MHC polymorphism in the three-spined stickleback and its role in host-parasite co-evolution

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität

zu Kiel

vorgelegt von

Tobias L. Lenz

Kiel

2008

Referent:	Prof. Dr. Manfred Milinski
Korreferent:	Prof. Dr. Günther Hartl
Tag der mündlichen Prüfung:	02.02.2009
Zum Druck genehmigt:	Kiel,

Der Dekan

“Nothing in biology makes sense except in the light of evolution.”

Theodosius Dobzhansky (1900 – 1975)

Table of content

Summary

Natural genetic variation is the raw material for evolution and in vertebrates the most polymorphic part of the genome is represented by the major histocompatibility complex. This complex is part of the adaptive immune system and plays a key role in the recognition of invading pathogens. The exceptional polymorphism of these genes is potentially the result of an ongoing arms race between vertebrate hosts and their parasites, described by the Red Queen Hypothesis, which is facilitated by genetic recombination and therefore maintains sexual reproduction. Genes in this region are known to influence parasite resistance and mating decisions in several species, including the three-spined stickleback *Gasterosteus aculeatus*. This fish has long been a model species for behavioural studies and is currently also becoming a model system in evolutionary ecology and genetics. With the available knowledge about its behavioural ecology, parasite community and first insights into the genetics of the stickleback MHC, it provides a perfect tool to investigate and disentangle the mechanisms that drive host-parasite co-evolution and potentially the maintenance of sexual reproduction.

Previous work on the MHC in the three-spined stickleback has mainly focussed on overall genetic diversity, but already found first hints for allele-specific host-pathogen interactions. To follow up on these results and further unravel the mechanisms of host-parasite co-evolution, I started my PhD with the development of a new MHC typing protocol, which is optimised for highly polymorphic loci and based on reference strand-mediated conformation analysis. The new protocol includes new primers that reliably amplify all known MHC class IIB alleles and was thoroughly verified by cloning and sequencing. It enables reliable detection and recognition of individual MHC alleles to the sequence level and therefore provides an important new tool for evolutionary studies in the stickleback. The first pilot screens gave already interesting results in terms of chromosomal organisation of the MHC with tight linkage between amplified MHC

class IIB loci and potential polymorphism in the number of loci between haplotypes.

The allele specificity also allowed a new experimental level, which was applied in an experiment with semi-natural enclosures. We selected wild-caught individual sticklebacks with specific MHC IIB genotypes and let them reproduce under semi-natural conditions where they encountered their natural parasites, but were protected from predators. During the full reproductive period, we collected eggs and determined parenthood from microsatellite typing. Parental assignment in combination with MHC genotyping revealed mate choice for intermediate genetic distance between mates, but not for specific haplotypes. Further analysis also showed effects of MHC IIB diversity and a haplotype on individual conditional and immunological parameters as well as parasite resistance.

Sticklebacks are known for their potential to adapt to new environments and show high diversification in many different habitats across their distributional range. This diversification potentially also includes adaptation to new parasite communities. In another experiment, we therefore tested two local stickleback populations for levels of local immunogenetic adaptation. These populations had been shown in previous studies to be differentiated and reproductively isolated and it has been suggested that this is based on locally specialised cycles of host-parasite co-evolution. We reciprocally exposed lab-bred hybrid and pure line F_2 fish to their native and to the foreign habitat. We found MHC-linked as well as genome-wide local adaptation. However the interactions turned out to be more complex than expected and were dependent on the life-history strategy of the parasites. These results show that more focussed experiments are needed to understand the interaction of host and parasite genotypes.

In the last and still ongoing project, we are increasing the focus and test experimentally, whether frequency-dependent selection between a host and a single parasite can occur. Although this is a widely hypothesised mechanism for the maintenance of MHC polymorphism, it could not be verified in vertebrates yet. Here we infected two groups of lab bred fish each with one of two sympatric nematode parasites, let them reproduce in a common garden set up and re-

infected their offspring. We hypothesize that under natural selection and amplified by sexual selection, the more resistant genotypes will show increased reproductive success and that the resistance is transmitted to the offspring. This will lead to an increase of the resistant host genotype and a decrease of the parasite in the offspring generation and by that fulfil the definition of negative frequency-dependent selection. The data collection has been finished, but the analysis is still ongoing. We will also include host MHC genotypes in the analysis to understand the extent of host allele-parasite associations, which could ultimately explain the exceptional polymorphism in the MHC.

The here presented studies provide a more detailed view on the extensive polymorphism of the MHC in the three-spined stickleback and give further insights into the key role that this gene complex has in the arms race between hosts and parasites, mediated through its pleiotropic role in mate choice and parasite resistance.

Zusammenfassung

Natürliche genetische Variabilität ist die Basis für Evolution und bei Wirbeltieren finden sich die variabelsten Gene im Haupthistokompatibilitätskomplex (Major Histocompatibility Complex, MHC). Dieser Genkomplex ist Teil des erworbenen Immunsystems und spielt eine wichtige Rolle bei der Erkennung von körperfremden Erregern. Die außergewöhnliche Variabilität dieser Gene ist vermutlich das Ergebnis eines ständigen Wettrüstens zwischen Wirten und ihren Parasiten, auch als Red Queen Hypothese bekannt, das durch genetische Rekombination erleichtert wird und dadurch auch sexuellen Fortpflanzung bevorzugt. Die Gene in dieser Region sind dafür bekannt in mehreren Arten Parasitenresistenz und Partnerwahl zu beeinflussen, so auch beim dreistachligen Stichling (*Gasterosteus aculeatus*). Dieser Fisch ist seit langem ein Modellorganismus für Verhaltensstudien und entwickelt sich derzeit auch zu einem Modellsystem für evolutionäre Ökologie und Genetik. Durch das verfügbare Wissen über ihre Verhaltens-Ökologie, ihre Parasiten und erste Einblicke in die Genetik des Stichlings-MHC bietet diese Art ein perfektes Werkzeug zur Untersuchung und Entschlüsselung der Mechanismen, die Wirt-Parasit-Koevolution und potenziell auch die Erhaltung der sexuellen Fortpflanzung steuern.

Frühere Arbeiten über den MHC im dreistachligen Stichling haben sich größtenteils mit der genetischen Diversität befasst, aber bereits erste Hinweise für allel-spezifische Wirts-Pathogen-Interaktionen gefunden. An diese Erkenntnisse anschließend und mit dem Ziel, die Mechanismen der Wirt-Parasit Koevolution besser zu verstehen, begann ich meine Promotion mit der Entwicklung eines neuen MHC Typisierungsprotokolls. Diese neue Methode wurde optimiert für hoch variable Loci und basiert auf der Referenzstrang-vermittelten Konformationsanalyse (RSCA). Das neue Protokoll basiert zusätzlich auf neuen Primern, die alle bisher bekannten MHC Klasse IIB Allele des Stichlings amplifizieren und wurde durch Klonierung und Sequenzierung verifiziert. Es ermöglicht die zuverlässige Auftrennung und Erkennung individueller MHC Allele bis auf Sequenzebene und bietet dadurch ein wichtiges neues Werkzeug für evolutionäre Studien mit dem dreistachligen Stichling. Erste Untersuchungen

ergaben bereits interessante Ergebnisse im Bezug auf die chromosomale Organisation des MHC, mit enger Kopplung zwischen einzelnen MHC Klasse IIB Loci und potenzieller Variabilität in der Anzahl der Loci zwischen Haplotypen.

Die Allelspezifität der neuen Methode erlaubt auch eine neue experimentelle Ebene, die in einem Experiment unter naturnahen Bedingungen angewandt wurde. Wir selektierten wild-gefangene Stichlinge mit spezifischen MHC IIB Genotypen und ließen sie unter semi-natürlichen Bedingungen reproduzieren, wo sie auf ihre natürlichen Parasiten trafen, aber vor Räubern geschützt waren. Während des gesamten Fortpflanzungszeitraums sammelten wir die befruchteten Eier und bestimmten anschließend die Elternschaft anhand von Mikrosatelliten. Die elterliche Zuordnung in Kombination mit der MHC-Genotypisierung ergab für die Partnerwahl eine Bevorzugung für Paare mit mittlerer MHC-genetischer Distanz zwischen den Partnern, aber keine Bevorzugung für bestimmte Haplotypen. Weitere Analysen zeigten auch Auswirkungen der MHC IIB Diversität und eines Haplotyps auf einzelne Konditions- und immunologische Parameter sowie auf Parasitenresistenz.

Stichlinge sind bekannt für ihr Potential zur Anpassung an neue Umweltbedingungen und weisen eine hohe Diversifizierung in viele verschiedene Lebensräume innerhalb ihres Verbreitungsgebietes auf. Diese Diversifizierung beinhaltet vermutlich auch Anpassung an neue Parasitengemeinschaften. In einem weiteren Experiment haben wir daher zwei Stichlingspopulationen auf ihre lokale immungenetische Anpassung hin untersucht. Diese Populationen hatten sich schon in früheren Studien als genetisch unterschiedlich und reproduktiv isoliert herausgestellt, und eine Aufspaltung aufgrund von lokal spezialisierten Zyklen der Wirts-Parasit Koevolution war hypothetisiert worden. Wir setzten Hybride und reine Linien einer laborgezüchteten F2 Generation ihren heimischen und reziprok den jeweils fremden Lebensräumen aus. Wir fanden sowohl MHC-basierte als auch genomweite lokale Anpassung an die jeweilige Parasitenfauna. Die Interaktionen erwiesen sich als komplizierter denn erwartet, und waren abhängig von der Infektionsstrategie der jeweiligen Parasiten. Diese Ergebnisse zeigen, dass weitere Experimente notwendig sind, um die Interaktion von Wirts- und Parasitengenotypen zu verstehen.

Im letzten, noch andauernden Projekt Erhöhen wir die Fokussierung und testen experimentell, ob die negativ-frequenzabhängige Selektion zwischen einem

Wirt und einem einzigen Parasiten auftreten kann. Obwohl es sich hierbei um einen weithin hypothetisierten Mechanismus für die Aufrechterhaltung der MHC-Variabilität handelt, konnte er bisher bei Wirbeltieren noch nicht verifiziert werden. Wir infizierten hierfür zwei Gruppen von laborgezüchteten Stichlingen mit jeweils einer von zwei parasitischen Nematodenarten, ließen sie unter 'common garden' Bedingungen reproduzieren und infizierten ihre Nachkommen mit den gleichen Parasiten. Unsere Hypothese war, dass unter natürlicher Selektion und verstärkt durch sexuelle Selektion, die parasitenresistenteren Genotypen einen erhöhten reproduktiven Erfolg aufweisen würden und dass diese Resistenz an die Nachkommen vererbt würde. Dies würde zu einem Anstieg des resistenten Wirtsgenotyps und einem Rückgang des Parasiten in der Generation der Nachkommen führen und damit die Definition der negativen frequenzabhängigen Selektion erfüllen. Die Datenerhebung für dieses Experiment ist abgeschlossen, aber die Auswertung ist noch im Gange. Zusätzlich werden wir auch MHC-Genotypen in die Analyse einbringen, um das Ausmaß der Wirtsallel zu Parasit Assoziation zu verstehen, welche letztlich den außergewöhnlichen Polymorphismus im MHC erklären könnte.

Die hier vorgelegten Studien liefern eine detailliertere Sicht auf die außerordentliche Variabilität des MHC im dreistachligen Stichling und geben weitere Einblicke in die entscheidende Rolle, die dieser Gen-Komplex im Rüstungswettlauf zwischen Wirt und Parasiten, vermittelt durch seine pleiotrope Rolle in Partnerwahl und Parasitenabwehr, erfüllt.

Introduction

The objective of evolutionary biology is to investigate and understand the mechanisms and processes that have led to the extraordinary biological diversity in nature. However, while 149 years of successful research by uncountable scientists, since Charles Darwin published his theory of evolution, have brought a tremendous insight and understanding of evolutionary processes, there are still many unresolved questions. Two major concerns in evolutionary biology until today are the maintenance of genetic variation in populations and the widespread occurrence of sexual reproduction.

Natural genetic variation

One of the prerequisites for the course of evolution as defined by Charles Darwin in his 'On the origin of species' (1859) is the presence of heritable variation in a population. This includes variation in any form that affects the individual itself or its interaction with the environment. Darwin speculated that the slightest difference between individuals could ultimately translate into differential survival and reproduction, due to natural selection and competition between individuals in the ever ongoing 'struggle for life'. Therefore only a fraction of the individuals in a population contribute to the next generation (Darwin 1859).

There are different levels of variation in nature. The most obvious is species diversity, which can (still) be observed by anyone willing to take a closer look at his or her surroundings. Undoubtedly, this diversity is a result of the vast number of different habitats, which are in turn a result of the climatic and geological variation on our planet (Fig 1, Gaston 2000). Less obvious and sometimes invisible is genetic variation within species (Ridley 1996; Oleksiak *et al.* 2002; Gibson & Dworkin 2004; Barrett & Schluter 2008). It is however the material for adaptation and therefore enables occupation of new niches as well as response to changing conditions in established habitats.

Intuitively, knowing that nearly all living beings use the same genetic code and species – at least within their kingdom – share many of their basic cellular

molecules (Barton *et al.* 2007; Alberts *et al.* 2008), one would expect the fixation of the most successful type of each molecule at least on the species level. In many instances this is indeed the case and an efficient cellular machinery has evolved to keep mutations at a minimum and even repair damages in the genetic code, which is equally ubiquitous as the genetic code itself (Alberts *et al.* 2008). This is reflected in the genetic identity between for example humans and chimpanzees, who share 99 % of their coding DNA (Wildman *et al.* 2003). Nevertheless, as outlined above, genetic variation on the species and population level persists and evolutionary mechanisms must exist that counteract the effects of directional selection and random genetic drift (Hedrick & Thomson 1983; Hedrick 2006; Barrett & Schluter 2008).

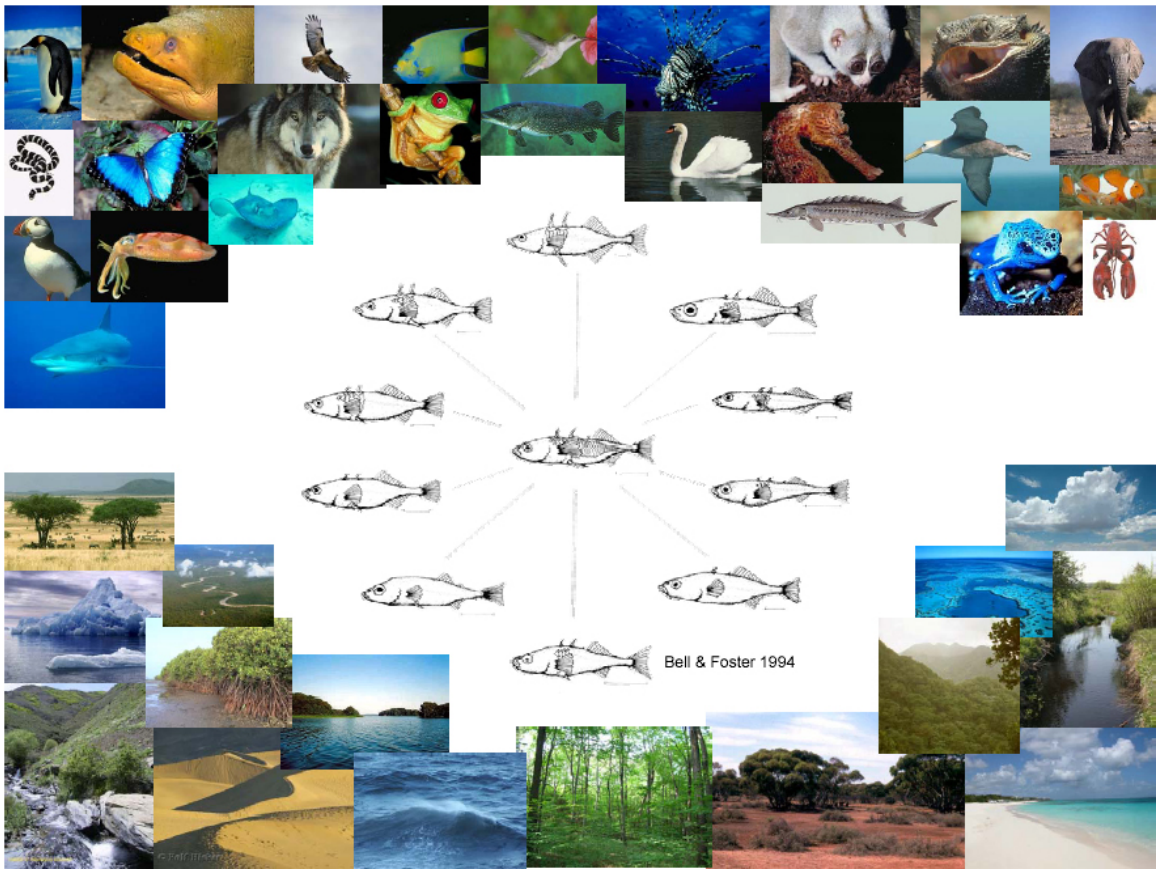


Figure 1: Examples for the vast natural diversity of organisms and habitats on earth in general and of the Three-spined stickleback in particular (pictures are courtesy of various anonymous artists).

Sexual reproduction

One mechanism that maintains genetic variation in a population is sexual reproduction, where two chromosomal sets get reshuffled to produce new offspring genotypes. However, the wide-spread occurrence of sexual reproduction in nature provides a paradox to evolutionary biologists (Maynard Smith 1978; Hamilton *et al.* 1990; Barton & Charlesworth 1998): Asexual organisms can produce twice the number of offspring of sexual ones, because they do not afford one sex that is only needed to fertilise the eggs of the other sex. Due to this higher reproductive output, an asexually reproducing organism would always outcompete its sexual counterpart, except if there were an advantage for the sexual organism. This at least two-fold advantage, however, which is necessary to counteract the cost of 'affording' two sexes, has not yet been satisfactorily resolved (Ridley 1996; Barton & Charlesworth 1998; Ryan 2004; Milinski 2006). The most commonly brought hypotheses, which address this paradox, propose that the clonally replicating organisms are suffering either from an increased accumulation of deleterious mutations without the possibility to purge them through recombination (Muller 1964) or from a limited overall genetic variability, which potentially decreases the chance to adapt to changing environments. These mechanisms might also work nonexclusively (Kondrashov 1993; West *et al.* 1999), however, neither the accumulation of mutations (e.g. Keightley & Eyre-Walker 2000) nor ordinary abiotic environmental changes are likely to provide such a detrimental effect that sexual reproduction provides a more than two-fold advantage to a sexual organism over its clonal counterpart (Maynard Smith 1978; Ridley 1996; Milinski 2006).

