown), again reflecting previous results where locus affiliation could only be detected by intronic rather than exonic sequences (Reusch and Langefors, 2005). We found a pattern of strong positive selection (dN/dS = 4.08) within the total set of sequences. Looking at codon resolution it becomes obvious that some sites of the gene experience strong positive selection while others rather experience purifying selection (Fig. 1A), which is the typical picture observed for MHC class IIB genes, and most likely reflects sites that are involved in antigen binding versus sites that are important for structural integrity of the protein.

We did not find significant differentiation between cohorts or groups based on alleles with >5% frequency (inertia of correspondence analysis: 0.086, $p = 0.308$ obtained from 999 random permutations), supporting previous results reporting no genetic differentiation on large geographical scales as well as between cohorts (Hoarau et al., 2002). Rarefaction analysis of the existing data set showed that a total of 176 sequences could be expected in the whole population and that a sample size of >250 individuals would be needed to reach saturation (Fig. 1B). Separating the analysis between 0+ and 1+ age classes indicated that diversity was substantially higher in 0+ fish than in older age classes (asymptotic values for 0+ = 258.4 and 1+ = 105.8, respectively, Fig. 1B). This difference originated from a higher frequency of singleton alleles in 0+ fish. Consequently, with 0.81 ± 0.15 common alleles (frequency >5%) per individual, 0+ fish had significantly less common alleles than 1+ fish (1.56 ± 0.27, $F_{1,38} = 6.72$, $p = 0.014$).

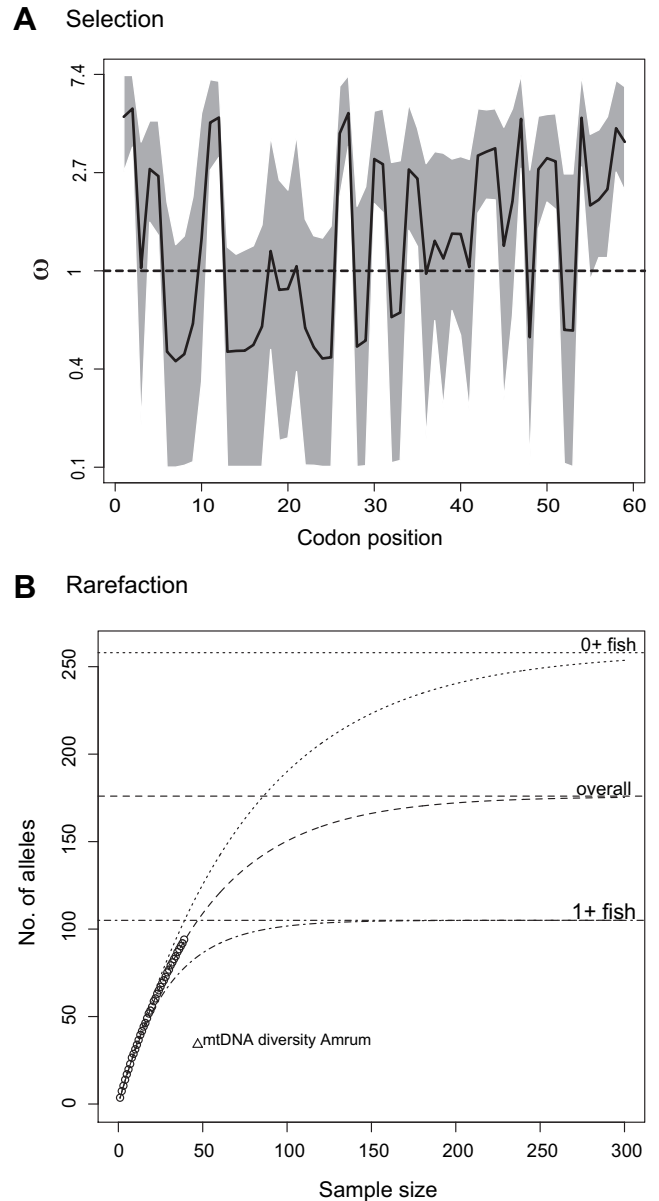**Fig. 1.** Selection and diversity in European plaice MHC class IIB genes. A) Signature of selection across codons expressed as the non-synonymous to synonymous substitution rate (ω = dN/dS). The solid line gives the highest posterior density for each codon position and the shaded area represents the 95% C.I. The dashed line shows the rate of $\omega = 1$ representing neutral evolution. B) Relationship of sampling effort to allele recovery using a rarefaction analysis. Circles show the means of 100 random resamples of the whole data set with a given sample size. The dashed line shows the best fit function of $y = 175.81 * (1 - e^{(-x/51.75)})$ and the horizontal line shows the asymptotic value of the expected number of alleles present in the population. The dotted lines show the rarefaction curve and asymptotic value obtained when only 0+ fish were considered ($y = 258.40 * (1 - e^{(-x/75.23)})$) and the dash-dotted lines show the corresponding results for 1+ fish ($y = 105.04 * (1 - e^{(-x/28.94)})$), indicating that allelic diversity was higher in 0+ fish. The triangle shows haplotype diversity of the neutral mitochondrial D-loop marker used in Hoarau et al. (2004) for comparison.

Observed diversity of MHC class IIB genes was substantially higher than diversity of mitochondrial D-loop sequences of plaice stemming from a nearby location, the island of Amrum, (Hoarau et al., 2004, Fig. 1B). But this may probably not support contemporary balancing selection maintaining MHC diversity, as effective size of the two nuclear MHC loci will be roughly eight times higher than that of a mitochondrial marker.