Host – parasite co-evolution

The origin for most genetic variation are mutations in the broadest sense, i.e. changes in the genetic code (Alberts *et al.* 2008). As these mutations usually occur randomly, the vast majority of them has either no or deleterious effects for the individual and gets purged again from the gene pool by means of natural selection (Frankham *et al.* 2002). However, there is a mechanism that conserves

variation through balancing selection, which is the process of frequency dependent co-evolution between individuals and species. The best example for such a process is the co-evolution of hosts and their parasites (Sorci *et al.* 1997). From the host point of view, parasites are part of the environment, but in contrast to abiotic factors, they are not fixed and therefore provide a 'moving target' for host adaptation processes. In addition, they do not vary randomly, for instance in the frequency of their genotypes, but they respond themselves adaptively to resistance mechanisms in the host. Parasites constantly drive their hosts to evolve new or recycle old variants of their defence mechanisms and vice versa, host resistance constantly urges the parasites to evolve new or recycle old variants of their exploitation mechanisms (Fig 2). This is often depicted as an 'arms race' and termed the '*Red Queen Hypothesis*' (Van Valen 1973; Ebert & Hamilton 1996):

"... it takes all the running you can do, to keep in the same place."
– Red Queen in L. Carroll's *Through the Looking-Glass* 1871

This arms race between hosts and parasites requires rapid response on both sides. Mutations would be one way to 'invent' new defence variants, but they occur at a too low rate to provide substantial innovations. New combinations of existing variants, however, would be comparably effective and organisms that can reshuffle their alleles between generations thus possess a significant advantage. Taken together, the need to counteract co-evolving parasites might explain the paradox of sexual reproduction, which newly recombines existing variation and therefore provides an advantage over

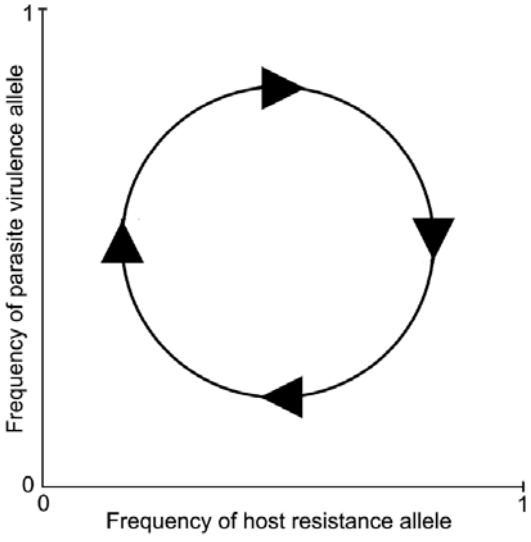


Figure 2: Co-evolutionary cycle of host resistance and parasite virulence.

clonally reproducing organisms in this arms race (Hamilton *et al.* 1990; Lively 1996).

This effect can be amplified by the process of mate choice during sexual selection. This, at first sight costly and time consuming behaviour, provides an elegant mechanism to select mates, which are most adapted to the environment (including parasites) and therefore promise a higher fitness potential than a random partner (Maynard Smith 1991; Ryan 2004). By exerting mate choice for 'honest signals', which are again costly and can only be presented by high quality mates (Zahavi 1975; Hamilton & Zuk 1982), the choosing sex increases the condition of its offspring and therefore its own fitness. In fact it has been shown that male sexual secondary signals, which affect female mate choice, can reflect the male's parasitisation (Hamilton & Zuk 1982; Read 1987; Milinski & Bakker 1990; Møller 1990a; Zuk *et al.* 1990).

Parasites have thus been proposed to be one of the major factors driving sexual selection and ultimately also explaining the paradox of sexual reproduction (Hamilton 1980; Lively 1987; Read 1988; Hamilton *et al.* 1990; Møller 1990b; Lively 1996; Milinski 2006). However, while numerous theoretical models describe the possible co-evolutionary scenarios of host-parasite interactions and several studies found support for necessary prerequisites, direct experimental evidence is still scarce and additional studies are needed to disentangle and understand the different mechanisms which maintain sex and genetic variation (Sorci *et al.* 1997; Salathé *et al.* 2008).

In addition to the fluctuation of advantageous and deleterious genotypes over time, there is also a geographical component to host-parasite adaptation. In structured metapopulations with reduced gene flow between subpopulations, host-parasite co-evolution can lead to local adaptation (Kaltz & Shykoff 1998; Lively 1999). Such ongoing, but geographically separated cycles of co-evolution are thought to be a driving force for ecological speciation of both parasites and hosts (Summers *et al.* 2003). This is of particular interest, since comparison of different populations can reveal divergent adaptive processes and allow for

evolutionary conclusions, which are otherwise only possible through long-term studies.

Adaptive immunity in vertebrates – The MHC

On the host side, parasite driven selection acts on genes that are involved in the immune response against these co-evolving pathogens. Jawed vertebrates possess a two-edged immune system, with two major subdivisions. The first line of defence is the innate immune system, which is unspecific and provides an immediate protection against a broad range of invading pathogens (Janeway *et al.* 2005). It does, however, not provide long-term immunity. The second line of defence is represented by the acquired immune system, which is antigen-specific and therefore provides a more sophisticated mechanism against pathogen invasion. It also involves the memorization of pathogen-specific antigens through memory cells and enables subsequently a faster response upon second infection (Janeway *et al.* 2005). A key component of the acquired immune system is the major histocompatibility complex (MHC, Klein 1986). The MHC represents one of the most polymorphic regions in the vertebrate genome (Klein 1979; Horton *et al.* 1998; Robinson *et al.* 2003) and comprises a large set of genes, many of which are involved in the initiation and maintenance of the adaptive immune response (Klein 1986; Trowsdale 1995; Beck & Trowsdale 2000). This complex is known to show recombination hot-spots as well as conserved regions in strong linkage disequilibrium (de Bakker *et al.* 2006), both potentially the result of strong selective forces. The high polymorphism in this system is composed of a high total number of alleles at the population level as well as high allelic divergence on the sequence level (Klein 1979; Robinson *et al.* 2003). It is thought to provide resistance against the vast repertoire of naturally occurring parasites and by that facilitates populations to adapt to and resist the various pathogen communities in natural habitats (Apanius *et al.* 1997; Bernatchez & Landry 2003; Piertney & Oliver 2005; Sommer 2005a; Milinski 2006). The most outstanding in terms of polymorphism are the classical MHC class I and class II genes, which are responsible for the presentation of intra- and extracellularly derived antigens

respectively. The class I genes are expressed in basically all nucleated cells and their products present antigens, derived from the cytosol, to the extracellular space. Here they are bound by T-cells, which recognise non-self antigens and can trigger a specific immune response (Trowsdale 1993; Janeway *et al.* 2005). This mechanism allows recognition and elimination of cells that have been invaded by intracellular pathogens, like for instance HIV (Moore *et al.* 2002), the Epstein Barr Virus (de Campos-Lima *et al.* 1993) or plasmodium, i.e. malaria (Hill *et al.* 1991). Analogously, the class II genes are expressed in so-called antigen presenting cells, i.e. macrophages, dendritic cells and B-lymphocytes, which take up and present antigens from the extracellular space (Trowsdale 1993; Janeway *et al.* 2005). These specialised cells are therefore involved in the detection and eradication of extracellular pathogens, like for instance helminth macroparasites (Froeschke & Sommer 2005; Harf & Sommer 2005; Schwensow *et al.* 2007; Tollenaere *et al.* 2008), but also microparasites (Thursz *et al.* 1995; Langefors *et al.* 2001).

Each respective MHC molecule binds only a certain range of antigens, a specificity that is mostly determined by special residues in the amino acid sequence of the binding grove, the antigen binding sites (ABS, Fig 3, Brown *et al.* 1993). Antigens in this context are short peptides of 8-10 amino acids (for class I) or longer (for class II) and are derived from self and non-self proteins via specific intracellular pathways (Janeway *et al.* 2005). Several bioinformatic prediction tools have been developed to estimate the binding properties of a given allele and tested on

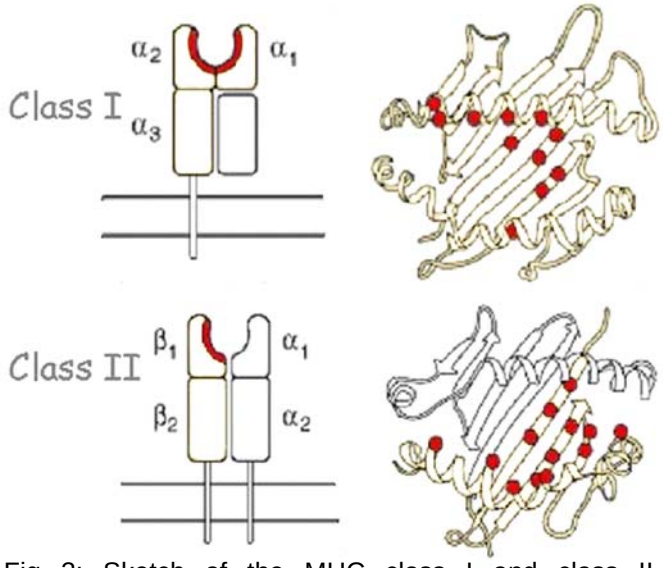


Fig 3: Sketch of the MHC class I and class II molecules. Shown are the different extracellular domains (Left: from the side, Right: antigen binding grove from above). Marked in red are the amino acid residues responsible for antigen binding.

experimentally acquired data (for example Altuvia & Margalit 2004; Bui *et al.* 2006; Peters *et al.* 2006; Karpenko *et al.* 2008). However, these are only available for human and mouse MHC and the extent to which certain alleles overlap in the range of peptides they bind is still an open question.

The MHC in the wild

The major histocompatibility complex has been studied extensively in model species under laboratory conditions (Apanius *et al.* 1997), but also became the focus of an increasing number of studies on natural populations because of its functional importance for the immune system and mate choice.

Its role in mate choice has first been investigated in mice (Yamazaki *et al.* 1976; Potts *et al.* 1991), but was subsequently supported in many other species (reviewed in Milinski 2006). It is based on olfactory cues that are linked to MHC peptide ligands (Milinski *et al.* 2005; Spehr *et al.* 2006). The explanations for MHC-dependent mate choice are based on its unique polymorphism and the fact that nearly all individuals differ in their MHC genotype to certain extents, which makes these genes a perfect candidate for gene based mate choice mechanisms. The proposed mechanisms range from inbreeding avoidance to optimizing the immunogenetic repertoire of the next generation (reviewed in Potts & Wakeland 1990; Edwards & Hedrick 1998; Penn & Potts 1999; Bernatchez & Landry 2003; Milinski 2006). Several studies found that females choose mates according to their own MHC repertoire, either selecting for optimal diversity (e.g. Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Bonneaud *et al.* 2006a; Forsberg *et al.* 2007) or maximum diversity (e.g. Landry *et al.* 2001; Neff *et al.* 2008; Schwensow *et al.* 2008b) in the progeny. In both cases, females were looking for '*genetic compatibility*' in their partners (Mays & Hill 2004). Other studies found no evidence for mate choice for compatible MHC genes, but instead for '*good genes*', which increase the fitness of their bearer, irrespective of the choosing sex' constitution (Ekblom *et al.* 2004; Richardson *et al.* 2005). However, many studies on MHC-dependent mate choice focussed on diversity in terms of number of alleles or heterozygosity and only rarely single alleles or haplotypes

have been addressed (e.g. von Schantz *et al.* 1996; Ekblom *et al.* 2004). As each MHC allele can only present a limited range of antigens, there are differences between distinct alleles in resistance against a given pathogen (e.g. Hill *et al.* 1991; Thursz *et al.* 1995; Langefors *et al.* 2001; Harf & Sommer 2005; Bonneaud *et al.* 2006b). So far it is still an open question to which extent '*good genes*' and '*compatible genes*' contribute to the evolution of MHC-based mate choice (Mays & Hill 2004; Neff & Pitcher 2005).

The function of the MHC in pathogen resistance is more obvious, because of its crucial role in the adaptive immune system. A number of hypotheses have been put forward to explain the high polymorphism in the MHC based on parasite-driven selection that counteracts the effects of random genetic drift and fixation (reviewed in Potts & Wakeland 1990; Apanius *et al.* 1997; Edwards & Hedrick 1998; Summers *et al.* 2003; Piertney & Oliver 2005; Sommer 2005a). The '*negative frequency-dependent selection*' hypothesis (Clarke & Kirby 1966; Bodmer 1972; Takahata & Nei 1990) is based on the cyclic patterns of host parasite co-evolution, stating that frequencies of host alleles constantly change with the frequency of adapted and non-adapted parasite genotypes. New or rare alleles are being maintained and promoted within a population through frequency fluctuations of pathogens. The prerequisite for this co-evolution in terms of reciprocal adaptation is the direct interaction of individual alleles and distinct pathogenic morphotypes. Evidence from a variety of taxa for such an interaction exists (e.g. Hill *et al.* 1991; Thursz *et al.* 1995; Langefors *et al.* 2001; Harf & Sommer 2005b; Bonneaud *et al.* 2006b). Another hypothesis, the *'heterozygosity advantage*' has been proposed as a general mechanism to maintain high allele numbers in populations by an advantage for individuals carrying more than one allele and by that enabling them to present a broader range of pathogens (Doherty & Zinkernagel 1975b; Klein & Figueroa 1986). Some experiments support this hypothesis when animals have been exposed to multiple pathogens (Penn *et al.* 2002; up to an optimum in Wegner *et al.* 2003a) but not if exposed to single pathogens (Langefors *et al.* 2001; Wedekind *et al.* 2005; Wedekind *et al.* 2006), which seems plausible, as the advantage would most likely result from a

combined resistance of individual alleles each presenting a certain range of parasites (‘overdominance’, Hughes & Nei 1988, 1989). Nevertheless the exact mechanisms are still a matter of debate (Lipsitch *et al.* 2003; Woelfing *et al.* 2009). A recent study showed for instance a negative effect of extreme heterozygosity at the MHC (Ilmonen *et al.* 2007) and De Boer *et al.* (2004) studied the degree of MHC polymorphism arising when ‘*heterozygote advantage*‘ is the only selection pressure by using mathematical models. The simulations revealed that the advantage for an individual that simply carries two alleles instead of one is not sufficient to explain the high population diversity of the MHC. Wakeland *et al.* (1990) proposed a special case of heterozygosity advantage, the '*divergent allele advantage*', which refers to the genetic distance between alleles and expects an advantage for more distant allele combinations, enabling the respective carrier to present more different antigens to its adaptive immune response. This mechanism could potentially explain selection for both, a large number of alleles at the population level and a high divergence of alleles at the individual level. A number of recent studies have found support for this hypothesis in mate choice (Landry *et al.* 2001; Consuegra & Garcia de Leaniz 2008; Neff *et al.* 2008; Schwensow *et al.* 2008a), but the expected direct effect in higher resistance against multiple infections has not yet been shown.

While maximum pathogen recognition selects for as many different MHC molecules as possible, another selective force drives allelic diversity on the individual level in the opposite direction. This is the negative T-cell selection during thymic development. The highly diverse T-cell receptor repertoire, which is produced by somatic recombination in a similar fashion as antibodies, plays the complementary part in the detection of non-self antigens besides the MHC molecules (Janeway *et al.* 2005). Its diversity is of key importance for the detection of non-self antigens as each T-cell clone carries a receptor for the detection of a unique MHC-bound antigen. T-cells develop in the thymus and undergo an MHC-dependent selective process in which self-reactive T-cells get eliminated (Boehm 2006). Theory predicts that an overall higher number of different MHC molecules will eliminate a larger range of T-cell clones, eventually

resulting in a depleted T-cell repertoire (Nowak *et al.* 1992). The number of alleles, under which this depletion becomes significant and disadvantageous, is however under debate (for review see Woelfing *et al.* 2009). Some field studies, which analysed MHC diversity over several loci, have found signs for an optimal diversity (Wegner *et al.* 2003b; Buchholz *et al.* 2004; Madsen & Ujvari 2006) and experimental approaches have supported this observation (Wegner *et al.* 2003a; Bonneaud *et al.* 2004a; Kurtz *et al.* 2004; Kalbe *et al.* In press). Understanding this effect could eventually contribute to the understanding of outbreeding depression, i.e. the phenomenon of inferior hybrids between two locally adapted populations, and ecological speciation (Eizaguirre *et al.* 2009).

The limited number of MHC loci, which is potentially the result of this trade of between maximising antigen presentation and minimising T-cell elimination, therefore determines the bottleneck for adaptation against parasite virulence. Each individual can express only a limited number of the alleles that are available in the population and consequently can only resist a limited range of pathogens. This bottleneck keeps the 'Red Queen' of host-parasite co-evolution running as it leaves space for parasites to evade MHC presentation. It might also drive host specialisation to specific parasite communities, i.e. local adaptation, as this potentially limits the number of different pathogens that the host has to fight (Eizaguirre *et al.* 2009).

The three-spined stickleback

Three-spined sticklebacks occur throughout the Northern hemisphere in a diverse array of habitats, ranging from saltwater to freshwater and have already been called "the cichlids of the North" (Bjarni Jónsson, 2006 Stickleback Conference). Their intriguing behavioural repertoire and extent of natural variation within and between populations has triggered an enormous number of studies, which contributed significantly to our knowledge about various evolutionary processes. It has provided revealing insights in the mechanisms that drive life-history trade offs (e.g. Milinski & Heller 1978), sexual selection (e.g. Milinski & Bakker 1990; Bakker *et al.* 1999; Reusch *et al.* 2001a; Milinski *et al.*

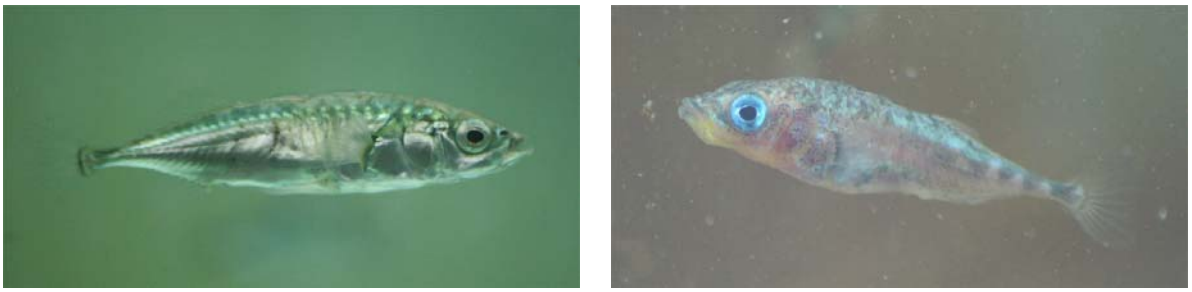


Figure 4: Female (left) and male (right) three-spined stickleback.

2005), natural selection (e.g. Wegner *et al.* 2003a; Barrett *et al.* 2008), speciation (e.g. Schluter 1994; Rundle *et al.* 2000; Boughman 2001; McKinnon *et al.* 2004) and development (e.g. Shapiro *et al.* 2004; Colosimo *et al.* 2005). It is therefore fair to say that it has become a well known and accepted model system (McKinnon & Rundle 2002; Gibson 2005; Cresko *et al.* 2007; Beaumont 2008; Cresko 2008). Since the stickleback genome has been sequenced, a large genetic resource is also available and new genetic tools such as a tight linkage map or expression arrays are being developed in parallel to take advantage of this fascinating and potent species (Peichel *et al.* 2001; Hubbard *et al.* 2007; Miller *et al.* 2007b). Three-spined sticklebacks are relatively easy to maintain and breed in the lab. A valuable resource for the understanding of host-parasite co-evolution is the well described makroparasite community of the stickleback (Wedekind 1997; Kalbe *et al.* 2002; Wegner *et al.* 2003b). Several of the better known parasite species, such as the digenean trematode Diplostomum pseudospathacaeum, the cestode Schistocephalus solidus or the nematodes Camallanus lacustris and Anguillicoloides crassus can be used for controlled infections and enable experimental tests of immunocompetence (Wegner *et al.* 2003a; Kurtz *et al.* 2004; Kurtz *et al.* 2007) or local immunological adaptation (Kalbe & Kurtz 2006; Scharsack *et al.* 2007). Several important life-history traits, like parasite resistance (Wegner *et al.* 2003a), survival (Wegner *et al.* 2008) and reproductive success (Kalbe *et al.* In press), have been shown to be influenced by optimal MHC genetics.

The three-spined stickleback's intriguing mating system includes on the male side territoriality, nest building, courtship behaviour and maintenance of the

eggs and on the female side extensive mate choice based on olfactory and visual cues. The red breeding coloration of the males for instance has been shown to be an honest signal, reflecting parasite resistance (Milinski & Bakker 1990). The breeding behaviour can be studied under controlled conditions in the lab and it has been shown that visual (e.g. nest quality or red coloration, Jäger *et al.* 2007) as well as olfactory cues are influenced by MHC genetics (Reusch *et al.* 2001a; Aeschlimann *et al.* 2003; Milinski *et al.* 2005). This species therefore provides a highly suitable model to study the evolutionary mechanisms that lead to sexual reproduction and local adaptation as well as maintain the high polymorphism in the vertebrate's MHC.

Outline

During my PhD I focussed on the genes of the evolutionary important major histocompatibility complex in the three-spined stickleback *Gasterosteus aculeatus*. I studied their exceptional polymorphism and the evolutionary mechanisms that maintain this variability in the context of host-parasite co-evolution. My thesis consists of five main projects, which are outlined below. All of these projects have been conducted in cooperation with different colleagues. Table O.1 provides a detailed overview of each author's contribution.

Chapter I

Simple approach to reduce PCR artefact formation leads to reliable genotyping of *MHC* and other highly polymorphic loci – implications for evolutionary analysis

To understand the mechanisms that lead to the genetic variability in the MHC, it is of utmost importance to use a reliable genotyping technique. Previous work has dealt mainly with the overall diversity at all MHC loci. For a more detailed view and understanding of the evolutionary mechanisms, an allele specific analysis is inevitable. The characterisation of multi-gene markers, such as the MHC with its duplicated loci, poses an increased risk of producing PCR and cloning artefacts. As a first step, we therefore developed a protocol for the amplification of multi-allele templates, which minimizes the risk of artefact production. In this chapter we also analysed the effects that such artefacts can have on the evolutionary interpretation of sequence data.

Chapter II

RSCA genotyping of MHC for high-throughput evolutionary studies in the model organism three-spined stickleback *Gasterosteus aculeatus*

Cloning and sequencing is commonly used for the characterisation of individual genotypes. This is a time-consuming procedure, which becomes even more complex when a higher number of alleles is to be expected in the template. To investigate evolutionary phenomena with sometimes small effect

sizes, it is however necessary to achieve large experimental sample sizes. We therefore developed and optimised a protocol for high-throughput genotyping of the MHC class II loci of the stickleback. This protocol is based on Reference Strand-mediated Conformation Analysis (RSCA) and its validity was confirmed by the most accepted and established method, i.e. cloning and sequencing. First results on the MHC constitution from the new method are also presented.

Chapter III

Effects of MHC diversity and specific haplotypes on immunocompetence, reproductive success and mate choice decisions of Three-spined sticklebacks under semi-natural conditions

In previous studies on three-spined sticklebacks, both, mate choice and parasite resistance have been shown to be influenced by MHC genetics. The focus, however, has so far mostly been on MHC diversity in terms of number of different alleles. In this project, we aimed at investigating the role of specific MHC haplotypes and diversity on the sequence level in mate choice under semi-natural conditions. Using the newly developed genotyping protocol, we selected wild-caught sticklebacks that carried specific MHC IIB genotypes. We then let them reproduce in an outdoor enclosure system, where they were protected from predators but still faced their natural parasites. The experimental design allows us to differentiate between '*good gene*' and '*compatible gene*' effects.

Chapter IV

MHC local adaptation: a field experiment

The polymorphic genes of the MHC have been proposed to enable individuals to adapt and therefore survive and reproduce in their local pathogenic environment. Previous studies have in deed found differences in the MHC diversity between habitats that also differ in parasite communities. However, such potential local MHC adaptation has so far not been experimentally tested. We therefore crossed three-spined sticklebacks from a river and a lake habitat, whose populations carry different parasites and posses almost totally distinct MHC allele pools. By intercrossing F1 hybrids, we obtained a second

fish generation, which segregated into pure river, pure lake and two hybrid MHC genotypes, while at the same time maintained a common and randomized genetic background. We used a double common garden experiment where fish were exposed simultaneously to natural lake and river conditions for ten consecutive months. The design aimed at studying the extent of local adaptation on the MHC level as well as on the whole genome level.

Chapter V

Parasites maintaining MHC polymorphism: an experimental test for negative frequency-dependent selection

One of the suggested mechanisms that maintain the exceptional polymorphism in the MHC is negative frequency-dependent selection by parasites. Host individuals with specific MHC alleles that confer resistance against currently predominating parasites should outperform others that do not posses these alleles. Consequently these individuals should also be favoured during mate choice. We designed an experimental test for negative frequency-dependent selection including parasite pressure and sexual selection. Six replicate populations of lab-bred three-spined sticklebacks were exposed to one of two nematode parasites. After natural reproduction in controlled outside tanks, we confronted the offspring with both parasites again. We hypothesise that fish from parents that faced one parasite will be less heavily infected by this parasite than those coming from parents that encountered the other parasite and vice versa. This project was not completed at the stage of thesis submission. Therefore only preliminary results are presented.

Table O.1: Individual contribution to the different chapters.

Chapters	**I**	**II**	**III**	**IV**	**V**
Idea & experimental design	SB, **TLL**	CE, SB, TBHR, **TLL**	CE, MK, MM, **TLL**	CE, MM, **TLL**	CE, MM, **TLL**
Research performance	**TLL**	**TLL**	CE, JS, MK, **TLL**	CE, CH, **TLL**, MK	CE, MK, JS, **TLL**
Data analysis	**TLL**	**TLL**	CE, **TLL**	CE	CE
Manuscript writing	SB, **TLL**	**TLL**	**TLL**	CE, **TLL**	CE, **TLL**

Authors are given in alphabetical order. CE: Christophe Eizaguirre, CH: Chris Harrod, JS: Jörn Scharsack, MK: Martin Kalbe, MM: Manfred Milinski, SB: Sven Becker, TBHR: Thorsten B. H. Reusch, TLL: Tobias L. Lenz.

Chapter I

Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci – implications for evolutionary analysis

Tobias L. Lenz & Sven Becker

Abstract

Genetic variation in coding regions is of strong interest for biologists as it represents an important factor that drives evolution. To analyse polymorphic loci, researchers usually rely on commonly used typing techniques such as cloning, SSCP, DGGE or RSCA. However, there are potential pitfalls in screening multi-allelic templates, which are mainly the formation of sequence chimeras during PCR amplification, and mosaic sequences during cloning. One of the most challenging genomic regions to explore is the Major Histocompatibility Complex (MHC), which codes for peptide-binding proteins of the vertebrate's adaptive immune system and is well known for its exceptional polymorphism. We compared the effect of two different PCR amplification approaches in a study of the MHC class IIB genes of the three-spined stickleback (*Gasterosteus aculeatus*). One approach used standard PCR conditions and the other a combination of several measures to eliminate PCR artefacts. In both approaches, the amplicons obtained were cloned and sequenced. In the first, established approach, 24% of the clones represented artefacts, while in the second approach the number of artefacts were reduced ten-fold. Furthermore, it enabled easy differentiation between real alleles and artificial sequences. We also analysed the potential effects of such artefacts in genetic analysis and evolutionary interpretation, and found a slight reduction in the signature of positive selection and an increase in recombination events. Consequently, we strongly recommend to apply the new PCR approach described in this study when genotyping MHC or other polymorphic genes.

This chapter has been published in *Gene* (2008) 427: 117-123.

Introduction

Genetic variation is an important factor in the course of evolution and therefore provides the basis of most research in evolutionary biology and similar fields. One of the outstanding examples for high genetic polymorphism is the Major Histocompatibility Complex (MHC, Klein *et al.* 1993b; Beck & Trowsdale 2000). Its genes code for cell surface receptor molecules that present pathogen-derived antigens to T-cells, and hence trigger the response of the adaptive immune system in vertebrates (Klein 1986). The variability in the MHC is represented at the population level in the number of alleles as well as in sequence divergence between alleles, and at the individual level in the number of locus duplications and excess of heterozygosity (Klein 1986). This polymorphism has been observed in most species and populations studied so far and is maintained by balancing selection, potentially mediated by the diversity and fluctuation of pathogens in the natural environment. The exact reasons for the origin and persistence of this polymorphism are still under discussion (reviewed in e.g. Apanius *et al.* 1997; Edwards & Hedrick 1998; Hedrick 1999; Bernatchez & Landry 2003; Piertney & Oliver 2005; Sommer 2005a; Milinski 2006). Many studies on wild animals as well as a number of experimental studies tried to understand these mechanisms, but the results are still inconsistent (Paterson *et al.* 1998; Penn *et al.* 2002; Lipsitch *et al.* 2003; Wegner *et al.* 2003a; Westerdahl *et al.* 2005; Axtner & Sommer 2007; Ilmonen *et al.* 2007; Schwensow *et al.* 2007).

A first step to test any of the proposed hypotheses is the characterization of individual MHC genotypes. Several techniques are currently in use, e.g. single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE, both discussed in Lessa & Applebaum 1993), reference strand-mediated conformation analysis (RSCA, Argüello *et al.* 1997), and cloning. Each of these relies on the amplification of the respective target genes by polymerase chain reaction (PCR, Mullis *et al.* 1986). In some mammalian species it is possible to amplify individual MHC genes separately while this is hardly possible in most fish, bird and reptile species, because of high sequence similarity between loci or because of recent locus duplication. Due to the high level of MHC heterozygosity there is usually more than one allele in the PCR template. This holds true especially

for species where loci cannot be targeted individually, hence there are usually several different alleles in the template that will be amplified in parallel.

There are potential pitfalls in screening multiallelic templates during PCR amplification, independent of the identity of the screened locus. These are mainly the formation of chimeras (amplicons that contain sequence motifs from two different alleles), and the formation of heteroduplexes, which become mosaic sequences through the *E. coli* mismatch repair system during cloning (Kanagawa 2003). The formation of chimeras is thought to mainly occur during the later cycles of a PCR thermal programme, when dNTP and primer concentrations have been reduced, while the number of complete, but also incompletely elongated, amplicons increased. Under such conditions, the incomplete amplicons function as "heteroduplex primers" by hybridising to the wrong templates due to partial homology and higher melting temperature than the original short primer oligonucleotides (Saiki *et al.* 1988; Meyerhans *et al.* 1990; Kanagawa 2003). Another possibility for such an artefact to occur is the methyl-directed mismatch repair system mutHLS in *E. coli*, which usually corrects DNA polymerase errors in newly replicated DNA strands that still lack methylation (Modrich 1987). This happens during cloning, when heteroduplexes of two non-complementary strands have been formed in the last PCR cycle and are ligated and transformed into *E. coli*. The mutHLS system then repairs arbitrarily both strands and hence produces allele 'mosaics', because in PCR amplicons both strands are unmethylated (Longeri *et al.* 2002; Thompson *et al.* 2002; Kanagawa 2003).

The possibility of such artificial amplicon (or allele) formation during PCR amplification has already been mentioned in the early history of PCR (Saiki *et al.* 1988; Nagamine *et al.* 1989), and since then a number of approaches have been reported to eliminate artefact formation during and after PCR (Table I.1). However, their application is still rare, although artefact formation has been highlighted in several important research fields other than genetics (reviewed in Kanagawa 2003). Most of the studies on such artefacts deal with microbial ecology, where microbial communities are characterized by amplifying the same gene from a pool of bacteria (Judo *et al.* 1998; Becker *et al.* 2000; Thompson *et al.* 2002; Acinas *et al.* 2005). In such cases unnoticed artefacts may lead to an overestimation of diversity (v.

Wintzingerode *et al.* 1997). PCR artefacts have also been identified in MHC research (L'Abbé *et al.* 1992; Longeri *et al.* 2002), but here artificial alleles are hard to detect because they usually resemble original alleles that were established through recombination of alleles (which is exactly what is happening during PCR artefact formation *in vitro*).

Unnoticed artificial alleles can lead to overestimation of allele numbers and population diversity. As these are crucial parameters for understanding the mechanisms of MHC evolution, we pursued a thorough investigation of the frequencies of such PCR artefacts and how they can be minimized. We tested two different amplification approaches for the highly polymorphic MHC class IIB genes of the three-spined stickleback *Gasterosteus aculeatus* (2-4 loci, Reusch & Langefors 2005): one with standard PCR conditions, and one approach with a combination of several conditions to reduce PCR artefact formation. These conditions are the use of independent amplification reactions, a reduction of PCR cycle number, increased elongation time and a reconditioning step. Additionally, we discuss the probability of artefact formation in different typing techniques and sum up with a recommendation for future studies.

Table I.1: Studies on PCR artefact reduction in the literature.

Template	Template diversity	Proposed PCR conditions	Reference
HIV genes	2 strains	- Increased elongation time	(Meyerhans *et al.* 1990)
Murine Protease inhibitor (α_1-PI)	Multigene family	- Third primer in additional PCR (PCR+1)	(Borriello & Krauter 1990)
HLA DRB	Multigene family	- Third primer in additional PCR (PCR+1)	(L'Abbé *et al.* 1992)
Alcohol dehydrogenase (Adh-1)	Heterozygous single locus	- Use proof-reading polym. (Vent)	(Bradley & Hillis 1997)
pUC-based plasmid sequence	2 artificial sequence variants	- Reduced number of PCR cycles - Increased elongation time - Avoid proof-reading polym. (Vent)	(Judo *et al.* 1998)
Murine immunoglobulin germline genes V	Allele mixes (4-11 alleles)	- Reduced number of PCR cycles - Avoid proof-reading polym. (Pfu)	(Zylstra *et al.* 1998)
16s rDNA	3 bacteria species	- Reconditioning PCR	(Thompson *et al.* 2002)
16s rRNA	Bacterioplancton sample (> 500 ribotypes)	- Reduced number of PCR cycles - Reconditioning PCR	(Acinas *et al.* 2005)

Materials and Methods

Genetic material

For this study we used the genomic DNA of ten three-spined sticklebacks (*Gasterosteus aculeatus*) from Northern Germany. The fish were caught in Lake Großer Plöner See in January 2004 and preserved in 70% ethanol. The DNA for this study was freshly extracted from muscle tissue, using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to the protocol, and eluted in buffer (included in the kit). The DNA was normalized to a concentration of 30 ng/µl for subsequent PCR amplifications.

Amplification

For both approaches evaluated in this study we used a hot-start DNA polymerase (AmpliTaq Gold®; Applied Biosystems, Darmstadt, Germany) that has no proof-reading capacity, because two studies have shown that proof-reading enzymes enhance PCR artefact formation (Judo *et al.* 1998; Zylstra *et al.* 1998). The oligonucleotide PCR primers (Metabion, Martinsried, Germany) for both approaches were GA11F (Sato *et al.* 1998) for forward, and GAIIExon2R_std or GAIIExon2R_mod (Reusch *et al.* 2001a) for reverse amplification. The amplicons were a 198 bp fragment of the highly polymorphic exon 2 of the MHC class IIB genes, including 14 of the 24 putative antigen binding sites (Brown *et al.* 1993). We used two different reverse primers in independent reactions, because they are supposed to amplify two subsets of alleles in the sampled population (Binz *et al.* 2001; Reusch *et al.* 2001a). This led to at least two independent amplification reactions per individual. The two reaction volumes were pooled after completion of the thermal programme, and the amplicons with different reverse primer sequences were identified by sequencing in a later stage of our study.

The PCR reaction volume was 20 µl and consisted of 2 µl template DNA, 1x GeneAmp® PCR Buffer II, 5 mM $MgCl_2$, 50 µM of each dNTP, 0.5 µM of each primer and 1 U of AmpliTaq Gold®. The PCR was conducted on the thermal cycler PC-200 (Bio-Rad, Munich, Germany). For the first approach we employed no specific conditions to reduce artefacts and used the following protocol: 95°C for 10 min (activation of the hot-start polymerase), 33 cycles of

94°C for 30 s, 58°C for 30 s and 72°C for 30 s with a final extension step of 72°C for 3 min.

In the second amplification approach we combined several conditions from the literature to reduce the formation of PCR artefacts: eight independent amplification reactions per individual, a reduction of the PCR cycle number to 25 in the first round of PCR, a longer elongation time within each PCR cycle (60 instead of 30 s), and a reconditioning step in which diluted product from the first PCR is used as a template in a second PCR. Independent amplification reactions have been proposed to eliminate PCR bias and random artefacts, because they would only occur in one of the replicate reactions (Wagner *et al.* 1994; Kanagawa 2003). We used four replicates of the two reverse primer reactions which led to eight independent amplification reactions. The four replicates were cloned and sequenced independently. The reduction of PCR cycles has been proposed by several studies (Judo *et al.* 1998; Zylstra *et al.* 1998; Acinas *et al.* 2005), because conditions like low primer to amplicon ratio and increased occurrence of incompletely elongated fragments at the plateau phase of the PCR are avoided. We tested the minimum number of cycles needed in our study by agarose gel electrophoresis and ethidium bromide staining of fragments. Agarose gels (1.5%) were loaded with 4 µl of the PCR volume and run for 3.5 hours at 80 V. Gels were then stained for 40 min in an ethidium bromide solution, rinsed in HPLC water and documented with a Gel Imager system (Intas, Göttingen, Germany). After 25 thermal cycles product bands were usually visible in the gel (data not shown). Additionally, we increased the elongation time within each cycle to 60 s to avoid the formation of incompletely elongated amplicons (Meyerhans *et al.* 1990; Judo *et al.* 1998). Finally, we employed a reconditioning PCR step of 6 cycles after the initial PCR as recommended by Thompson et al. (2002). We diluted the product of the first 25 cycle-PCR 5-fold in HPLC water and used 2 µl of it as template in the second PCR, together with a normal reaction mix (see above) and the same thermal programme as before, but only 6 cycles. This led to a readjustment of the primer to amplicon ratio and decreased the chance of over-amplifying rare chimeras.

Cloning and sequencing

For both PCR approaches, the volumes of the two reverse primer reactions per sample or replicate were pooled and 30 µl were loaded on an agarose gel (1.5%) and separated overnight at 45 V. The gel was stained with SYBR Gold (Invitrogen, Karlsruhe, Germany) for 35 min and bands were excised and eluted in 30 µl HPLC water (NucleoSpin Extract II kit by Macherey Nagel, Düren, Germany). Fragments were cloned with the TA cloning kit (Invitrogen, Karlsruhe, Germany), according to the manufacturer's protocol. Plasmids were extracted with the Invisorb Plasmid HTS 96 kit (Invitek, Berlin, Germany) on the Freedom EVO® working station (Tecan, Crailsheim, Germany). Plasmid DNA was sequenced on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany) using the universal plasmid primer M13 and the BigDye 3.1 terminator kit (Applied Biosystems, Darmstadt, Germany).