3.2. Microbial diversity

We obtained a total of 71 430 reads for the 16S rRNA PCR product. After sequence cleaning we could assign 40 210 reads to a bacterial genus with certainty >0.5. We observed a high microbial diversity covering a total of 304 bacterial genera out of which 23 were previously described to contain pathogenic species/strains. The most commonly observed genus was *Sphingomonas* (α-proteobacteria) with a high proportion of 71% of all assignable reads. α-proteobacteria were also found to be the most common order. Out of the 276 genera, 269 only occurred rarely with a frequency $<1\%$ of all reads, reflecting inherent bias towards the "rare biosphere" in marine microbial communities (Kemp and Aller, 2004; Pedros-Alio, 2006). High abundance of rare phylotypes potentially leads to under sampling of the true diversity at higher taxonomic levels. Therefore, Fig. 2 shows the taxonomic distribution of all bacteria at the order level for all stations and higher taxonomic resolution was only generated for potentially pathogenic genera (shaded area of Fig. 2). The orders showing the strongest differences among sites were *Caulobacterales*, *Pseudomonales* and *Burkholderiales* dominating site no. 8, while *Thiotrichales*, *Alteromonodales* and *Hydrogenophilales* dominated tidal channel samples from Blidsel bight (Fig. 2). A formal test of these differences using a discriminate analysis on the principal components of order level diversity confirmed that microbial communities differed significantly among sites (Fig. 3A). Communities at stations no. 3 and 5 were similar, probably reflecting the similar catch dates and therefore similar abiotic conditions in terms of temperature (Table 1). Due to similar depth profiles at these stations, physical conditions (e.g. currents, tides) reflect additionally shared parameters that may have resulted in higher community similarity. All other stations differed significantly from each other whereas variation within sites was comparatively small (Fig. 3A). The overall diversity was highest in Blidsel bight and lowest at station no. 3 (Fig. 3C), whereas no significant difference was found for load calculated as the natural logarithm of the observed sequence reads per individual (Fig. 3E).

Similarly, the abundance of bacterial genera containing potentially pathogenic strains differed significantly among sites with *Acinetobacter* being most common in station no. 8, while *Photobacterium* and *Mycoplasma* showed high abundances in station no. 5 and *Francisella* was nearly exclusively found in Blidsel bight (Fig. 2). Consequently, the ordination also showed site-specific grouping strongly differentiating among station no. 5, no. 8 and no. 3/Blidsel (Fig. 3B). As with total bacterial diversity, pathogen diversity was also highest in Blidsel bight, although this pattern was less pronounced (Fig. 3D) and we did not find any significant difference among sites in potential pathogen load (Fig. 3F).

We did not find a significant relationship between body condition and pathogen or total bacterial load, nor between body condition and any diversity measure of pathogen or general bacterial richness (all p values > 0.05). Pathogen load also did not correlate significantly with total load ($F_{1,38} = <0.01$, $p = 0.981$). For the co-inertia analysis linking MHC matrices (cases: individual fish, variables: presence/absence of alleles) to bacterial matrices (cases: individual fish, variables: abundance of bacterial genera) we only considered common MHC sequence variants ($>5\%$ in the total sample) and concentrated on the subset of potentially pathogenic bacteria to achieve at least a low level of statistical replication. Overall, the two matrices correlated significantly to each other (RV: 0.349, $p = 0.041$ obtained from 999 resamples), which is not self-evident given that alleles can either confer resistance or susceptibility. To detect potential associations between abundance of bacterial genera and MHC alleles responsible for the overall matrix correlation, we used the normed column scores of both matrices. Fitting univariate linear models, we found 20 significant

associations when evaluating all 276 possible combinations of alleles and bacteria, which corresponds to 7% of all tests, and thus mainly reflects the false-positive discovery of multiple testing. Considering only normed column scores matching between both matrices (i.e. positive/negative scores for both principal axes in both matrices) reduced the total number of tests from 276 to 69, out of which 15 tests were significant (=21%). All parameter estimates of the significant allele associations were positive indicating that these alleles did not confer resistance, which might explain the overall correlation between the matrices despite rather weak associations. Table 2 gives the statistical details of those associations remaining significant after adjustment of the false discovery rate (Benjamini and Hochberg, 1995).

4. Discussion

Balancing selection driving the polymorphism of immune genes is considered one prime example where genetic diversity exceeds neutral expectations. A case in point are classical MHC genes which have been found to be more polymorphic than under neutral expectations in many species (Bernatchez and Landry, 2003; Wegner, 2008) despite evolutionary forces like genetic drift eroding variability (van Oosterhout et al., 2006). But what if natural population sizes are so large like in many marine fish species (DeWoody and Avise, 2000; McCusker and Bentzen, 2010)? Exceeding the neutral expectation in such population calls for tremendous amounts of selectively maintained genetic variation. On the first look, with a total of 97 alleles in 40 individuals, the cohorts of European plaice *P. platessa* from the Northern Wadden Sea described here seem to fulfil this requirement. Especially given that our sampling effort was too small to fully characterise the amount of genetic variation at MHC class IIB genes in this population where a total number of >150 alleles can be expected (Fig. 1C).