Genetic analysis

Sequences were aligned manually with the programme BioEdit 7 (Hall 1999). We tested the effect of the sequence artefacts on commonly used genetic analysis, i.e. tests for selective pressure (CODEML, Z-test) and gene conversion events (GENECONV). CODEML is included in the PAML4 package (Yang 2007) and calculates the likelihood for models assuming different selection regimes by estimating the ratio of non-synonymous and synonymous nucleotide substitutions (ω) for different site classes. We tested the model M1a against the model M2a. M1a represents a nearly neutral selection scenario, by assuming one class of sites under purifying selection ($\omega < 1$) and another class of sites under neutral selection ($\omega = 1$). Model M2a assumes positive selection and adds a third class with $\omega > 1$ (Yang 2007). Using likelihood ratio statistics, the best explanatory hypothesis can be selected, and thus signs for positive selection can be detected (Yang *et al.* 2000). Another method to test for signatures of selection is the codon-based Z-test, which is implemented in the programme MEGA4 (Tamura *et al.* 2007). It also estimates the overall value for ω and tests whether it is significantly different from one. For the calculation we used the Nei-Gojobori method with Jukes-Cantor correction for multiple substitutions at the same site (Nei &

Gojobori 1986), and a bootstrap replication of 10,000. The software GENECONV (Sawyer 1999) screens a set of alleles for stretches of similar sequences and detects traces of former gene conversion events as well as gives information about recombination between alleles (Sawyer 1989). The reported p-values were calculated based on 10,000 permutations, and we used the number 123 as the starting seed as proposed by the author.

We also tested whether the formation of chimeras between two alleles is linked to the genetic distance between them. For this we calculated the nucleotide distance for allele pairs using the similarity matrix function in BioEdit. The number of chimeras between each two alleles was divided by the number of fish in which these alleles occurred and then correlated with the distance. Statistical tests on all results were done with the programme SPSS (SPSS GmbH, Munich, Germany).

Results

First approach with standard PCR protocol

From the first amplification approach, 439 clones were sequenced (between 38 and 47 per individual). By excluding irreproducible single nucleotide polymorphisms that are caused by polymerase errors, 82 distinct sequences were retrieved. All of them aligned very well without any indels, and produced only MHC class II hits in an NCBI BLAST search. However, a large number of these sequence variants was only represented by a small number of clones, and after within-individual sequence comparison were identified as artificial chimeras. This distinction could only be achieved due to the large

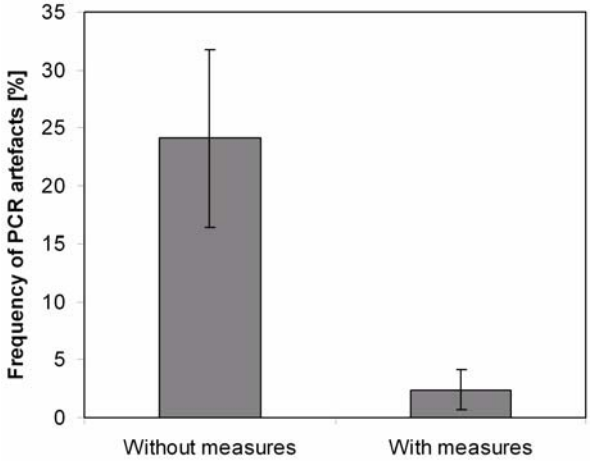


Figure I.1: Frequency of artefact sequences without and with PCR conditions to reduce artefact formation, mean and 95% CI. PCR amplification of MHC class IIB genes of 10 individuals of three-spined stickleback, *Gasterosteus aculeatus*.

number of clones picked. Finally, all original alleles were identified and chimeras were assigned to the allele pair from which they originated, because the original alleles were usually represented by a higher number of clones than the chimeras (on average 8.8 clones per original allele versus 1.6 clones per chimera; Mann-Whitney U test, $Z = 6.4$, $p<0.001$). This led to a reduced pool of 13 real alleles in the ten fish, with a range of 2 to 5 per individual. As a summary, in this first amplification approach with standard conditions, on average 24% of the sequenced clones represented artificial allele chimeras (Fig. I.1). These chimeras were hybrids of usually two (see Fig. I.2 for an example), in rare cases also three different alleles that occurred in the same individual. Of the 69 detected chimeras, 26% were found in at least two independent PCR reactions (18 sequence variants represented by on average 2.7 clones; Fig 3), and 50% of those even in different individuals which had the same allele pair. The occurrence in two independent reactions within an individual was implied from the fact that both reverse primer sequences were represented by at least one of the clones respectively, indicating that they had been formed in both independent reverse primer reactions. The number of clones for each of the repeatedly detected chimeras ranged from 2 to 7. The number of chimeras did not correlate significantly with the number of alleles within an individual (non-parametric correlation, Spearman's rho = 0.38, $p=0.27$).

12 of the 13 detected alleles could be found in GenBank (accession no: DQ016399, AY687842, AF395722, AY687843, DQ016400, AF395711, DQ016410, DQ016417, AY687846, AJ230191, DQ016404 and DQ016402). The new allele was submitted as well (GenBank: EU541449). None of the artefacts detected in our study could be found in GenBank.

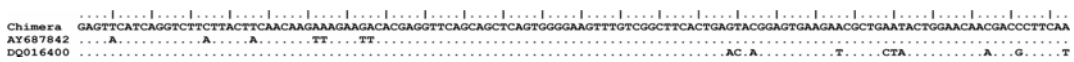


Figure I.2: Example of a chimera PCR product from 2 different alleles (accession numbers: AY687842 and DQ016400). Dots in the alignment represent consensus with the chimera sequence in the top line. Only a part of the amplified fragment is shown for better scaling.

Second PCR approach with conditions to reduce artefact formation

In the second amplification approach with a combination of several conditions to reduce artefact formation, 24 clones from each of the four independent amplification reactions per individual were picked and in total 891 clones were sequenced (between 84 and 94 clones per individual). We chose a higher number of clones in this approach to detect more rare sequence variants. In total 31 unique sequence variants were identified and all of them showed MHC class II identity in a NCBI BLAST search. After manual alignment, 13 distinct exon 2 alleles and 18 chimeras could be differentiated. Of those chimeras 15 were only represented once, and 3 others twice, all of them in independent PCR reactions. It was not possible to determine whether the chimeras had occurred during PCR or through mismatch repair during cloning. The exon 2 alleles were represented by on average 23.4 clones and therefore were differentiated unambiguously from the PCR artefacts. The proportion of clones that represented an artefact was only 2.3%, a ten-fold reduction of the proportion found in the first approach (Mann-Whitney U test, $p<0.001$; Fig. I.1). The 13 alleles identified and the individuals in which they were found were the same as detected in the first approach.

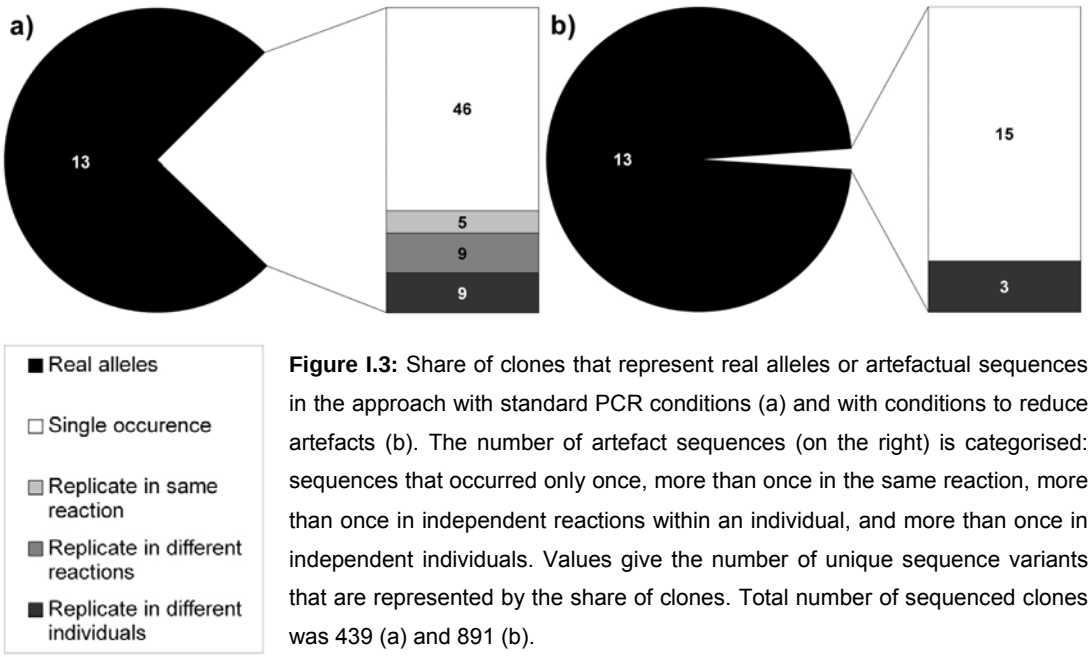


Figure I.3: Share of clones that represent real alleles or artefactual sequences in the approach with standard PCR conditions (a) and with conditions to reduce artefacts (b). The number of artefact sequences (on the right) is categorised: sequences that occurred only once, more than once in the same reaction, more than once in independent reactions within an individual, and more than once in independent individuals. Values give the number of unique sequence variants that are represented by the share of clones. Total number of sequenced clones was 439 (a) and 891 (b).

Effects of chimeras on genetic analysis

We analyzed the effect of the chimeras on the d_N/d_S ratio (ω) of the derived sequences. For this we compared the group of 13 real alleles against the group of sequences that occurred at least in two independent reactions (13 real alleles + 18 chimeras). First we used the programme CODEML, which is based on a maximum likelihood calculation (Yang 2007). For both groups the neutral model M1a was rejected in favour of the selection model M2a, which fitted the sequence data significantly better (alleles only: $2\Delta\ell = 40.66$, $p<0.001$; alleles + chimeras: $2\Delta\ell = 35.97$, $p<0.001$). A slight difference can be observed in the extent of the site class that shows ω>1 and in the ω of this site class, which is 38% of the sites with ω = 9.53 in the real alleles, and only 29% of the sites with ω = 7.28 when including the chimeras. This means that in our data set the inclusion of the artefacts reduces the signature of positive selection slightly. The same pattern was observed with MEGA using the codon-based Z-test. The group of real alleles showed a strong trend towards positive selection over all sites (Z-value = 1.973, p=0.051), while the inclusion of the chimeras reduced the signature of selection to a level at which it is not significant anymore (Z-value = 1.679, p=0.092).

We also applied the software GENECONV, which can detect sequence fragments that result from former recombination events. We again compared the group of real alleles against a combination of real and artificial alleles to see whether the inclusion of the artefacts would skew the result. Within the real alleles, 5 global inner fragments were detected that may result from intra- or even inter-locus recombination. When including the artefacts, the number of global inner fragments rose to 17 (all Bonferroni-corrected sim p-values < 0.05). The number of chimeras between any two alleles did not correlate significantly with the nucleotide distance of the two alleles (non-parametric correlation, Spearman's rho = 0.22, p=0.21).

Discussion

In this study, we developed a protocol to minimize PCR artefacts when amplifying the polymorphic genes of the MHC class IIB in the three-spined stickleback *Gasterosteus aculeatus*. Additionally, we tested the effect of

artefactual allele sequences with respect to evolutionary analyses and interpretations. We also did an extensive literature search to determine standard PCR conditions currently used for the amplification of MHC loci and the application frequency of conditions to reduce artefact formation. Overall, we found 19 studies in Web of Science, which were published in 2007 or 2008 and had analysed MHC diversity in an ecological context using PCR amplification and either cloning, SSCP or DGGE. In two of the studies, the exact amplification protocol was not described. The average number of PCR cycles in the remaining 17 studies was 33.5, with a range of 30 to 40 cycles. Although several studies acknowledged the possibility of PCR artefacts, the application of additional conditions to reduce them was scarce. Only one study used two independent amplification reactions for all samples and none used a reconditioning second PCR.

In our first amplification approach with average PCR conditions, we detected a high number of unique sequences that were all confirmed by BLAST search to originate from MHC class IIB loci. After alignment and sequence comparison, we distinguished 13 real alleles from 69 artefacts. Most of those artefacts were represented only once, and were easily identified due to the high number of clones we sequenced. However, several sequence variants were represented by two or more clones, and even in independent reactions. Representation of a sequence by more than one clone, preferably in independent PCR reactions, is used as a threshold to officially assign new alleles of the human MHC (Marsh *et al.* 2005). With these criteria, in our first amplification approach this would have been 26% of the chimeras, leading to 18 new alleles and hence more than double the initial pool of 13 real alleles.

By using a combination of non-average PCR conditions in the second amplification approach, we reduced the occurrence of artefacts ten-fold. We still observed a small number of suspicious alleles, but they were differentiated easily from the real alleles because of their low representation. Therefore the application of the non-average PCR conditions significantly improved the genotyping reliability for MHC class IIB genes in our study.

We also analyzed the potential effects such artefacts can have on genetic analysis and evolutionary interpretation, by comparing the real alleles alone against the pool of both real and artificial alleles. Using the programmes

CODEML and MEGA, we analyzed the ratio of non-synonymous to synonymous nucleotide substitutions in the amplified section of exon 2. With CODEML it is not possible to compare the fit of the models between the two groups, and both groups showed a significantly better fit with the selection model M2a. Using the Z-test for selection in MEGA, however, we showed that the signature for positive selection was slightly reduced when the artificial sequences were included in the analysis. This is probably due to the arbitrary nature of artefact formation, which counteracts the effect of positive selection. The repair system of the cloning host *E. coli* for instance arbitrarily 'repairs' mismatches of heteroduplexes and therefore may contribute to an increase in the number of synonymous substitutions.

When using the software GENECONV to detect traces of gene conversion, we found an increase in the number of fragments that appear to originate from recombination events. This is an expected result as the chimeras consist of parts of the real alleles and therefore share long sequence stretches with the original alleles. This result shows that undetected PCR artefacts can lead to an overestimation of recombination.

Another aspect applies to studies where the MHC is used as a population genetic marker. It seems possible that an artificial increase in local sequence variants, such as it was observed here, leads to significant differentiation between populations while there is none in reality, because chimeras will increase the representation of local, i.e. private sequence motifs. Unfortunately, we could not test this aspect because all our tested fish originated from the same local population.

We cannot differentiate whether the observed artefacts occurred during PCR amplification or through mismatch repair during cloning. This seems to be non-trivial as most studies do not differentiate between the two mechanisms. We are also unable to tell which of the new PCR conditions had the strongest effect in reduction of artefacts, but based on other reports (Judo *et al.* 1998; Zylstra *et al.* 1998) we assume that it is the reduced number of PCR cycles. Especially when using indirect typing methods like RSCA or SSCP, for which the PCR product is denatured before screening, the reconditioning PCR step might be omitted. Heteroduplex formation in the last PCR cycle is neutralized in these methods, and the risk of mosaic-like

sequence artefacts from the mismatch repair system of *E. coli* only occurs during cloning only (Thompson *et al.* 2002). However, if a high number of PCR cycles were needed to amplify a certain fragment, it would probably be advisable to use a reconditioning PCR with a small amount of diluted template instead of increasing the cycle number.

In typing techniques, which use the separation of alleles due to sequence-specific denaturation characteristics of double stranded PCR products, e.g. DGGE (Myers *et al.* 1987) and constant denaturant capillary electrophoresis (CDCE, Khrapko *et al.* 1994), the formation of heteroduplexes can increase the number of detected alleles, as was shown by Thompson et al. (2002). Even techniques that are not using target-specific primers, but are based on the linkage of restriction or other recognition sites, are also affected by heteroduplex formation. This has been shown for restriction fragment length polymorphism (RFLP, Jansen & Ledley 1990) and random amplification of polymorphic DNA (RAPD, Ayliffe *et al.* 1994). When using such techniques, the reconditioning PCR step is highly recommended to avoid overestimation of allelic diversity.

There is, however, a probability of missing rare alleles when applying the reconditioning step, because only an aliquot is taken from the PCR product, hence sequence variants with low concentration may be missed. Such an under-representation can occur when primers target a variable region and, amplify alleles with different efficiency. This is not uncommon in MHC research, because in some species there are hardly any conserved areas in this polymorphic region (e.g. in the three-spined stickleback, Binz *et al.* 2001). We tried to consider this by using a lower dilution of the template for the second PCR than initially recommended by Thompson et al. (2002). This made it possible to use a larger aliquot of the first PCR product. Another possibility for the under-representation of alleles can be an arbitrary PCR bias (Wagner *et al.* 1994). Such random effects can be avoided by pooling the product of several independent amplification reactions before analysis (Kanagawa 2003).

Even if it would be possible to amplify single loci only, this reduction of template diversity was not sufficient to avoid artefacts. We demonstrated this with the two individuals that had only two alleles but still showed a substantial

number of chimeras. In general, the number of chimeras did not correlate with the number of original alleles in the template. A study that employs the amplification of the DRB3 locus in the bovine MHC revealed artefacts when only two alleles were present (Longeri *et al.* 2002). Similar results were found by Judo et al. (1998) when testing two artificial sequence variants. An additional source for artefacts is low quality DNA, attained for instance through non-invasive sampling techniques, such as faecal sampling (Knapp 2005).

The studies shown in Table 1, which addressed the formation of artefacts and pointed out ways to avoid them, used a variety of templates and applications. This indicates that the issue is not a problem in MHC research only, but it rather is a general challenge that applies to multi-template amplifications. Most of the studies listed in Table 1 aimed at describing the diversity of the template analysed, with respect to the number of bacteria species in a microbial sample (Thompson *et al.* 2002; Acinas *et al.* 2005), number of HIV strains (Meyerhans *et al.* 1990), multigene families (Borriello & Krauter 1990; L'Abbé *et al.* 1992) and single genes in heterozygous individuals (Bradley & Hillis 1997). Two studies even used artificial templates to investigate the formation of artefacts (Judo *et al.* 1998; Zylstra *et al.* 1998). Amplification of the MHC, especially if several loci are addressed simultaneously, is technically the same as amplifying one gene from for instance a pool of bacterial strains. The PCR conditions that were tested by these studies and that we combined in our second amplification protocol in this study have therefore a general applicability to the MHC, but also to other highly polymorphic loci.

There are examples where PCR artefacts have been mistaken as polymorphism and were annotated in GenBank, e.g. MHC-linked olfactory receptor genes (discovered by Ziegler *et al.* 2000), and possibly also MHC DRB genes (discovered by Kennedy *et al.* 2002). The artefactual sequences that were detected in this study were not found in GenBank, although most of the detected original alleles are available in the database. Those alleles were uploaded from four different cloning projects, indicating that they probably had not been found together in the same individual before. This would explain why none of our artefacts had been detected previously.

To estimate how widespread the use of PCR conditions for the reduction of artefact formation is, we screened the most recent literature that deals with MHC loci and ecological questions. We found that the large majority of studies on MHC loci with respect to ecological, population genetic or evolutionary questions do not apply any such conditions. It is impossible to speculate how many previous results have been skewed by artificial alleles from PCR reactions, but the application of the conditions described in this study will help to avoid such biases in future studies. Furthermore, it is advisable to verify PCR-based genotyping of highly polymorphic loci with independent reaction assays and even typing techniques if possible. Based on the results in this study and earlier observations by other scientists, these are crucial points in evolutionary studies on multi-allelic templates like for example the MHC.

Acknowledgements

We would like to thank Manfred Milinski, Christophe Eizaguirre, Thorsten Reusch and Irene Samonte for fruitful discussion and helpful comments during the writing of this manuscript. We appreciate the comments of two anonymous reviewers to improve an earlier version of the manuscript. Thorsten Reusch and Martin Kalbe provided the fish for this experiment. We are very grateful to Sybille Liedtke for excellent help in the lab. This study was supported by Max Planck Society.