The high dN/dS ratio additionally demonstrates that selection shaped this repertoire at parts of the gene putatively involved in antigen binding (Fig. 1A). Such a signature of selection might, however, represent past selection events and not necessarily contemporary balancing selection (Garrigan and Hedrick, 2003; Wegner, 2008). Indeed, when comparing diversity at the MHC genes typed here to haplotypic diversity of the mitochondrial D-loop of a nearby population (Hoarau et al., 2004), we found that MHC diversity may not have exceeded neutral expectations given that the two nuclear MHC loci and a single mitochondrial locus only differ by the constant factor of eight in N_e (Birky et al., 1989). On the other hand, the observed diversity was only 2.3 times higher for 1+ fish and 3.2 times higher for 0+ fish with comparable sampling effort (Fig. 1C). The amount of genetic variation maintained at a single locus is determined by $\theta = 4N_e u$, indicating that not only the effective size N_e is important, but also the mutation rate, which can be substantially higher in mitochondrial genes than in nuclear genes (Haag-Liautard et al., 2008). For MHC genes, intra- and intergenic gene conversion events are common (Zangenberg et al., 1995; Reusch and Langefors, 2005), probably increasing the mutation rate far above that of point mutation alone. A further similarity between variation found in plaice MHC class IIB genes and mitochondrial D-loop haplotypes is the large number of singleton sequences (76% for MHC vs. 64% for mtDNA). Consequently, the distribution of allele frequencies also looked similar, further indicating that large amounts of selectively neutral genetic variation can be maintained in large populations without selection acting on it.

Eventhough plaice show signs of genetic impoverishment due to overfishing (Hoarau et al., 2005), the lack of any population structure throughout the continental shelf coupled with high genetic
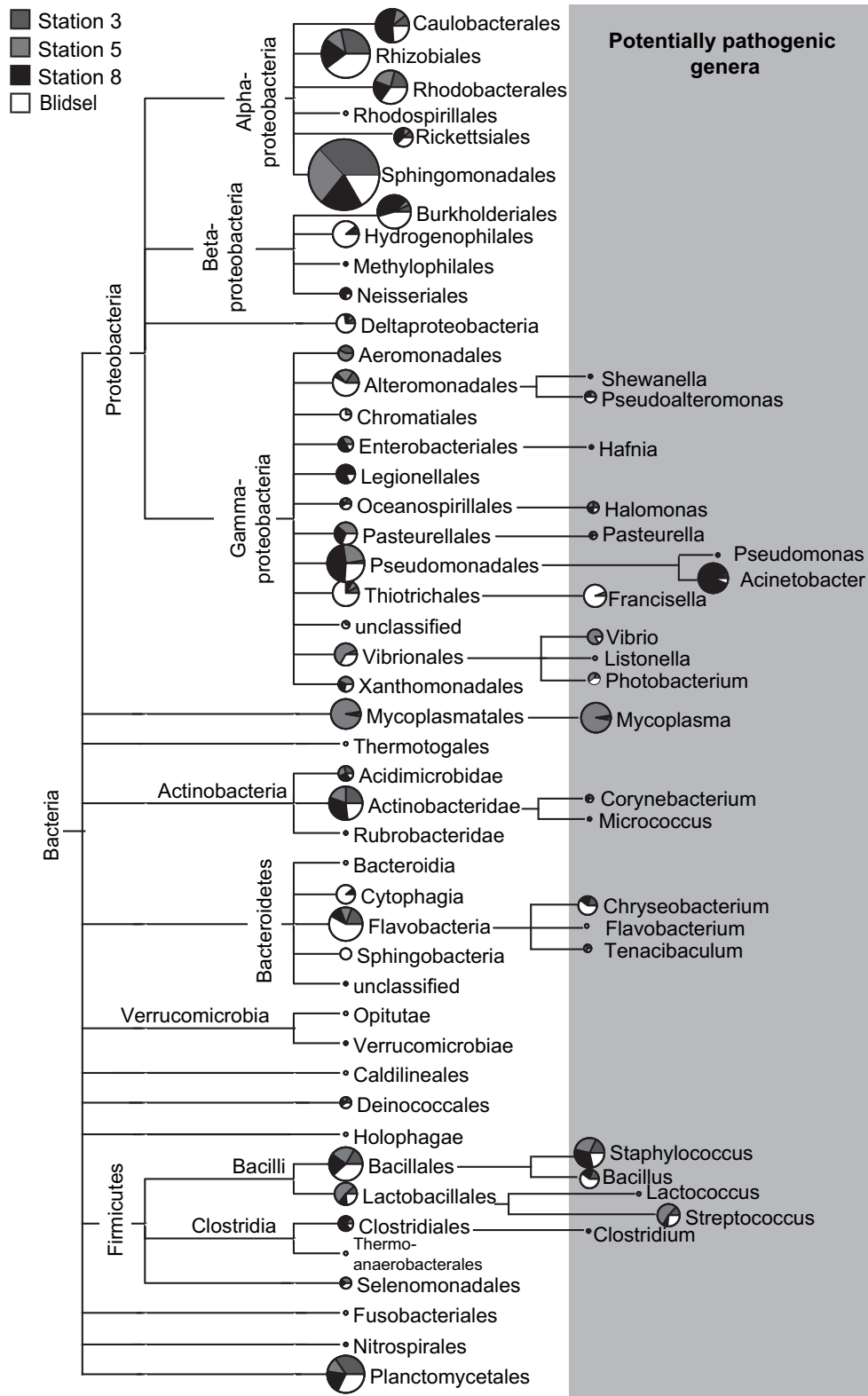
Fig. 2. Bacterial diversity on gills of European plaice determined by 16S rRNA deep amplicon sequencing. Total diversity is shown on the order level, while potentially pathogenic genera are shown on the genus level (shaded area). Different sites are coded by shades of grey.