Chapter II

RSCA genotyping of MHC for high-throughput evolutionary studies in the model organism three-spined stickleback *Gasterosteus aculeatus*

Tobias L. Lenz, Christophe Eizaguirre, Sven Becker & Thorsten B. H. Reusch

Abstract

The highly polymorphic genes of the major histocompatibility complex (MHC) code for antigen presenting molecules of the adaptive immune system in all jawed vertebrates. The polymorphism is composed of a high total number of alleles as well as high nucleotide diversity at the sequence level between any two alleles. It provides means of resistance against the natural parasite fauna, adaptation to new pathogen threads, and co-evolution with sympatric parasites. To be able to analyse this diversity and understand the underlying mechanisms that maintain it, requires a reliable genotyping technique, optimised for such highly variable genes.

We present a genotyping protocol that uses Reference Strand-mediated Conformation Analysis (RSCA), optimised for the duplicated MHC class IIB genes of the three-spined stickleback, *Gasterosteus aculeatus*. In addition, we use a comprehensive plasmid library of MHC class IIB alleles to determine the nucleotide sequence RSCA allele peaks. Verification of the RSCA typing by cloning and sequencing shows high congruency between both methods and provides new insight into the polymorphism of the stickleback MHC. Analysis of the plasmid library additionally reveals the high resolution and reproducibility of this technique.

This new RSCA genotyping protocol offers a fast, sensitive and reliable way to determine the MHC allele repertoire of three-spined sticklebacks. It therefore provides a valuable tool to employ this highly polymorphic and adaptive marker in future high-throughput studies of host-parasite co-evolution and ecological speciation in this emerging model organism.

Manuscript in review in BMC Evolutionary Biology

Introduction

Natural genetic variation is the basic material for evolution, and to study its origin and persistence helps understand how organisms, populations and ecosystems function. The major histocompatibility complex (MHC) is one of the most polymorphic regions in the vertebrate genome (Bodmer 1972; Klein 1979). It consists of genes which code for cell surface molecules that present self and non-self antigens to T-cells and therefore play an essential role for the recognition of pathogens invading the body (Klein 1986). The high polymorphism of MHC class I and II genes is usually reflected in high total number of alleles as well as exceptional nucleotide diversity at the sequence level between any two alleles, It probably resulted from sexual selection through mate choice, and from natural selection due to co-evolving parasites, bacteria and viruses (Bernatchez & Landry 2003; Piertney & Oliver 2005; Sommer 2005a; Milinski 2006).

The population genetics of classical MHC genes has stimulated a growing body of research and is increasingly being studied in natural populations. Yet, its polymorphism also presents a serious challenge as reliable genotyping methods for such variable and often duplicated loci are difficult to apply and optimise. In many mammalian species it is nowadays possible to target single loci, because of relatively stable structural organisation and gene orthology within the mammalian MHC (Edwards *et al.* 1995b; Nei *et al.* 1997). In contrast, non-mammalian species show a substantially lower conservation in their MHC structure (Edwards *et al.* 1995b). This leaves a large black box concerning the organisation of the MHC and gene orthology in most non-model species and locus specific typing is hardly possible (Edwards *et al.* 1995a; Binz *et al.* 2001; Reusch *et al.* 2001; Bonneaud *et al.* 2004b; Westerdahl *et al.* 2004b; Huchard *et al.* 2006). Therefore it is of major concern for our understanding of the MHC and its role in evolution and ecology, to find a reliable typing method for MHC diversity that is not depending on detailed knowledge about the respective species' organisation of the MHC and is able to cope with larger numbers of alleles.

The design of MHC primers for new species can be achieved by starting from sequence information of closely related species, because of certain sequence conservation of parts of the MHC genes (Edwards *et al.*

1995a; Klein *et al.* 2007). This approach facilitates the amplification of MHC genes in new species without excessive genome screens, but potentially risks lower specificity of the primers. Additionally, more and more species under investigation show signs of very recent gene duplications in the MHC. Both, conserved sequence motifs as well as sequence similarity due to recent duplication often lead to co-amplification of several loci and a large number of alleles with undistinguishable locus origin that have to be separated simultaneously (e.g. Reusch *et al.* 2001; Bonneaud *et al.* 2004b; Westerdahl *et al.* 2004b; Bryja *et al.* 2006; Huchard *et al.* 2006; Meyer-Lucht *et al.* 2008).

The most reliable and trusted method for allele identification so far is sequencing with prior allele separation via cloning of PCR products. However, this is costly and requires substantial effort, which increases exponentially with the expected number of alleles. As sequenced clones represent only a small sub-sample of the amplified fragments, this method is more prone to PCR artefacts and overestimation of true allele number, especially if only a small number of clones of each individual is sequenced to save resources (Bower *et al.* 2005; Lenz & Becker 2008).

Faster and less expensive methods for genotyping of unknown alleles have been employed, the most common ones being denaturant gradient gel electrophoresis (DGGE, Myers *et al.* 1987) and single-strand conformation polymorphism analysis (SSCP, Orita *et al.* 1989). DGGE separates different double-stranded sequence variants according to their denaturation characteristics and under optimal conditions provides a single band per variant. SSCP analysis achieves separation of variants due to mobility differences of the two complementary single strands in a non-denaturing matrix and therefore provides two bands per allele (Both methods discussed in Lessa & Applebaum 1993). SSCP can also be run in a capillary system, using automated laser detection of labelled single strands and by that producing peaks instead of bands (CE-SSCP, Makino *et al.* 1992; Binz *et al.* 2001). Several factors can complicate allele identification with these methods, which produce only one and two characteristic values per allele respectively. These are for instance large numbers of distinct variants in the pool, which increase the likelihood of overlapping and therefore undistinguishable peaks,

or run to run variation in the gel matrix, which confounds the comparison of different cohorts of individuals that have been typed at different time points.

A disadvantage of indirect typing methods is also the lack of sequence information. The MHC genes comprise certain amino acid residues that are actively involved in antigen binding and presentation while most of the other residues have more conformational importance (Brown *et al.* 1993). The former residues usually face strong balancing selection, shown by a high rate of non-synonymous substitutions between alleles, while the latter are less variable. With increasing insight into the evolutionarily important mechanisms of antigen binding, it becomes a key aspect to resolve alleles at the sequence level, enabling sequence and even residue-specific analysis of adaptation (Blais *et al.* 2007; Forsberg *et al.* 2007; Neff *et al.* 2008; Schwensow *et al.* 2008a). While this lack of sequence information in SSCP and DGGE can be overcome by running the fragments on a physical gel and subsequently sequence the detected bands, this is not possible for high throughput typing protocols like CE-SSCP.

Reference Strand-mediated Conformation Analysis (RSCA, Argüello *et al.* 1998) is a more sophisticated PCR-based genotyping technique, which experiences increasing popularity. Here, all sequence variants are PCR-amplified simultaneously and hybridised to a given reference strand. Upon cooling, heteroduplexes are formed with distinct mismatches for each variant. The reference strand is a known sequence variant, produced by amplifying a single allele template (plasmid or homozygous individual). These heteroduplexes are then separated according to their specific mobility in a given non-denaturing environment, which depends on the distinct tertiary structure of each heteroduplex. This mechanism provides a significant advantage over other commonly used indirect typing techniques, because it produces as many mobility values (bands or peaks) per allele as desired by using several distinct reference strands and therefore creates a multi-dimensional coordinate for each allele. This enables differentiation of highly similar sequence variants (Corell *et al.* 1999).

Since its development by Argüello *et al.* (1998), RSCA has been used in a number of species to type MHC loci, namely in humans (e.g. Argüello & Madrigal 1999), other primates (e.g. Smith *et al.* 2005; Baquero *et al.* 2006;

Tanaka-Takahashi *et al.* 2007; Blasky *et al.* 2008) and in several non-primate mammalian species (e.g. Feichtlbauer-Huber *et al.* 2000; Kennedy *et al.* 2003; Brown *et al.* 2004; Drake *et al.* 2004; Kennedy *et al.* 2005). However, its application in non-mammalian species is still rare. To our knowledge the only studies so far have been performed in brown trout (Noakes *et al.* 2003) and red jungle fowl (Worley *et al.* 2008), where only one and two loci have been addressed at a time, respectively.

The three-spined stickleback *Gasterosteus aculeatus* is an emerging model species that is increasingly used to study evolutionary phenomena like, for instance, sexual reproduction (e.g. Reusch *et al.* 2001; Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Boughman 2007; Jäger *et al.* 2007), host-parasite co-evolution (e.g. Wegner *et al.* 2003a; Kalbe & Kurtz 2006), ecological speciation (e.g. Rundle *et al.* 2000; Bolnick 2004; Boughman *et al.* 2005; Vines & Schluter 2006; Blais *et al.* 2007; Gow *et al.* 2007; Svanback & Bolnick 2007) or evo-devo (e.g. Peichel *et al.* 2001; Shapiro *et al.* 2006; Cresko *et al.* 2007). In this study we focused on the class II genes of the stickleback MHC, as they were shown to have significant influence on parasite resistance (Wegner *et al.* 2003a; Kurtz *et al.* 2004), mate choice (Reusch *et al.* 2001; Aeschlimann *et al.* 2003; Milinski *et al.* 2005), survival (Wegner *et al.* 2008), lifetime reproductive success (Kalbe et al., submitted; Eizaguirre et al., submitted-b) and even speciation (Eizaguirre et al., submitted-a). Previously, MHC class IIB genes of the three-spined stickleback have been analysed using CE-SSCP (Binz *et al.* 2001; Reusch *et al.* 2001). A crucial point in genotyping is the number of different loci that have to be separated. In the past, the number of MHC class IIB loci in the stickleback had been reported to be as high as 6 (Sato *et al.* 1998), but this estimate was recently reduced to 2-4 (Reusch & Langefors 2005). Potentially due to recent origin of locus duplication or ongoing inter-locus recombination (Reusch *et al.* 2004), it is not possible to address these loci singly and therefore several alleles per individual have to be differentiated, independent of the genotyping technique used. This poses a strong challenge even to cloning, the gold standard genotyping method, due to the increased rate of sequence artefacts under certain conditions (Lenz & Becker 2008). The established CE-SSCP protocol for the stickleback has been known to provide limited resolution by not

separating all sequence variants (Reusch *et al.* 2001). It finally reached its limits when following a comprehensive breeding design with fish from two different natural populations, for which a new set of eight primer combinations had to be developed (Rauch *et al.* 2006). This increased the number of necessary PCR reactions and SSCP runs per fish substantially and inhibited high-throughput genotyping. We therefore developed a new genotyping protocol employing RSCA and constructed a comprehensive plasmid library of MHC class IIB alleles to determine the allele sequence represented by each RSCA peak. Results of the allele- and sequence-specific RSCA typing were verified using the gold standard technique of cloning and sequencing.

Methods

Primer design

It is a certain challenge to design primers for the MHC region. If primers within the most variable exon 2 are needed, they should on the one hand be as locus-specific as possible, and on the other hand they shall amplify all alleles that are present at this locus. In the tree-spined stickleback the constraint of locus specificity can be neglected, because the different MHC loci do not show locus-specific patterns due to concerted evolution (Reusch *et al.* 2004; Reusch & Langefors 2005). Previous studies have used different approaches to study the MHC class IIB genes of the stickleback and a number of different primer pairs were designed (Binz *et al.* 2001; Reusch *et al.* 2001; Wegner 2004); none of the primer pairs alone, however, amplifies all present sequence variants. Therefore we strove to develop a primer pair that amplifies all exon 2 alleles simultaneously and, together with the higher resolution of RSCA, would provide a fast and reliable genotyping technique for future research on MHC in sticklebacks.

Our criteria for the primer pair used for RSCA genotyping were, first, that they align in a conserved region to include all possible sequences in the detected allele pool, second, that they enclose most of the highly variable exon 2, and third, that both primers hybridise within the exon 2. The last criterion avoids any length variation of the amplicon because no intron sequences are included. The presence of amplicons of different length may

cause PCR conditions that outcompete longer amplicons (Becker *et al.* 2000), and in RSCA may lead to a hybridisation bias with the reference strand.

Due to the high sequence variation in exon 2 of the MHC class II loci (Ono *et al.* 1993; Sato *et al.* 1998; Reusch & Langefors 2005), the first criterion is a certain challenge with respect to the primer design. However, recently published sequence information (Reusch *et al.* 2001; Reusch *et al.* 2004; Reusch & Langefors 2005) and the stickleback genome (http://mar2008.archive.ensembl.org/Gasterosteus_aculeatus) provided good sources for the design of new primers. Additionally we designed a primer for the conserved exon 1 of the MHC IIB genes (GAIIEx1F: 5'- CAG CGT CTC CCT CCT CTT CAT – 3') and cloned the exon 1 to intron 2 sequence of a number of fish to obtain sequence information about the so far rarely addressed beginning of the exon 2. Based on this combined new sequence information, the new forward primer: GAIIEx2startF (5'- GTC TTT AAC TCC ACG GAG CTG AAG G -3') was then set in a fully conserved region at the beginning of exon 2 (Fig II.1) and is therefore optimised for the three-spined stickleback. The new reverse primer GAIIExon2R_RSCA (5'- ACT CAC CGG ACT TAG TCA G –3') spans the exon 2 – intron 2 boundary in a conserved region as well (Fig II.1). The partial connection to the exon 2 avoids any length polymorphisms, which occur frequently in the rest of the intron 2. Although the new reverse primer spans the exon-intron boundary, it can nevertheless be used for expression studies, because more than two thirds of it align within the exon 2 and the remaining one third is complementary to both the beginnings of intron 2 and exon 3, which are highly similar, leaving only 2 conserved mismatches at the 5'-end when used with reverse-transcribed cDNA. This new primer combination produces a fragment of 247 base pairs (203 bp without primers) and spans 88 % of the entire exon 2. These primers amplify all currently known MHC IIB loci. For simplicity reasons we refer to different sequence variants as alleles, although they may originate from different loci.

Selection of suitable reference alleles for RSCA

The selection of suitable reference alleles is a crucial step in the optimization of RSCA. Selection criteria were: First, they should represent

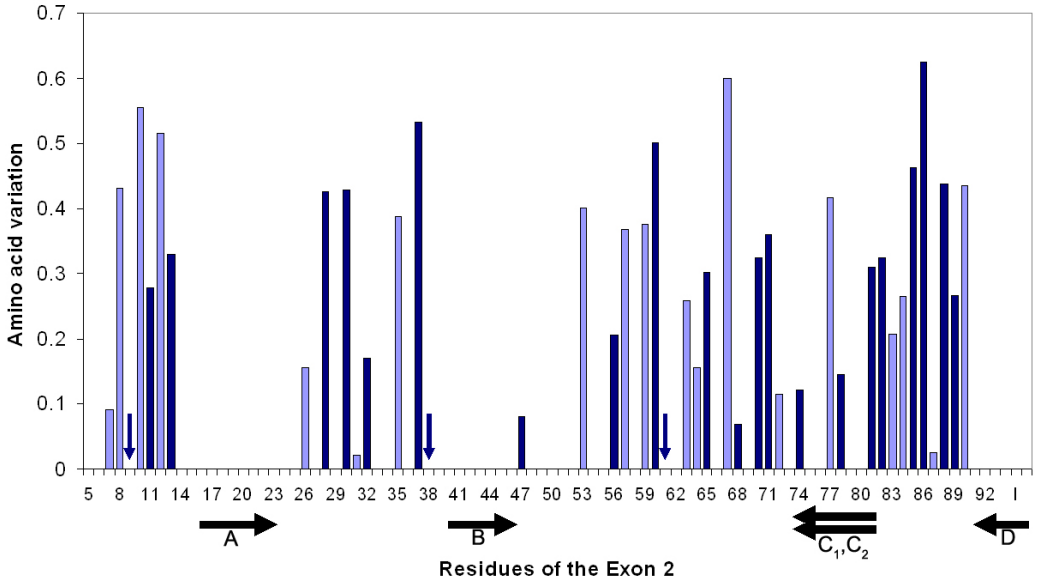


Figure II.1: Variation of amino acid residues in exon 2 of the MHC IIB genes in the three-spined stickleback *Gasterosteus aculeatus*. Included are 31 sequence variants from local populations. The variation for each residue is based on the number and frequency of substitutions and is calculated as y = 1 – Valdar01 score, as determined with the Scorecons server by Valdar (2002). Dark bars and arrows indicate residues involved in antigen binding, according to Brown *et al.* (1993). Black arrows at the bottom show positions of the primers used for amplification of the Exon 2: A (GAIIEx2startF) and D (GAIIExon2R_RSCA) are the new primers designed for this study; B and C_1,C_2 are the primers used in previous CE-SSCP typings of the stickleback MHC (Binz *et al.* 2001; Reusch *et al.* 2001). For simplification, the variation plot was extended by 6 bases into the adjacent intron 2 (I, right side of panel). Despite the fact that an amino acid variation score cannot be calculated for an intron sequence, these 6 base pairs were fully conserved in all local sequences investigated so far.

sequence variants that do not occur in the screened populations. Second, they should not be too genetically distant to assure reliable hybridisation. And

third, to increase resolution, the individual reference sequences should be as dissimilar from each other as possible. Due to the trans-species polymorphism of the MHC (Klein *et al.* 1998) it is even possible to use reference sequences from closely related species (Baquero *et al.* 2006), but this increases the risk that certain alleles do not hybridize, which again increases with the complexity of the template (number of alleles) because of competition between alleles during the hybridisation reaction. For our study we tested nine cloned MHC IIB sequence variants of three-spined sticklebacks from a West Canadian population (TBH Reusch & T Reimchen, unpublished data), which due to long divergence time (Mäkinen & Merilä

2008) are unlikely to carry any European alleles. Due to the existence of a database of more than 120 sequence variants that have been sequenced over the years ((Reusch *et al.* 2001; Reusch *et al.* 2004; Reusch & Langefors 2005) and TBH Reusch & KM Wegner, unpublished data), we were able to compare the genetic distance between the Canadian alleles and our local ones, to be able to choose candidate alleles that differed from each other more, but from the local alleles on average less than the population average. Additionally we estimated the distribution of the genetic distance between each reference and all the known alleles from our local populations to choose the flattest and broadest distribution (see Fig II.2 for examples). A broad distribution promises the highest resolution of alleles, because genetic distance and molecule mobility in the gel matrix are highly correlated due to the three-dimensional structure of the heteroduplex that increases in complexity with increasing mismatches between the reference and the target strand. A flat and broad distribution of the genetic distances between reference allele and the allele test pool therefore leads to a broader distribution of allele peaks and avoids overlap between them.

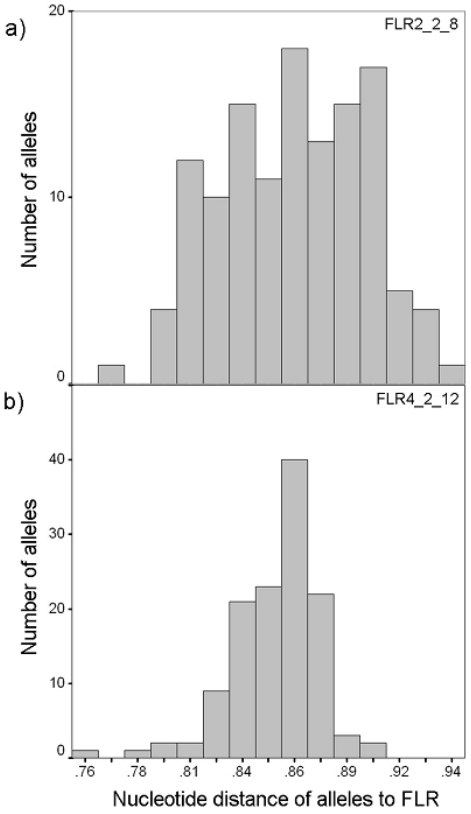


Figure II.2: Suitability of reference alleles (FLR). Distribution of genetic distances between alleles in the collected library and a) a suitable FLR with a broad distribution of genetic distances, b) an unsuitable FLR with a narrow distribution. Distance is measured in pair-wise nucleotide p-distance. Note the different scale of the y-axis between panel a) and b).