variability indicates that effective size is still quite large. Drift might thus only play a minor role and balancing selection might pass by undetected in such populations. Selection coefficients of overdominant selection were found to be smaller than 0.01 in large populations of Trinidian guppies (N_e = 2400, van Oosterhout et al., 2006), which is considerably smaller than the effective size of N_e = 20 000 estimated for the North Sea population of plaice (Hoarau et al., 2005). Similarly, no selection could be detected on
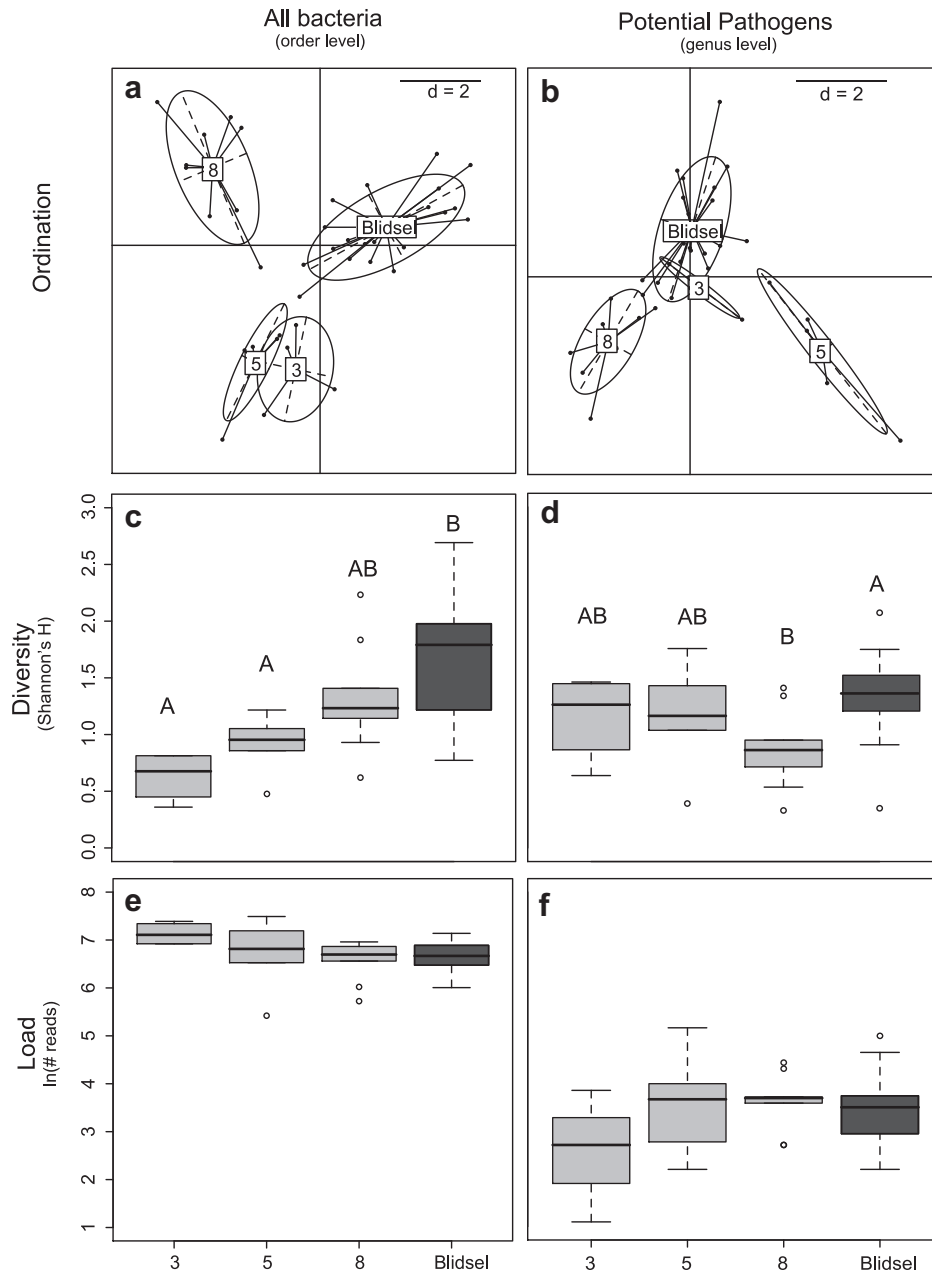
Fig. 3. Comparison of bacterial diversity in 40 230 successful taxonomy calls from gills sampled within and among different sites in the Sylt-Rømø bight representing different age cohorts. 0+ juveniles originated from Blidsel bight and older individuals were caught at the other 3 sites. a) Ordination of a principal component analysis on the order level for all bacterial taxa showing significant site specific bacterial assemblages (permutation test, 999 permutations, $p = 0.001$). Scale shows inertia (squared distance from the origin). b) Ordination of a principal component analysis on the genus level for potentially pathogenic bacteria also showing significant site specific bacterial assemblages (permutation test, 999 permutations, $p = 0.001$), yet showing different clustering than the total bacterial fauna. c) Bacterial diversity expressed as Shannon's H for all bacteria. Overall there was significant variation among sites ($F_{3.36} = 7.47$, $p < 0.001$). Sites not sharing a letter code were significantly different from each other. d) Shannon's H calculated for potentially pathogenic genera showing significant variation among sites ($F_{3.36} = 3.43$, $p = 0.027$). Sites not sharing a letter code were significantly different from each other e) Bacterial load expressed as log-transformed number of sequences per individual. No significant difference among sites was found ($F_{3.36} = 1.71$, $p = 0.184$) f) Pathogen load did not differ significantly among sites ($F_{3.36} = 1.59$, $p = 0.209$).

a marker linked to MHC class IIA genes of herring *Clupea harengus* (Andre et al., 2010), which probably possesses similar if not larger population sizes than plaice. This suggests that the observed allelic diversity in European plaice might reflect neutral diversity and not contemporary balancing selection maintaining diversity against erosion from drift.

Despite the lack of neutral marker population structure on larger geographic scales (Hoarau et al., 2002), we found that common MHC alleles of a frequency >5% occurred more frequently in 1+ fish than in 0+ fish from Blidsel bight. Consequently, allelic

diversity was larger in 0+ fish than in 1+ fish (Fig. 1B). Such different allelic compositions might indicate that different cohorts from several spawning grounds mix in nurseries of young fish (Hovenkamp, 1991), but split according to kin groups when migrating to adult feeding or spawning grounds. Such behaviour would, however, lead to differentiation on larger spatial and temporal scales, which obviously is absent (Hoarau et al., 2002). Such a pattern is also compatible with selection for beneficial alleles that increase in frequency in adults compared to juveniles, which is particularly relevant for species with type III survival

Table 2
Associations of MHC class IIB alleles with presence of potentially pathogenic bacterial genera.

Allele	Bacterial genus	Parameter estimate	F	Adjusted p value
Plpl DAB15	Flavobacterium	0.86	9.51	0.044
Plpl DAB20	Photobacterium	1.29	14.01	0.021
	Vibrio	0.34	9.70	0.044
Plpl DAB17	Vibrio	0.34	9.69	0.044
Plpl DAB32	Tenacibaculum	0.86	37.55	<0.001
Plpl DAB43	Lactococcus	0.42	11.07	0.044

p values were adjusted following (Benjamini and Hochberg, 1995).

curves characterised by high juvenile mortality rates. Although predation on larvae has been identified as one of the major sources of mortality accounting for up to 35% of juvenile mortality (van der Veer and Bergman, 1987), disease might still be another important factor. Due to the crucial role of MHC genes in mounting an immune response, pathogens seem to be likely agents of selection. Therefore, we focussed on microbiota found on gill tissue to determine whether a link between MHC genotypes and microbial colonisation especially by potentially pathogenic genera actually exists. We found significant differences in the communities of associated microbes (Fig. 3A) as well as in potential pathogens (Fig. 3B), even while acknowledging that assigning pathogenicity to a 16S rRNA phyloptype remains uncertain as many species within a genus as well as many strains within a species will not necessarily be virulent. Due to higher relatedness within a genus it is nevertheless conceivable that an immune response mounted against a pathogenic strain will also affect congeners to a larger degree and vice versa, a fact that has been utilised in fish vaccinations using live avirulent strains to gain protection against virulent strains (Norqvist et al., 1989).

Clustering of microbial communities according to station differed between total and pathogenic microbiota, indicating that pathogenic genera are not only a random subsample of total bacterial diversity and that environmental or host factors influence their distribution (Fig. 2). Total microbial community composition seemed to be influenced by ecological parameters because microbial communities found in fish from stations no. 3 and 5 had very similar community composition (Fig. 3A), probably owing to the fact that fish were caught at the same time and at a similar depth of ~2 m (Table 1). Community composition at the other stations was very different, but we cannot completely disentangle the effect of different catch dates and locality as they partly confound each other. Effects of both season as well as locality on a larger scale were shown to correlate with different bacterial communities (Wilson et al., 2008), and repeated sampling at these sites would be necessary to differentiate between both effects. It has nevertheless to be noted that catch dates and localities are representative for the respective size classes and it is therefore interesting to note that both total and pathogenic diversity was highest in the 0+ fish from Blidsel bight (Fig. 3C, D). This could partly be caused by the later sampling date, but similarly high diversity indices at station no. 5 along with a significantly different community structure (Fig. 3B, D) suggests that the high diversity in 0+ fish cannot only be attributed to late sampling.

Overall diversity was considerably larger than that for bacterial communities described by previous culture based (Liston, 1957) or PCR based (Wilson et al., 2008) studies of fish species, which can easily be explained by much higher per individual coverage using a next-generation sequencing approach. Diversity was largest in the γ−Proteobacteria (Fig. 2), which matches previous studies characterising bacterial diversity in flatfish (Liston, 1957) and cod (Wilson et al., 2008). In contrast to these studies, the numerically most abundant groups belonged to the α-proteobacteria, especially to the

genus Sphingomonas. α-proteobacteria have found to be dominant in open sea water (Galand et al., 2009), but the numerical abundance of α-proteobacteria in open sea water was partly attributed to PCR bias by preferential amplification of sequences from this taxonomic group (Gonzalez and Moran, 1997). We also cannot rule out the presence of PCR bias in our samples, but Sphingomonas sp. were previously found to be dominant on freshly prepared cod from Iceland (Reynisson et al., 2009), indicating that this genus can reach high numbers under natural conditions. Interestingly, some Sphingomonas isolates possess antimicrobial activities (Romanenko et al., 2008), and might thus represent a beneficial microbe protecting the host from infection by harmful bacteria. Alternatively, the fast disappearance on fish meat (Reynisson et al., 2009) might indicate that Sphingomonas sp. can benefit from the host's immune system, keeping superior competitors from the γ−Proteobacteria with pathogenic potential like Vibrio sp., Photobacterium sp., and Pseudomonas sp. at bay, which were otherwise shown to be common in fish mucus (Wilson et al., 2008).