General RSCA protocol

Fluorescent labelled reference strands (FLR)

To enable high-throughput genotyping via capillary electrophoresis on an automated sequencer, the reference strands were fluorescently labelled. The templates

for the FLRs were plasmids with a single sequence variant that were obtained by cloning three-spined sticklebacks from a Canadian population (TBH Reusch & T Reimchen, unpublished data). Suitable plasmids were selected according to the criteria outlined above and amplified with the same primer pair as the unknown alleles, except that the forward primer was fam-labelled. The PCR mix for a 50 µl reaction volume contained 5 µl diluted plasmid (~10 ng/µl), 1x GeneAmp PCR Buffer II (Applied Biosystems), 5 mM $MgCl_2$, 50 µM of each dNTP, 0.5 µM of each primer and 2.5 units of AmpliTaq Gold (Applied Biosystems). The following PCR program was used: 95°C for 10 minutes to activate the hot-start polymerase, 33 cycles of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 60 seconds with a final extension step of 72°C for 5 minutes. We ran our programs on the thermal cyclers PC-200 (Bio-Rad, Munich, Germany) or LabCycler (SensoQuest, Göttingen, Germany). The PCR products were purified with the NucleoSpin Extract II Kit (Macherey Nagel, Düren, Germany) and eluted in 100 µl HPLC grade water (Mallinckrodt Baker, USA). The purified FLRs were kept at -20°C until further use.

Amplification and hybridisation

For genotyping of new three-spined stickleback individuals, we amplified exon 2 of the MHC class II loci from genomic DNA. DNA was extracted from ethanol- preserved muscle tissue with the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany). For amplification of the exon 2 we used the newly developed primer pair GAIIEx2startF and GAIIExon2R_RSCA (see above, 247 bp product). A reaction volume of 25 µl contained 90 ng of DNA, 1x GeneAmp PCR Buffer II, 5 mM $MgCl_2$, 50 µM of each dNTP, 0.5 µM of each primer and 1 unit of AmpliTaq Gold polymerase. The following PCR program was used: 95°C for 10 minutes to activate the hot-start polymerase, 27 cycles of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 60 seconds with a final extension step of 72°C for 5 minutes. The low number of PCR cycles was chosen to avoid formation of PCR artefacts (Judo *et al.* 1998; Lenz & Becker 2008). A reconditioning PCR step against artefact formation [20(Thompson *et al.* 2002) was omitted in our protocol, because heteroduplexes that form after the last PCR cycle can be neglected due to the subsequent hybridisation reaction of target and FLR in the RSCA protocol.

Next, the appropriate amount of PCR product was mixed with the FLRs. This has to be adjusted for each FLR, because the hybridisation efficiency is altered by the average genetic distance and GC content between target alleles and FLR. The ratio depends also on the concentrations of the PCR product and the FLR. Eventually, a good ratio should give equal heights of homo- and heteroduplexes in RSCA. In our study we used 6 or 8 µl PCR product, depending on the FLR (1 µl each). The hybridisation started with a denaturation step at 95°C for 10 minutes, then the heteroduplex formation is facilitated by a slow cooling of 2°/sec to 55°C, which is subsequently held for 20 minutes. A final cooling step at 4°C for 15 minutes assures stabilisation of the heteroduplexes. This hybridisation product is stable for several hours at 4°C and for several days at -20°C.

Capillary electrophoresis

The separation of heteroduplexes was performed on a model ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) with a 36 cm capillary. The Conformation Analysis Polymer (CAP, Applied Biosystems) provided the non-denaturing matrix in which the heteroduplexes migrate according to their tertiary structure. We used a 5% polymer: 5 g 9% CAP (prewarmed to room temperature), 2.16 g Urea (Sigma-Aldrich, Steinheim, Germany), 0.95 g HPLC grade water and 0.9 g 10x Running Buffer (Applied Biosystems). For each sample 1.5 µl of the hybridisation product was mixed with 0.3 µl GS1000 Rox size standard (Applied Biosystems) and 9.7 µl HPLC grade water. The internal size standard ensures proper alignment of heteroduplex peaks and minimizes between-run variation. The running conditions were: 18°C run temperature, 15 kV injection voltage, 15 sec injection time, 10 kV run voltage. The hybridisation products of each FLR had to be run separately, otherwise the antisense strand of one FLR would hybridise with the labelled strand of another FLR and produce a heteroduplex peak that cannot be differentiated from real alleles.

Plasmid library for sequence specific genotyping

A plasmid library of MHC class II exon 2 sequence variants has been collected over the years from cloned three-spined sticklebacks (TBH Reusch,

KM Wegner, C Eizaguirre & TL Lenz, unpublished data). It comprises currently 83 distinct variants that differ to varying extents (1 to 47 of 203 bp). Of the 3,403 potentially possible allele combinations in a pairwise comparison, only 0.7% differ by less than 3 bp (17 by 1 bp and 8 by 2 bp). We recorded mobility values for each sequence variant with each of the three FLRs. This was done four times independently using different polymer lots to determine the run-to-run variability of the mobility values. To estimate the resolution and specificity of the new typing method, we compared the difference in mobility values of alleles pair-wise in all 3,403 possible combinations. Those allele pairs that differed for all three FLRs in their mean mobility values by less than their combined standard deviation were assigned to be undistinguishable.

RSCA typing procedure

MHC II exons 2 of individual three-spined sticklebacks were amplified with the newly developed primers and processed according to the protocol outlined above. Mobility values for each heteroduplex peak were recorded after alignment of the internal size standard. Using the mobility values from the three FLRs and the library with mobility values of the collected alleles, we assigned allele identity to the heteroduplexes (see supplementary Fig II.4 for an example). However, the identity of an allele was only assigned if all three corresponding mobility values (+/- 1 SD) from the library were found in the individual. In cases where not all three values from the allele library could be found in an individual, we marked the heteroduplex peak as new allele. Individuals with such unknown heteroduplexes were subsequently cloned to identify the sequence of the new allele.

Cloning of fish to verify RSCA typing

To verify our RSCA genotyping protocol, we compared it with the result of 23 cloned and sequenced sticklebacks from different locations and time points. Ten of these fish had been genotyped in one of our former studies to establish a reliable amplification and cloning protocol for multi-locus templates (Lenz & Becker 2008). That study involved a protocol to avoid artefact formation during PCR (with other primers for amplification of a shorter exon 2 fragment), cloning and then sequencing of on average 89 clones per

individual. We also cloned an additional set of 13 three-spined sticklebacks, some of which, according to RSCA typing possessed so far unknown alleles. For the amplification and cloning of these fish, we followed the protocol in Lenz & Becker (2008), but with the new primer pair developed for this study (see above), only two independent amplification reactions and 48 clones per fish for sequencing. Forty-eight clones is a threshold that we determined by applying accumulation curves on data from the first cloning set (Lenz & Becker 2008), which resulted in 99.99% probability to have typed all present sequence variants (data not shown).

Applied software

For alignment and estimation of genetic distance between sequences, we used BioEdit7 (Hall 1999). RSCA chromatograms were aligned along the peaks of the internal size standard and analysed with GeneMarker 1.6 (SoftGenetics, PA State College, USA). Due to the non-denaturing feature of the polymer, the manufacturer's established values of the size standard did not match the peak pattern in our analysis, therefore the values for the size standard peaks were assigned arbitrarily, starting with 1,000 for the longest fragment. The software Primer 6 (Clarke & Gorley 2006) was used to calculate accumulation curves based on clone data from a previous cloning project (Lenz & Becker 2008). To obtain a bootstrap estimate of confidence, we computed 999 times the clones from two individuals with five alleles. We used the ScoreCons online server (Valdar 2002) to determine variation for amino acid residues of the exon 2. The software MultiLocus 1.22 (Agapow & Burt 2001) was used to estimate linkage disequilibrium between detected alleles; 1,000 randomizations were run.

Results

Plasmid library

The three mean mobility values for each of the 83 allele variants averaged over four independent runs are shown in Fig II.3. To estimate the resolution of the RSCA typing method, we compared all 83 alleles pair-wise. The mean mobility difference between alleles was 46.5, 42.3 and 27.3 units for the three FLRs. In comparison, the average standard deviation of the

independently obtained mobility values of an allele was 1.2, 1.0 and 1.4 units. In seven (0.2%) of 3,403 possible allele pairs, the two alleles were not distinguishable according to our definition outlined above. Six of them differed by 1 bp and one by 2 bp. The other 11 pairs with 1 bp and seven with 2 bp difference were distinguished unambiguously by using the three different FLRs (Fig II.3).

RSCA typing of selected fish

We detected 86 alleles by RSCA-typing 23 fish. The number of alleles per fish ranged from 2 to 5 with a median of 4 (Suppl. Tab 1). Within the 86 alleles, we found 28 distinct sequence variants, 15 of which occurred at least twice in different individuals. Using the allele library, the mobility values of 23 alleles were identified unambiguously, and therefore their sequences were determined (Tab 1). Five alleles were assigned as new according to our stringent selection criteria (see Methods section), and because their mobility value combination was not identified with the existing allele library. Four individuals carrying these five new alleles were therefore subjected to cloning to identify the sequence of the new alleles.

Comparison between cloning and RSCA

Cloning and sequencing of 23 individuals revealed 86 alleles, among them 27 distinct ones. All sequences showed highest similarity to stickleback MHC IIB exon 2 variants in a NCBI-BLAST search. Sixteen of the 27 alleles have been deposited already in GenBank (Suppl. Tab 1), and the remaining ones have been submitted during this study [GenBank:FJ360531 - FJ360541]. An additional very divergent sequence variant [GenBank:AF395709], which has been described before (Reusch & Langefors 2005), was detected by cloning and RSCA typing in every individual investigated. Due to sequence conservation, this variant was not addressed in this study. This sequence, which can also be found in the recently published genome sequence of the Alaskan stickleback, potentially originates from an invariant MHC locus that may have antigen processing function, similar to the invariant H2-M locus in mice (Hermel *et al.* 1995).

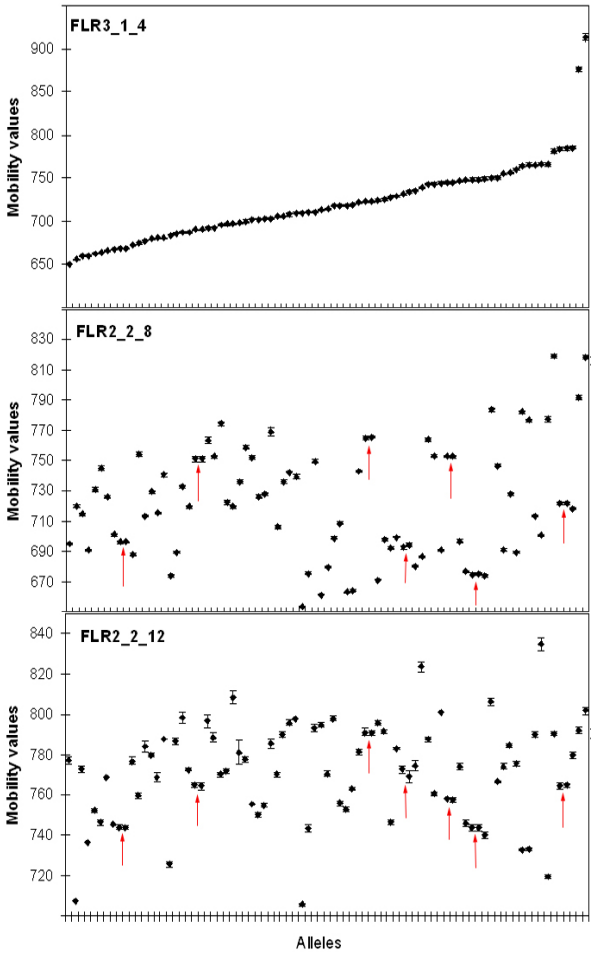


Figure II.3: Mobility values of 83 alleles with the selected three reference alleles (FLR). Mean and SD of each allele from four independent runs are shown. Arrows indicate the allele pairs that can not be differentiated with the three FLRs. Alleles are presented in the order of mobility with the reference FLR3_1_4 (top panel).

The remaining alleles differed in 1 to 46 nucleotides with a mean of 26, and in 1 to 27 amino acids (mean = 15.7; d_N/d_S ratio = 1.96; Z-test, p = 0.018). The results of the cloning confirmed the previous RSCA typing in 22 out of 23 individuals for allele number and allele identity, i.e. known alleles were recognised and unknown alleles were assigned as new and differentiated from each other. In one individual, a known allele (Neu51) was not recognised and assigned as new, although it was present in the allele library and had even been recognised in another fish of the same set (Suppl. Tab 1). This finding represents a conservative ambiguity that was resolved by following the proposed protocol, i.e. cloning the respective individual to resolve the possible new sequence variant. In this set of 23 individuals therefore 16,240 bases were typed correctly and another 1,218 bases (5 different alleles in 4 individuals) were scheduled for cloning by the proposed protocol.

The detected alleles showed strong linkage disequilibrium (I_A = 0.38, p<0.001). All alleles that occurred in more than one individual showed obvious haplotypic patterns with one to three variants per linkage group (Suppl. Tab 1). In total, 15 distinct haplotypes were differentiated, two of which shared one allele.

We also RSCA-typed the first ten fish, which had been cloned with a different primer pair in the earlier project [20], using the same primer pair to

enable direct comparison between methods. The same number of alleles per individual was observed in both methods, which shows that the RSCA typing protocol is reliable, independent of the primer pair used.

Discussion

In this study we developed a protocol for reliable genotyping of the highly polymorphic MHC class IIB genes of the three-spined stickleback using Reference Strand-mediated Conformation Analysis. To verify the results obtained from RSCA typing, we cloned and sequenced a total of 23 individuals to saturation. The congruence between RSCA and cloning in number as well as identity of detected alleles shows the reliability of the new typing protocol.

While the process of amplification, cloning, plasmid extraction, sequencing and alignment is time consuming and laborious, our RSCA protocol can handle a high number of fish in substantially less time and for less cost. Incorporating a library of collected sequence variants, we were able to unambiguously identify the sequence information underlying each peak or assign new alleles if applicable. Only in one case out of 86 alleles was an allele assigned as new, although its sequence was already known. This is due to the adopted stringent typing procedure and a conservative error, which was resolved by following the typing protocol, i.e. cloning and sequencing this individual.

The relatively high number of individuals with unknown alleles presented here (22% carried one or more new alleles) is due to the fact that several of these individuals were chosen for cloning, because they contained new alleles. In pilot screens of local populations (data not shown), the average fraction of individuals with new alleles ranged from 5-7%, i.e. only this fraction of a given sample set has to be cloned to obtain complete sequence information for all individuals. Evidently, the more complete the underlying allele library, the lower the fraction of fish with novel alleles.

In addition to the RSCA protocol, we also present a new primer combination, which is more universal than the ones previously used, because it amplifies all present alleles in a single reaction. These primers performed reliably in the local populations as seen in the data presented here and

enabled tracing of haplotypes over two generations in a large breeding project of several families from two different populations (data not shown). It remains to be tested whether they work satisfyingly across all stickleback populations from the Northern hemisphere, considering the divergence between European populations (Mäkinen & Merilä 2008) and even more between the Atlantic and the Pacific clade (Orti *et al.* 1994). The currently available MHC sequence information from the genome of an Alaskan individual and sequences of some individuals from British Columbia / Canada (TBH Reusch & T Reimchen, unpublished data), however, support the universality of our new primers. Using these primers in the current study, we found a median of 4 alleles over 23 individuals, which is slightly lower than the previously reported 5.8 alleles per individual detected by CE-SSCP (Wegner *et al.* 2003b). This might be due to the limited number of individuals in this study, but it might also indicate that the previously used combination of two reverse primers (Reusch *et al.* 2001) overestimated the total MHC diversity. This would be in agreement with a recent finding, which estimates the number of MHC class IIB loci in the three-spined stickleback to only 2-4 (Reusch & Langefors 2005).

We did not test whether or not all detected alleles are expressed at the mRNA level. However, a previous screen of several families revealed that over 90% of the alleles detected by CE-SSCP are expressed (Wegner *et al.* 2003a), and there is no reason why this should not apply to the alleles detected by RSCA. In combination with the new typing technique, it remains to be tested whether all alleles are expressed in all organs of an individual, or whether there is locus-specific transcription regulation. This appears even more interesting in the light of previous results, which show substantial variation in overall MHC expression between families, organs and even haplotypes (Wegner *et al.* 2006). Such transcriptional variation receives increasing attention by evolutionary biologists, as it provides a potentially underestimated source for adaptability in natural populations (Oleksiak *et al.* 2002; Whitehead & Crawford 2006; Miller *et al.* 2007; Schlichting 2008; Wilhelm *et al.* 2008).

The first of the two only studies employing RSCA in non-mammalian species so far directly compared RSCA with SSCP typing, when genotyping the MHC class IIB locus in lake trout (Salvelinus namaycush Noakes *et al.*

2003). In the end, the authors favoured SSCP over RSCA, because it detected some additional alleles in the screened population. However, the existence of these additional alleles was not verified via sequencing. Moreover, the lake trout carries only one MHC IIB locus, which limits the number of detectable alleles to two per individual. The authors of the second study developed an RSCA typing protocol for both MHC class I and II in the red jungle fowl (*Gallus gallus*) and addressed two loci at a time (Worley *et al.* 2008). The authors tested congruency with cloning in eight individuals and conclude that RSCA is a reliable technique for MHC typing in the red jungle fowl.

In many species, the MHC consists of several loci in classes I and II (Kelley *et al.* 2005), which makes the allele pattern more complicated and increases the chance of overlapping signals using SSCP, a problem that occurred already with the single locus in the Lake trout (Noakes *et al.* 2003). Here, RSCA has a substantial advantage over other indirect typing techniques, because it provides several mobility values per allele, reducing the chance of two overlapping allele values and increasing resolution.

Reproducibility is also a major concern for indirect genotyping techniques and can be confounded by polymer lot variation, temperature fluctuation and other factors (Hennessy *et al.* 1998). Therefore, we measured the variation of mobility values for each allele across independent runs over several months. To test realistic laboratory conditions, we used different polymer lots, the most likely source of between-run variation, and a capillary that is in turns used also for fragment analysis with POP-4™ polymer (Applied Biosystems). By this we were able to show that between-run variation with RSCA is limited and can be overcome by the advantage of getting several mobility values from the different labelled references. We conclude that three different well-chosen reference strands are enough to differentiate 99.8% of the 3,403 potentially possible allele pairs and still 72% (18/25) of those allele pairs that differ only by one or two nucleotides (see Fig II.3). Employing more reference alleles would probably also differentiate the last alleles pairs, but we consider a resolution of 99.8% as high enough. Nevertheless, regular re-runs of known alleles / allele libraries are advisable to keep the mobility values “up to date” and counteract unreported chemistry changes by suppliers.

Overall, the close and reliable correspondence of the electrophoretic signal from capillary electrophoresis with the DNA sequence is a clear advantage of the new RSCA protocol over the previously used CE-SSCP typing. This sensitivity will enable future comparisons of allele pools from different populations and over different time points as well as analysis of evolutionary important parameters such as residue specific patterns of adaptation or the effect of genetic distance between alleles and individuals. Recent studies have revealed intriguing results on the functional aspects of sequence- and even residue-specific differences for mate choice (Forsberg *et al.* 2007; Neff *et al.* 2008; Schwensow *et al.* 2008a), parasite resistance (Consuegra & Garcia de Leaniz 2008) and adaptive divergence of species (Blais *et al.* 2007).

A new observation due to the high resolution in allele detection by RSCA is the fact that alleles seem to occur in linked haplotypes, differ in the number of sequence variants per haplotype and in one case share an allele between haplotypes. This provides a hint for strong linkage disequilibrium between loci and low but yet occurring recombination in nature, an observation that is in agreement with previous results on the organisation of the MHC region in the stickleback (Reusch *et al.* 2004). More important, it might indicate variation in the number of loci between haplotypes and addresses a phenomenon that has already been described for other species (Malaga-Trillo *et al.* 1998; Bontrop *et al.* 1999; Doxiadis *et al.* 2000; Bowen *et al.* 2004; Horton *et al.* 2008). This finding would explain the earlier reported large variance in allele numbers among individuals (Wegner *et al.* 2003b), and it reveals a potential mechanism of adaptation to changing pathogenic environments, which has first been termed by Klein *et al.* (1993a) as '*The Accordion Model of MHC Evolution*' and was later elaborated by Nei *et al.* (1997) to the Birth-and-Death-Model of Evolution. Nevertheless, these findings need further research at the genomic level, including more thorough analysis of the haplotype specific chromosomal organisation.

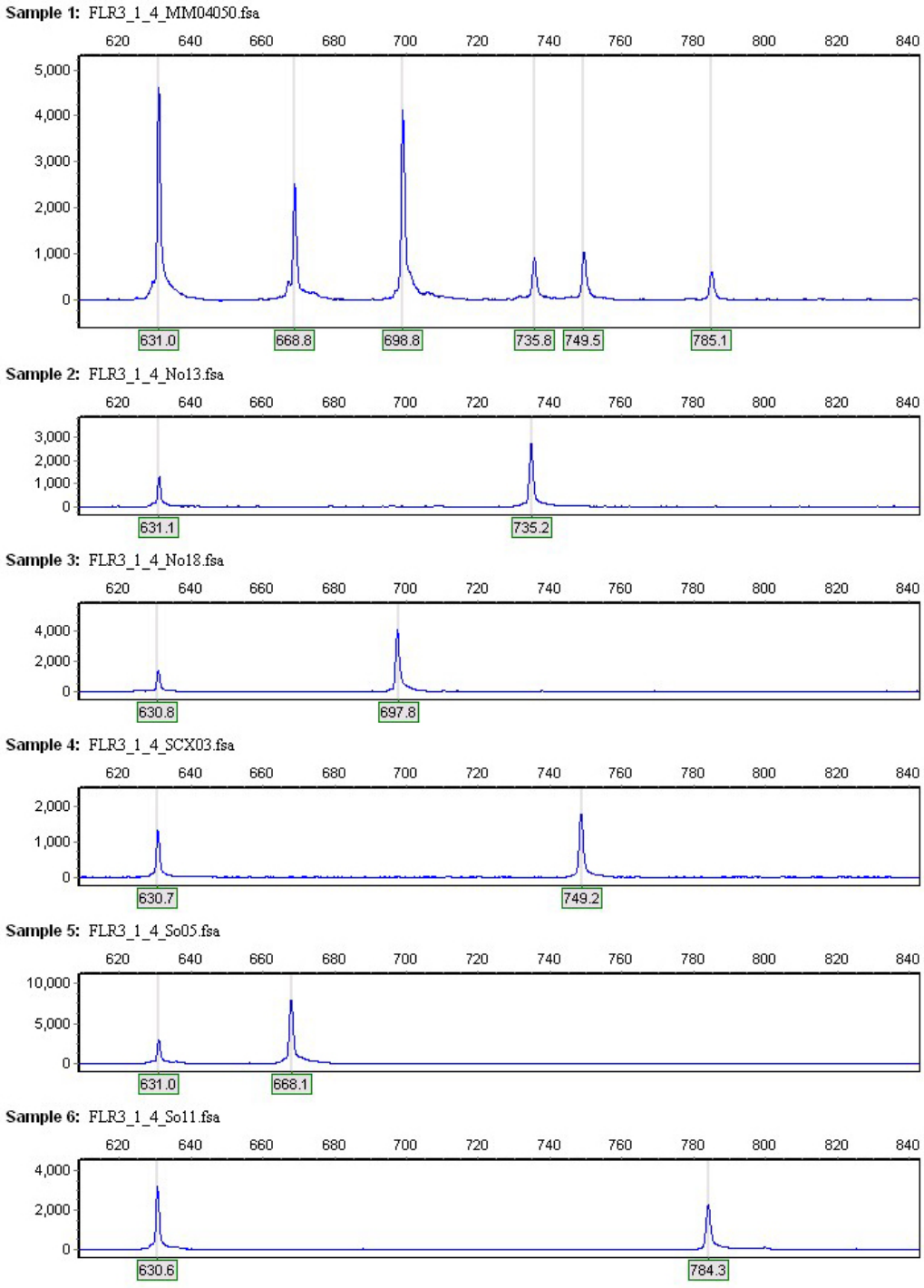
Conclusion

Here we present a new RSCA genotyping protocol for the highly polymorphic MHC genes of the three-spined stickleback *Gasterosteus*

aculeatus, which in combination with an established allele library provides sensitive and reliable allele data at the sequence level. Verification of the RSCA typing by cloning and sequencing shows high congruency between both techniques. Together with new insights from the polymorphic MHC of the three-spined stickleback, an emerging model system, this offers a resource for researchers to address questions of host-parasite co-evolution, local adaptation and ecological speciation.

Acknowledgements

We thank Manfred Milinski for fruitful discussions as well as support, and Thomas Boehm for proficient advice and helpful comments on the method development and the manuscript. Irene Samonte gave helpful comments on an earlier version of the manuscript. We are also grateful to Sybille Liedtke for invaluable help in the lab.



Supplementary Figure II.4: RSCA chromatograms. Top panel: typing of an individual with 5 alleles hybridised to one (of three) FLR. Panels below: five plasmids from the allele library hybridised to the same FLR. Each chromatogram shows the homoduplex of the FLR (leftmost peak) and five (top panel) or one (panels below) heteroduplex(es) for the individual alleles.

Supplementary Table 1: Three-spined stickleback individuals typed for comparison of RSCA and cloning. Given are Fish ID, the population and year in which they were caught, alleles as determined by RSCA and cloning, match between both typing methods and known accession numbers in GenBank. Note the strong linkage with one to three alleles per haplotype.

Fish ID	Population/Year	RSCA	Cloning	Match	AccessionNo
MM04031	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No15	No15	✓	DQ016410
		No16	No16	✓	DQ016417
MM04039	GPS/2004	No13	No13	✓	AF395711
		No18	No18	✓	AY687846
MM04047	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No13	No13	✓	AF395711
		No18	No18	✓	AY687846
MM04050	GPS/2004	No13	No13	✓	AF395711
		No18	No18	✓	AY687846
		So05	So05	✓	DQ016402
		So11	So11	✓	DQ016404
		SCX03	SCX03	✓	AJ230191
MM04060	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No13	No13	✓	AF395711
		No18	No18	✓	AY687846
MM04080	GPS/2004	No13	No13	✓	AF395711
		No18	No18	✓	AY687846
		No08	No08	✓	AY687842
		SCX15	SCX15	✓	EU541449
MM04083	GPS/2004	No15	No15	✓	DQ016410
		No16	No16	✓	DQ016417
MM04086	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No10	No10	✓	AF395722
		No11	No11	✓	AY687843
MM04087	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		So05	So05	✓	DQ016402
		So11	So11	✓	DQ016404
		SCX03	SCX03	✓	AJ230191
MM04131	GPS/2004	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No08	No08	✓	AY687842
		SCX15	SCX15	✓	EU541449
R027	MA/2005	So06	So06	✓	unpublished
		So05	So05	✓	DQ016402
		SCX03	SCX03	✓	AJ230191
		So11	So11	✓	DQ016404
R110	MA/2005	No05	No05	✓	AY687829
		Neu51	Neu51	✓	AY687833
		SCX20	SCX20	✓	unpublished

S2MA07005	MA/2007	So02	So02	✓	DQ016426
		So05	So05	✓	DQ016402
		SCX03	SCX03	✓	AJ230191
		So11	So11	✓	DQ016404
S2MA07061	MA/2007	So06	So06	✓	unpublished
		So01	So01	✓	unpublished
		So10	So10	✓	unpublished
S2MA07066	MA/2007	No05	No05	✓	AY687829
		No42	No42	✓	unpublished
		new5	No45	new	unpublished
En07284	GPS/2007	No08	No08	✓	AY687842
		SCX15	SCX15	✓	EU541449
		No13	No13	✓	AF395711
		No18	No18	✓	AY687846
TL06121	GPS/2006	No05	No05	✓	AY687829
		No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
TL06165	GPS/2006	No08	No08	✓	AY687842
		SCX15	SCX15	✓	EU541449
		No15	No15	✓	DQ016410
		No16	No16	✓	DQ016417
E10	GPS/2006	No08	No08	✓	AY687842
		SCX15	SCX15	✓	EU541449
		new4 *	Neu51	⊗	AY687833
		SCX20	SCX20	✓	unpublished
E47	GPS/2006	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No13	No13	✓	AF395711
		new1	No46	new	unpublished
E54	GPS/2006	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No13	No13	✓	AF395711
		new1	No46	new	unpublished
E62	GPS/2006	No01	No01	✓	DQ016399
		No12	No12	✓	DQ016400
		No43	No43	✓	unpublished
		No44	No44	✓	unpublished
E77	GPS/2006	No15	No15	✓	DQ016410
		No16	No16	✓	DQ016417
		new2	No48	new	unpublished
		new3	No49	new	unpublished
No of distinct alleles		28 *	27		
No of alleles typed		86	86		

* In one fish (E10), the allele Neu51 was determined as ‘new’ by RSCA typing, although it was present in the plasmid library. The same allele was correctly identified in another fish (R110). GPS – Großer Plöner See, MA – Malenter Au.

Chapter III

Effects of MHC diversity and specific haplotypes on immunocompetence, reproductive success and mate choice decisions of Three-spined sticklebacks under semi-natural conditions

Tobias L. Lenz, Christophe Eizaguirre, Jörn P. Scharsack, Martin Kalbe & Manfred Milinski

Abstract

Understanding the ubiquitous presence of sexual selection is still a major concern in evolutionary biology. One of the proposed driving forces for the evolution of mate choice, a crucial aspect of sexual selection, is its potential to increase immunocompetence in the offspring. The highly polymorphic major histocompatibility complex (MHC) with its pleiotropic role in mate choice and parasite resistance is therefore a perfect candidate to investigate the mechanisms that favour sexual selection. In the three-spined stickleback *Gasterosteus aculeatus*, both, mate choice and parasite resistance have been experimentally shown to be influenced by MHC genetics. To investigated the role of MHC based '*good genes*' and '*compatible genes*', we let sticklebacks with specific MHC IIB genotypes reproduce in an outdoor enclosure system, where they were protected from predators but encountered their natural parasites. In addition to a positive effect of MHC diversity on body size, which directly translated in reproductive success, we found mate choice for an intermediate genetic distance between parental MHC genotypes. To understand the underlying mechanisms, we also investigated immunological parameters and found the less efficient innate immune system to be more active in individuals with a genetically more diverse MHC allele repertoire, potentially explaining the mate choice for intermediate MHC distance. This important function of the MHC in natural mating decisions has potentially evolved to increase offspring immunocompetence in the ongoing arms race between hosts and their parasites.

Unpublished manuscript.

Introduction

Sexual selection in the context of mate choice is a costly, but common behaviour in natural populations and still represents one of the central concerns in evolutionary biology (Maynard Smith 1991; Ryan 2004). It has been proposed that one of the driving forces promoting sexual selection is the co-evolution between hosts and parasites (Hamilton & Zuk 1982; Read 1987; Milinski & Bakker 1990; Møller 1990a). Mate choice for detectable honest signals that reflect for instance heritable parasite resistance would provide a direct benefit for the next host generation by providing it with increased resistance (Read 1988; Hamilton *et al.* 1990; Møller 1990b). Two different components of such heritable genetic quality have been identified and described: '*good genes*' and '*compatible genes*' (reviewed in Neff & Pitcher 2005). The benefit of '*good genes*' is independent of the rest of the genome and provides the bearer of these genes with an additive advantage in lifetime reproductive success (Zahavi 1975; Hamilton & Zuk 1982; Møller & Alatalo 1999; Kokko 2001). The effect of '*compatible genes*' in contrast is non-additive and depends on the genetic environment: Only a combination of certain alleles of both partners is advantageous (Tregenza & Wedell 2000). Consequently, this requires a more direct way to assess genetic quality in the mate with respect to the own genotype and mate choice only for honest signals would not always be an optimal strategy in this context.

One of the rare traits that are simultaneously involved in parasite resistance and mate choice is the major histocompatibility complex (MHC, reviewed in Apanius *et al.* 1997; Edwards & Hedrick 1998; Penn & Potts 1999; Milinski 2006). The MHC is a gene dense region in the vertebrate genome and its most polymorphic genes code for cell surface molecules that are responsible for the presentation of endogenously and exogenously derived antigens to T-cells (Klein 1986; Janeway *et al.* 2005). Due to this function, the MHC plays a vital role in the recognition of pathogens invading the body and is a key component of the acquired immune system. It has been studied extensively in model species under laboratory conditions (Apanius *et al.* 1997), but also became the focus of an increasing number of studies on natural populations because of its functional importance for the immune system and mate choice (Reviewed in Bernatchez & Landry 2003; Piertney &

Oliver 2005; Sommer 2005a; Milinski 2006). The high polymorphism in the MHC is composed of a high total number of alleles on the population level as well as heterozygosity on the individual level. It is thought to be maintained by balancing selection and to facilitate populations and individuals to adapt to and resist the diverse pathogen communities in all natural habitats. A number of hypotheses have been proposed to explain the forces that counteract the effects of random genetic drift and fixation (reviewed in e.g. Potts & Wakeland 1990; Edwards & Hedrick 1998; Bernatchez & Landry 2003; Piertney & Oliver 2005; Sommer 2005a; Milinski 2006). The '*negative frequency-dependent selection*' hypothesis (Clarke & Kirby 1966; Takahata & Nei 1990; Slade & McCallum 1992) describes the cyclic patterns of host parasite co-evolution ('*Red Queen hypothesis*', Van Valen 1973), stating that frequencies of resistance alleles constantly change with the frequency of adapted and non-adapted pathogens. The prerequisite for this co-evolution in terms of reciprocal adaptation is the direct interaction of individual alleles and distinct parasites. Growing evidence from a variety of taxa supports this prerequisite (e.g. Hill *et al.* 1991; Thursz *et al.* 1995; Langefors *et al.* 2001; Harf & Sommer 2005; Westerdahl *et al.* 2005; Young *et al.* 2005; Bonneaud *et al.* 2006b; Schwensow *et al.* 2007; Loiseau *et al.* 2008). Another hypothesis, the *'heterozygosity advantage*' has been proposed as a general mechanism to maintain high allele numbers in populations by an advantage for individuals carrying more than one allele, enabling them to present a broader range of antigens to T-cells (Doherty & Zinkernagel 1975b; Klein & Figueroa 1986). There is experimental evidence that this is valid for animals exposed to multiple pathogens (Penn *et al.* 2002; McClelland *et al.* 2003) but not to single pathogens (Wedekind *et al.* 2005), which seems plausible, as the advantage would most likely result from a combined resistance of individual alleles each presenting a certain range of parasites (Hughes & Nei 1988, 1989). However, there is also a constraint against maximal MHC diversity, which is the negative T-cell selection during thymic development (Janeway *et al.* 2005) that increases selection against a too diverse MHC repertoire (Lawlor *et al.* 1990; Nowak *et al.* 1992; Woelfing *et al.* 2009). In fact a number of studies has shown an optimal instead of a maximal MHC diversity to be advantageous (Wegner *et al.* 2003a; Buchholz *et al.* 2004; Madsen & Ujvari

2006; Wegner *et al.* 2008) and one study even found a disadvantage for maximal MHC diversity (Ilmonen *et al.* 2007). Potentially, these two mechanisms are acting in synteny (Apanius *et al.* 1997) and might vary in their contribution, for instance in response to habitat heterogeneity or life history strategies, which both can affect parasite exposure and therefore immunogenetic adaptation.

The role of MHC in mate choice potentially amplifies and accelerates the effects of these different mechanisms and enables a faster adaptive host response in the co-evolutionary arms race against parasites (Milinski 2006). The influence of MHC on mating decisions has been shown in various species (but see also Westerdahl 2004; Sommer 2005b), however the nature of selection is not yet fully understood as evidence exists for both heterozygote advantage (Potts *et al.* 1991; Penn & Potts 1998; Consuegra & Garcia de Leaniz 2008; Neff *et al.* 2008; Schwensow *et al.* 2008a) and immunogenetic optimality (Bonneaud *et al.* 2006a; Forsberg *et al.* 2007). The Three-spined stickleback (*Gasterosteus aculeatus*) is a highly suitable model system to explore the role of mate choice in the maintenance of MHC diversity, because a number of experimental studies have already shown that MHC dependent olfactory mate choice for '*compatible genes*' plays a crucial role in this species (Aeschlimann *et al.* 2003; Milinski *et al.* 2005).

As an immunological advantage for certain MHC alleles has been shown in several species, the question arises whether females also select mates with '*good genes*' to increase the immunocompetence of their offspring. First evidence has been reported from birds (Ekblom *et al.* 2004), but the link to parasites is lacking. In a previous experiment we found first evidence that females are combining the '*good genes*' and the '*compatible genes*' strategies in their mate choice (Eizaguirre *et al.* Submitted-b, Annex I). In this study we therefore aimed at investigating the role of specific MHC haplotypes ('*good genes*'), but also distinct MHC dissimilarity ('*compatible genes*') in mate choice under semi-natural conditions. We stocked six enclosure cages in a natural breeding area of Three-spined sticklebacks with individuals that were only selected for specific MHC genotypes. The enclosure system allows natural behaviour, such as mate choice and reproduction under natural parasite pressure and food supply, but without confounding effects of

predation by birds or fish (Kalbe *et al.* In press). Eggs were collected over five weeks and mate-choice decisions inferred from parenthood analysis. To better understand the underlying mechanisms, through which the MHC affects parasitisation and mate choice, we also assessed immunological parameters, such as ratio of granulocytes to lymphocytes in the head-kidney, lymphocyte proliferation and respiratory burst activity. These parameters reflect the state of the innate and adaptive immune system more directly and are known to be influenced by parasite infection (Kurtz *et al.* 2004; Scharsack *et al.* 2004).

In local latitudes, three-spined sticklebacks usually experience only one reproductive period during their life, which is the summer following their first winter (Wootton 1984). Therefore the here presented data from a full reproductive period gives a good estimate of the lifetime reproductive success of an individual and should reflect the naturally evolved mating strategy (Kalbe *et al.* In press).