Finding possible correlations between immune genes of the MHC and bacterial abundance was difficult because most alleles were singletons and no statistical replication across individuals was available for testing the entire data set. Instead, we concentrated on common alleles with a frequency >5% offering at least a minimal amount of replication. We found a significant overall relationship between the MHC matrix and the potential pathogen abundance matrix, which can probably be attributed to the fact that all alleles were positively associated with the abundance of certain bacterial genera, i.e. presence of the allele was correlated to higher bacterial loads (see Table 2 for association remaining significant after correcting for false discovery rate). That all associations had the same sign could indicate that the overall correlation of bacteria and MHC matrix was rather weak, owing to the fact that we did not find any protective alleles. Previous studies tended to find more protective alleles against bacterial disease agents. Resistance against Aeromonas salmonicida in Atlantic salmon could, for example, clearly be attributed to certain MHC class IIB alleles (Langefors et al., 2001; Lohm et al., 2002; Grimholt et al., 2003). Also in flatfish, associations between MHC alleles and bacterial infections have previously been described, and contained a mix of protective alleles and alleles associated with infection (Zhang et al., 2006).

Presence of alleles associated with infection is usually explained by pleiotropic effects when such an allele also gives protection against other pathogens, which then prevents selection from removing it from the population. The fact that no significant protective effect could be found here is consistent with a scenario of negative-frequency dependent selection, since we only considered common alleles, which are often over-parasitised as a result of parasites tracking common host genotypes (Dybdahl and Lively, 1998; Lively and Dybdahl, 2000). We cannot, however, test the potential advantage of rare alleles due to the aforementioned lack of replication, which highlights the need for experimental studies using breeding schemes to test for allele specific effects (Arkush et al., 2002; Zhang et al., 2006). As the microbiota were sampled from gill tissue, we automatically sampled a proportion of the water column, resulting in a few abundant bacteria specific to each fish type and many transient rare bacteria. Alternatively, the expressed virulence upon infection might be low in general and strongly depend on environmental conditions (Wegner et al., 2008), given that many of the identified genera contain opportunistic pathogens (Austin and Austin, 2007) and often express virulence only in the case of multiple infections (Pujalte et al., 2003). Consequently, low selection coefficients for resistance exerted by pathogenic strains on their own will not increase the frequency of protective alleles. As pathogen diversity did not decrease significantly towards older cohorts, the questions of why

common alleles are mainly observed in older cohorts of fish cannot be directly answered by infection patterns of the potential pathogens investigated here. Larger-scale phylogeographic studies and especially experimental studies including other disease agents would be needed to clarify patterns of selection acting on MHC genes in European plaice.

In conclusion, our results show the utility of next-generation sequencing techniques (Binladen et al., 2007; Babik et al., 2009; Galan et al., 2010; Kloch et al., 2010) to simultaneously type several markers from the same individual in parallel. The two markers chosen here can, in theory, be complemented by many more and result in detailed characterisation of individuals, thereby increasing the cost-efficiency of these approaches (Wegner, 2009). The strong correlation between the two independent replicate PCRs for MHC typing indicates that the genotype calls are robust and our technique reliably discriminates true alleles from artefacts. Our results further suggest that in large and highly polymorphic populations like European plaice and probably other marine fish species (Andre et al., 2010), selection coefficients of contemporary balancing selection maintaining MHC polymorphism may be small, because its main opposing selective force, genetic drift, is small as well. This does however not mean that there is no selective advantage for the individual resulting from high degrees of population-wide polymorphism, it simply is a consequence of balancing selection being different from directional selection. Directional selection works more efficiently in large populations, whereas selection coefficients of balancing selection decrease with population size once a large degree of polymorphism is achieved. Nevertheless, our data clearly demonstrates that selection shaped the functional diversity of the allelic repertoire (Fig. 1A) and may select for certain alleles in older age classes. Allele frequency changes observed on small, local scales might nevertheless be too small to overcome the levelling effects of gene flow and/or sweepstake recruitment (Hauser and Carvalho, 2008) for following cohorts in highly panmictic species like European plaice. However, if these species continue to be overharvested, small effective population size will also erode MHC variability thereby increasing the importance of balancing selection (Sommer, 2005; Radwan et al., 2010). Monitoring of genes under balancing selection in conjunction with selective agents like pathogenic bacteria might then become an important tool to measure potential selective responses of populations under threat of overfishing.

Acknowledgement

References

Andre, C., Larsson, L.C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G.R., Dahlgren, T.G., Hutchinson, W.F., Mariani, S., Mudde, K., Ruzzante, D.E., Ryman, N., 2010. Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106, 270—280.

Arkush, K.D., Giese, A.R., Mendonca, H.L., McBride, A.M., Marty, G.D., Hedrick, P.W., 2002. Resistance to three pathogens in the endangered winter-run chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histocompatibility complex genotypes. Canadian Journal of Fisheries and Aquatic Sciences 59, 966—975.

Austin, B., Austin, D.A., 2007. Bacterial Fish Pathogens — Disease of Farmed and Wild Fish, fourth ed. Springer, Berlin.

Babik, W., Taberlet, P., Ejsmond, M.J., Radwan, J., 2009. New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. Molecular Ecology Resources 9, 713—719.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57, 289—300.

Bernatchez, L., Landry, C., 2003. MHC studies in non-model vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology 16, 363—377.

Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R., Willerslev, E., 2007. The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. PLoS ONE 2, e197.

Birky, C.W., Fuerst, P., Maruyama, T., 1989. Organelle gene diversity under migration, mutation, and drift — equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121, 613—627.

Bolle, L.J., Dickey-Collas, M., van Beek, J.K.L., Erftemeijer, P.L.A., Witte, J.I.J., van der Veer, H.W., Rijnsdorp, A.D., 2009. Variability in transport of fish eggs and larvae. III. Effects of hydrodynamics and larval behaviour on recruitment in plaice. Marine Ecology Progress Series 390, 195—211.

Deter, J., Bryja, J., Chaval, Y., Galan, M., Henttonen, H., Laakkonen, J., Voutilainen, L., Vapalahti, O., Vaheri, A., Salvador, A.R., Morand, S., Cosson, J.F., Charbonnel, N., 2008. Association between the DQA MHC class II gene and Puumala virus infection in *Myodes glareolus*, the bank vole. Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 8, 450—458.

DeWoody, J.A., Avise, J.C., 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fish Biology 56, 461—473.

Dionne, M., Miller, K.M., Dodson, J.J., 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.

Doherty, P.C., Zinkernagel, R.M., 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.

Dybdahl, M.F., Lively, C.M., 1998. Host-parasite coevolution: evidence for rara advantage and time-lagged selection in a natural population. Evolution 52, 1057—1066.

Galan, M., Guivier, E., Caraux, G., Charbonnel, N., Cosson, J.F., 2010. A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. BMC Genomics 11.

Galand, P.E., Casamayor, E.O., Kirchman, D.L., Lovejoy, C., 2009. Ecology of the rare microbial biosphere of the Arctic Ocean. Proceedings of the National Academy of Sciences 106, 22427—22432.

Garrigan, D., Hedrick, P.W., 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution 57, 1707—1722.

Gonzalez, J.M., Moran, M.A., 1997. Numerical dominance of a group of marine bacteria in the alpha-subclass of the class Proteobacteria in coastal seawater. Applied and Environmental Microbiology 63, 4237—4242.

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebo, S., Stet, R.J.M., 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.

Haag-Liautard, C., Coffey, N., Houle, D., Lynch, M., Charlesworth, B., Keightley, P.D., 2008. Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. PLoS Biology 6, e204.

Hauser, L., Carvalho, G.R., 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries 9, 333—362.

Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Palsson, J., Van der Veer, H.W., Rijnsdorp, A.D., Stam, W.T., Olsen, J.L., 2005. Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa* L.). Proceedings of the Royal Society, Series B: Biological Sciences 272, 497—503.

Hoarau, G., Piquet, A., Van der Veer, H., Rijnsdorp, A., Stam, W., Olsen, J., 2004. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: a comparison of resolving power between microsatellites and mitochondrial DNA data. Journal of Sea Research 51, 183—190.

Hoarau, G., Rijnsdorp, A.D., Van der Veer, H.W., Stam, W.T., Olsen, J.L., 2002. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. Molecular Ecology 11, 1165—1176.

Hovenkamp, F., 1991. Immigration of larval plaice (*Pleuronectes platessa* L.) into the western Wadden sea: a question of timing. Netherlands Journal of Sea Research 27, 287—296.

Huber, J., Morrison, H., Huse, S., Neal, P., Sogin, M., Mark Welch, D., 2009. Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. Environmental Microbiology 11, 1292—1302.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Research 17, 377—386.

Kemp, P.F., Aller, J.Y., 2004. Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. FEMS Microbiology Ecology 47, 161—177.

Kloch, A., Babik, W., Bajer, A., Sinski, E., Radwan, J., 2010. Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole *Myodes glareolus*. Molecular Ecology 19, 255—265.

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.

Lenz, T.L., Becker, S., 2008. Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci — Implications for evolutionary analysis. Gene 427, 117–123.

Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., Gordon, J.I., 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nature Reviews Microbiology 6, 776–788.

Liston, J., 1957. The occurrence and distribution of bacterial types on flatfish. Journal of General Microbiology 16, 205–216.

Liu, Z., Desantis, T., Andersen, G., Knight, R., 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucleic Acids Research 36, e120.

Lively, C.M., Dybdahl, M.F., 2000. Parasite adaptation to locally common host genotypes. Nature 405, 679–681.

Lohm, J., Grahn, M., Langefors, A., Andersen, O., Storset, A., von Schantz, T., 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.

Luttikhuizen, P.C., Drent, J., van Delden, W., Piersma, T., 2003. Spatially structured genetic variation in a broadcast spawning bivalve: quantitative vs. molecular traits. Journal of Evolutionary Biology 16, 260–272.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.

McCusker, M., Bentzen, P., 2010. Positive relationships between genetic diversity and abundance in fishes. Molecular Ecology 19, 4852–4862.

McKay, J., Latta, R., 2002. Adaptive population divergence: markers, QTL and traits. Trends in Ecology & Evolution 17, 285–291.

Norqvist, A., Hagstrom, A., Wolf-Watz, H., 1989. Protection of rainbow trout against vibriosis and furunculosis by the use of attenuated strains of *Vibrio anguillarum*. Applied and Environmental Microbiology 55, 1400–1405.

Pedros-Alio, C., 2006. Marine microbial diversity: can it be determined? Trends in Microbiology 14, 257–263.

Pujalte, M.J., Sitjà-Bobadilla, A., Alvarez-Pellitero, P., Garay, E., 2003. Carriage of potentially fish-pathogenic bacteria in *Sparus aurata* cultured in Mediterranean fish farms. Diseases of Aquatic Organisms 54, 119–126.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Radwan, J., Biedrzycka, A., Babik, W., 2010. Does reduced MHC diversity decrease viability of vertebrate populations? Biological Conservation 143.

Reusch, T.B.H., Langefors, Å, 2005. Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined Stickleback *Gasterosteus aculeatus*. Journal of Molecular Evolution 61, 531–541.

Reynisson, E., Lauzon, H.L., Magnosson, H., Jonsdottir, R., Olafsdottir, G., Marteinsson, V., Hreggvidsson, G.O., 2009. Bacterial composition and succession during storage of North-Atlantic cod (*Gadus morhua*) at superchilled temperatures. BMC Microbiology 9, 250.

Romanenko, L.A., Uchino, M., Kalinovskaya, N.I., Mikhailov, V.V., 2008. Isolation, phylogenetic analysis and screening of marine mollusc-associated bacteria for antimicrobial, hemolytic and surface activities. Microbiological Research 163, 633–644.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75, 7537–7541.

Sommer, S., 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology 2, 16.

Suzuki, M.T., Giovannoni, S.J., 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Applied and Environmental Microbiology 62, 625–630.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Tollenaere, C., Bryja, J., Galan, M., Cadet, P., Deter, J., Chaval, Y., Berthier, K., Ribas Salvador, A., Voutilainen, L., Laakkonen, J., Henttonen, H., Cosson, J.F., Charbonnel, N., 2008. Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. Journal of Evolutionary Biology 21, 1307–1320.

Turner, S.M., Faisal, M., DeWoody, J.A., 2007. Zygosity at the major histocompatibility class IIB locus predicts susceptibility to *Renibacterium salmoninarum* in Atlantic salmon (*Salmo salar* L.). Animal Genetics 38, 517–519.

van Beek, F., Rijnsdorp, A., de Clerck, R., 1989. Monitoring juvenile stocks of flatfish in the Wadden Sea and the coastal areas of the southeastern North Sea. Helgoland Marine Research 43, 461–477.

van der Veer, H., Bergman, M., 1987. Predation by crustaceans on a newly settled 0-group plaice *Pleuronectes platessa* population in the western Wadden Sea. Marine Ecology Progress Series 35, 203–215.

van der Veer, H.W., Bolle, L.J., Geffen, A.J., Witte, J.I.J., 2009. Variability in transport of fish eggs and larvae. IV. Interannual variability in larval stage duration of immigrating plaice in the Dutch Wadden Sea. Marine Ecology Progress Series 390, 213–223.

van Oosterhout, C., Joyce, D.A., Cummings, S.M., Blais, J., Barson, N.J., Ramnarine, I.W., Mohammed, R.S., Persad, N., Cable, J., 2006. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). Evolution 60, 2562–2574.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73, 5261–5267.

Wegner, K.M., 2008. Historical and contemporary selection of teleost MHC genes: did we leave the past behind? Journal of Fish Biology 73, 2110–2132.

Wegner, K.M., 2009. Perspective: massive parallel MHC genotyping: titanium that shines. Molecular Ecology 18, 1818–1820.

Wegner, K.M., Kalbe, M., Milinski, M., Reusch, T.B.H., 2008. Mortality selection during the 2003 European heat wave in three-spined sticklebacks: effects of parasites and MHC genotype. BMC Evolutionary Biology 8, 124.

Wegner, K.M., Kalbe, M., Rauch, G., Kurtz, J., Schaschl, H., Reusch, T.B.H., 2006. Genetic variation in MHC class II expression and interactions with MHC sequence polymorphism in three-spined sticklebacks. Molecular Ecology 15, 1153–1164.

Wilson, B., Danilowicz, B.S., Meijer, W.G., 2008. The diversity of bacterial communities associated with Atlantic cod *Gadus morhua*. Microbial Ecology 55, 425–434.

Wilson, D.J., McVean, G., 2006. Estimating diversifying selection and functional constraint in the presence of recombination. Genetics 172, 1411–1425.

Zangenberg, G., Huang, M.M., Arnheim, N., Erlich, H., 1995. New HLA-DPB1 alleles generated by interallelic gene conversion detected by analysis of sperm. Nature Genetics 10, 407–414.

Zhang, Y.X., Chen, S.L., Liu, Y.G., Sha, Z.X., Liu, Z.J., 2006. Major histocompatibility complex class IIB allele polymorphism and its association with resistance/susceptibility to *Vibrio anguillarum* in Japanese flounder (*Paralichthys olivaceus*). Marine Biotechnology 8, 600–610.

Zheng, D.H., Mai, K.S., Liu, S.Q., Cao, L.M., Liufu, Z.G., Xu, W., Tan, B.P., Zhang, W.B., 2004. Effect of temperature and salinity on virulence of *Edwardsiella tarda* to Japanese flounder, *Paralichthys olivaceus* (Temminck et Schlegel). Aquaculture Research 35, 494–500.