Materials and Methods

Experimental fish

The fish for this experiment were collected in January 2007 from the lake Großer Plöner See (54° 9'21.16"N, 10°25'50.14"E) in Northern Germany. Fish were trapped in the same shallow area in which the semi-natural enclosures for this experiment were located (see also Kalbe *et al.* In press). They were kept in the lab in a 90 l aquarium with constant freshwater supply and under winter conditions (6° C, 6/18 hours light/dark regime) until beginning of June. Feeding was ad libitum with frozen insect larvae. After two weeks of artificial spring conditions (12° C, 12/12 hours light/dark regime) these fish were brought to artificial summer (18° C, 15/9 hours light/dark regime) to adapt their physiology before getting out into the lake enclosures. From spring condition onwards the fish were isolated singly in 16 l tanks, but still with same feeding and constant water supply.

MHC genotyping and selection of fish

To stock the enclosures with fish that carry specific MHC class IIB genotypes, we determined the MHC genotype of 288 individual three-spined sticklebacks. DNA was extracted from the tip of a dorsal spine using the

DNeasy 96 well Tissue Kit from Qiagen (Hilden, Germany). The MHC typing was done using reference strand mediated conformation analysis (RSCA) as described in Lenz *et al*. (Chapter II). In brief, we amplified most of the exon 2 of the MHC class IIB genes, which are known to have undergone recent dupilcation in the three-spined stickleback (Reusch *et al.* 2004; Reusch & Langefors 2005). We took cautious measures to avoid artefact formation as described in Lenz & Becker (2008, Chapter I). In addition, we sex-typed these fish genetically to achieve a controlled sex ratio with an adjusted protocol from Griffiths et al. (2000), using fluorescently labelled primers for high-throughput typing on an automated sequencer (Eizaguirre *et al*., Chapter IV). 102 fish were finally selected and prepared for the enclosure experiment. All fish were measured for length, height and weight prior to release into the enclosures on 4th of July.

Experimental design

We used six enclosures to investigate MHC dependent mate choice under near-natural conditions. Using a newly developed genotyping protocol (Lenz *et al*., chapter II) we were able to screen a large number of fish for their MHC genotypes and select a subset for each enclosure. The new MHC typing protocol allows not only to identify alleles to the sequence level, but also to identify distinct haplotypes. In all enclosures, 9 males were paired with 8 females. Perfect replication in terms of genotype combinations for all six enclosures was not possible due to the natural diversity in the MHC. Nevertheless, the genotype composition did not differ substantially between Enclosures (Fisher's Exact Test, p = 1). Accordingly four enclosures received 6 heterozygous (4 alleles) and 2 homozygous (2 alleles) females together with 7 heterozygous and 2 homozygous males. The other two enclosures were stocked with three homozygous females, but only one homozygous male, nevertheless still keeping the same sex ratio (see Tab III.1). The MHC genotype composition was chosen such that most females were provided with mates of all three possible different genetic similarities: males with identical, semi-dissimilar or fully dissimilar MHC genotypes, which would result in minimal diversity, intermediate diversity and maximum diversity in the offspring respectively. In total only 7 different haplotypes entered the system,

a selection that reflects the most common haplotypes in the natural population at that time, as determined from the 288 randomly collected and screened individuals. Included was also a haplotype that had shown significant advantages against specific parasites and for reproductive success in the previous year (Eizaguirre *et al.* Submitted-b, Annex I). Additionally male allele diversity ranged from 2 to 5 MHC alleles (Tab III.1), representing approximately the natural range. This design therefore allowed females, themselves showing a limited diversity, to choose between different haplotypes (potential '*good genes*'), different similarity to themselves ('*compatible genes*') and a natural range of allelic diversity in the available mates.

Estimating mate choice and reproductive success

During the whole experimental period, we collected the egg clutches that were laid by the females and determined parenthood to estimate mate choice and reproductive success for each individual fish. Enclosures were checked weekly, to assure that all clutches were collected. Normal egg development until hatching at local temperatures (18-20°C) takes about eight

Table III.1: Stocking design for enclosures. a) Genotype distribution over the six enclosure cages. b) The MHCIIB allele composition of the selected haplotypes (given are GenBank accession numbers).

a) Genotypes per enclosure cage

Enclosure1		Enclosure2		Enclosure3		Enclosure4		Enclosure5		Enclosure6	
Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
AA	AA	AB	AA	AA	AA	AB	AA	AB	AA	AA	AA
AB	AB	AC	AB	AB	AB	AC	AB	AC	AB	AB	AB
AC	AC	AD	AC	AC	AC	AD	AC	AD	AC	AC	AC
AD	AD	AF	AD	AD	AD	BB	AD	AF	AE	AD	AD
AG	AE	BE	AE	AG	AE	BE	AE	BC	BB	AG	AE
BB	BB	BE	BB	BB	BB	BF	BB	BE	BB	BB	BB
BE	BC	BF	BB	BE	BC	BF	BC	BF	BC	BE	BC
BF	BE	CC	BC	BF	BE	CC	BE	CG	BE	BF	BE
CF		CG		CF		CG		DD		CF	

b) Alleles per genotype (accession numbers in GenBank)

A	AF395711, AY687846	E	AF395718, No31 (unknown in GenBank)
B (= F10)	DQ016399, DQ016400	F	AY687829
C	DQ016410, DQ016417	G	DQ016402, DQ016404, AJ230191
D	AY687842, EU541449		

days (Wootton 1984). Clutches were carefully collected from nests and brought to the lab for subsequent processing. Nests were placed back in the exact location where they had been found. The processing included separation of different developmental egg stages within each nest batch as they represent different clutches from different females (Kalbe *et al.* In press; Eizaguirre *et al.* Submitted-b, Annex I). Eggs were kept in aereated glass jars under Malachit-Green treatment (0.04 ppm) against fungus until they reached a developmental stage with visible dark eyes and neural tube to assure sufficient material for DNA extraction. Each clutch was weighed and photographed for later estimation of clutch size and DNA was extracted from a random subset of 24 eggs from each clutch to determine parenthood. Extraction was done with the Invisorb DNA Tissue HTS 96 Kit (Invitek, Germany) on the automated platform FreedomEvo (Tecan, Switzerland).

For parenthood analysis, we employed seven microsatellite markers (Stich4170, Stich1125, Stich5196, Stich1097, Stich7033, Largiadèr *et al.* 1999; Stn18, Stn75, Peichel *et al.* 2001), which are included in two multiplex PCR protocols (Reusch *et al.* 2001b; Kalbe *et al.* In press). Allele call was done in Genemarker 1.6 (Softgenetics). We used the software PAPA (Duchesne *et al.* 2002) to determine the most likely parental pair within each enclosure. For each nest the male with the highest proportion of fertilized eggs was assigned as the nest owner and therefore assumed to have been chosen by females. This designation has been verified by observations in an earlier experiment (Kalbe *et al.* In press). Eggs that were fertilized by another male than the nest owner were assigned as the product of sneaking, a common behaviour in sticklebacks (Wootton 1984).

Recapture and dissections

After five weeks the number of detected clutches had decreased substantially. The reproductive period was therefore ended and parental fish were recaptured and brought back into the lab. To check for parasite, we dissected the fish within three days after recapture as described by Kalbe et al. (2002). External and internal macroparasites were identified microscopically to the lowest feasible taxonomic level. The comprehensive screen included skin and fins, mouth, eyes and gills as well as body cavity

and internal organs including muscle tissue. Fish were measured again for length, height and weight.

Immune assays

During dissection, we isolated leucocytes from head kidney (HKL) according to Scharsack et *al.*'s protocol (2004). Total cell numbers in HKL isolates were determined with a standard cell dilution assay according to Scharsack et *al*. (2007). Additionally, flow cytometric measurements of freshly isolated HKL were used to determine proportions of granulocytes (FSC/SSC-high) and lymphocytes (FSC/SSC-low) in individual HKL samples.

We quantified the respiratory burst of HKL as important factor of the cell mediated innate immunity (Kurtz *et al.* 2004). As a parameter for activation of the adaptive immune system, frequencies of lymphocytes in G0-1, S and G2-M phases were acquired by DNA content analysis of red fluorescence intensity (propidium iodide labelling) of single cells from the lymphocyte gate (Scharsack *et al.* 2007).

Statistical analysis

- Parasite load

We calculated the Shannon Index over all detected parasite taxa, using the vegan package in software R statistics (Oksanen *et al.* 2008), to estimate the diversity of the parasite infection within individual sticklebacks. This index takes distribution and abundance of parasite infection into account and enables the comparison of infection between individuals. We also used the plain number of parasite taxa as an estimate for infection richness, which gives a more direct idea of the antigenic diversity each fish was able to fight against. These parasite parameters were analysed in linear mixed effect models, with enclosure as random effect to account for differences between enclosures. Independent parameters were sex, corrected initial body size, microsatellite heterozygosity, MHC diversity/haplotype and immunological parameters. To compare parasite assemblages between sexes or other grouping factors, we used an analysis of similarity (ANOSIM) on Bray-Curtis similarity matrix as implemented in PRIMER 6 (Clarke & Gorley 2006).

- MHC

As an estimate for individual MHC diversity besides the overall number of alleles, we used the mean genetic distance within an individual. For this we calculated the pairwise genetic distances between all alleles of an individual based on amino acid p-distance, using the software MEGA 4 (Tamura *et al.* 2007), and divided the sum by the number of pairwise distances. This gives a more precise estimate of the difference between alleles. Allele number and mean genetic distance were used as independent factors with linear and quadratic fit in models explaining parasite load, body size, reproductive success and immunological parameters.

- Body condition

We produced a principal component (PC) vector from body length and height using JMP 5.0 (SAS Institute Inc.), which expresses the body size more comprehensively than the simple length. However, all three body size parameters were highly correlated. To follow the development of the fish during the reproductive period we estimated the growth by subtracting initial length and height from the values measured at the end of the experiment and took again the first principal component of this parameter pair. The initial body size (before the fish were set out) was analysed in an ANCOVA with sex, microsatellite heterozygosity and number of MHC class II alleles or mean genetic distance as predictors. Growth was analysed in a similar way, but including the parasite parameters and using a mixed effect model with enclosure as random factor.

For all other post stocking parameters like for instance reproductive success or immunological parameters, we also controlled for the effect of differences between enclosures by using linear mixed effect models with enclosure as random effect. Test assumptions were verified and the according tests were conducted in R 2.6 (R Development Core Team 2007).

- Mate choice

To test whether mate choice was non-random, we simulated 1000 random pairings based on the observed mating distribution. We did the same

simulation for a random mating pattern based on the actually available mates within each enclosure. Using this approach we produced mating distributions for the MHC genetic distance as well as for the combined heterozygosity between mates (0 = identical genotypes, 1 = 1 haplotype difference, 2 = totally different genotypes). The simulated distributions based on the observed and theoretical matings were then compared with a Chi-square test (simulated p-value on 1000 replications). Additionally we analysed whether mate choice was influenced by overall relatedness. For this we used the method by Queller & Goodnight, implemented in the software IDENTIX (Belkhir 2002) and the same seven microsatellite markers that were used for parenthood analysis.

To investigate in which direction mate choice was aiming, we employed the approach taken by Forsberg et al. (2007) and compared the observed MHC diversity between mates against a simulated distribution which represented random mating. This distribution was produced by replicating 10,000 times the actual number of mating events from all theoretically possible pairs. Then we compared the observed mean against the distribution and referred significance from the 95% confidence intervals. This was done for the mean of the combined number of MHC alleles and for mean average genetic distance between mates This was calculated in a similar way as the individual genetic distance, just that here the basis were the pairwise distances between the mates and not the distances with an invidual. We did the same analysis for the variance of these two parameters since a lower than random variance would indicate specific choice.

Results

Of the 102 initially stocked individuals, we recaptured 86 alive. On the last day of egg collection 5 individuals were found dead. An additional 2 individuals had been found dead during egg collection in the previous weeks. The mortality did not differ between the sexes, nor was it influenced by any genetic parameter. For analysis of the initial body size all 102 individuals were used. The calculated PC for body size explained 97.9% of the covariance of body length and height. For analysis of all post stocking parameters, we used only the 86 survivors.

Body condition

The initial body size, i.e. the condition at the time when the fish were released into the enclosures, was higher in females ($F_{1,102}$= 10.38, p = 0.002) and correlated positively linear with individual number of MHC IIB alleles ($F_{1,102}$ = 7.47, p = 0.007; Fig III.1), but was independent of genome-wide heterozygosity as measured by microsatellites. When we looked at the effect of MHC diversity within sexes, we found that the linear positive correlation only stands in females ($F_{1,48}$ = 8.5, p = 0.005) while in males there was no effect of MHC diversity ($F_{1,54}$ = 1.2, p = 0.28). We checked for the same correlation in the full sample of 288 fish that were screened initially for this study and found again a strong trend in the linear correlation between body size and the number of MHCIIB alleles ($F_{1,288}$ = 3.8, p = 0.051). We did not find a significant effect of single haplotypes.

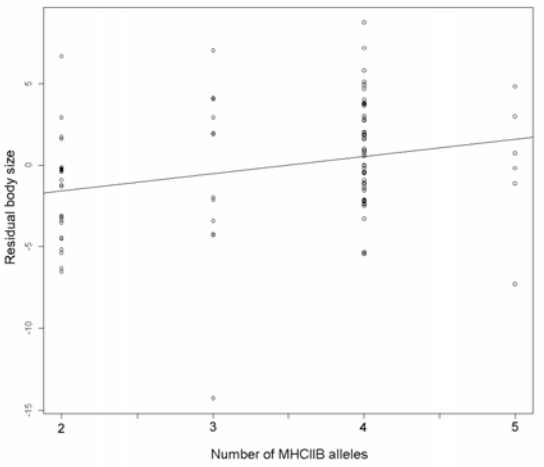


Figure III.1: Correlation between initial body size and individual number of MHC class IIB variants. Body size is corrected for sex.

The growth during the reproductive period was significantly different between sexes ($F_{1,86}$ = 45.2, p < 0.001) with males nearly stagnating (mean length gain 0.2 mm) and females increasing substantially (mean 2.0 mm). Growth was also negatively correlated to parasite diversity (F1,86 = 8.9, p = 0.004).

Parasitisation

We identified 21 distinct parasite taxa. The overall parasitisation in survivors, estimated by the Shannon Index, correlated positively with initial body size ($F_{1,86}$ = 13.2, p < 0.001) and was higher in males ($F_{1,86}$ = 38.4, p < 0.001), while MHC diversity did not show any effect on parasites. The parasite community did however not differ between sexes as determined by an ANOSIM (Global R = 0.002, p = 0.41). We tested whether the haplotype that

had provided a substantial advantage in the last year's enclosure experiment, was also influential in this year. In fact, we found a reduced parasitisation in females carrying this haplotype, which consists of the alleles No01 and No12 ($F_{1,48}$ = 4.78, p = 0.036, Fig III.2). This effect was not visible in males. However the possession of this haplotype did not bring an advantage against the infection of *Gyrodactylus* sp. (MWU-test, U = 941, p = 0.86), which had been the case in the previous year (Eizaguirre *et al.* Submitted-b, Annex I). Interestingly, individuals with this haplotype also showed a higher MHC genetic distance than those without it (MWU-test, U = 266, p < 0.001), but no difference in the number of alleles.

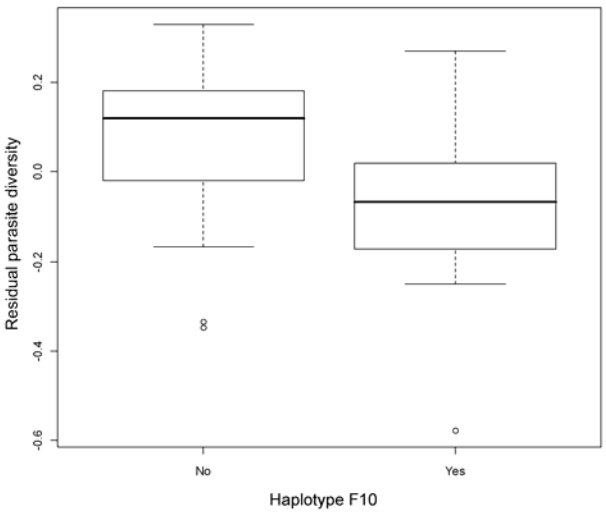


Figure III.2: Reduced parasite diversity in females without and with the haplotype F10. Parasite diversity is shown as Shannon Index, corrected for body size (Box shows median and quartiles).

Eggs and reproductive success

We collected eggs over five weeks, a period which had already been shown in previous experiments on the same population to cover the main reproductive period of three-spined sticklebacks (Kalbe *et al.* In press). A total of 152 clutches were collected. By genotyping 24 randomly chosen eggs per clutch, we could assign 3377 eggs unambiguously to their parental genotypes. 231 eggs could not be assigned or showed ambiguous parent assignment and were therefore left out of all subsequent analysis. As the 24 typed eggs do not represent the actual size of a clutch, we took photographs of the full clutch before processing and counted the eggs on the picture. With this information and the estimated proportion of fertilization for an individual clutch by parenthood analysis, we could determine the actual number of eggs laid by each female and fertilized by each male respectively. The number of eggs ranged from 25 to 986 (mean 452) for females and from 5 to 1298

(mean 402) for males and reached a total of 21,690. The variance of reproductive success (in number of eggs) in males was significantly higher than in females (Levene's test, $F_{1,102} = 17.3$, $p < 0.001$).

The overall number of total eggs sired per individual was positively correlated with the initial body size ($F_{1,86} = 18.8$, $p < 0.001$), but not with individual MHC diversity. It was also not influenced by the parasite load. When analysing only eggs sired by nest owners, i.e. eggs that were potentially the result of female choice, it appeared that in males the reproductive success was negatively correlated to the number of parasite species ($F_{1,54} = 4.7$, $p = 0.037$).

The number of sneaked eggs, i.e. sired by a male that was not the nest owner, showed an interaction effect between sex and body size ($F_{1,102} = 5.8$, $p = 0.018$), indicating that larger females were more likely to be sneaked and smaller males were more likely to sneak.

Mate choice

The design of the experiment was such that most females were able to choose from males with identical, half-dissimilar or fully dissimilar genotype. We first analysed the observed matings within each enclosure for patterns of non-random mate choice. When testing for combined heterozygosity, we found deviation from random in five of six enclosures (all 5 Chisquare, $X^2 > 11$, $p < 0.004$). In a test for mating with respect to the combined number of alleles against expectation under random mating, all six enclosures showed deviation from the expectation (all Chisquare, $X^2 > 97$, $p < 0.001$). This congruent observation over six enclosures supports the expectation that mate choice is not random with respect to MHC allele number (Binomial test, $p = 0.031$) and supports the result from a preceding study (Eizaguirre *et al.* Submitted-b, Annex I). To understand the pattern of MHC based mate choice, we then tested the overall observed mean against a random distribution. Here we did not find any deviation from random mating neither in the means nor in the variances. However, when we used only the subset of pairs with females that had the actual choice between MHC-identical, -intermediate and -dissimilar mates within their enclosure (31 instead of 48 females), the

variance in the observed genetic distances was outside the 95% confidence intervals and therefore significantly lower then expected by random mating (Fig III.3).

To test for good genes, we assigned the males into three groups, those that had not been chosen at all (n = 15), those that had been chosen less than 10 times (n = 20) and those that had been chosen ten or more times (n = 19). We compared the haplotype assembly between the not chosen and the highly chosen males in an ANOSIM, but did not find a significant pattern. Instead we found a difference in the parasite community between these two groups (ANOSIM, Global R = 0.132, p = 0.025), which was mostly attributed to the actively infecting trematode parasites *Echinochasmus* sp., *Cyathocotyle prussica* and *Gyrodactylus* sp., all of which showed higher infection intensity in the highly chosen males. The same pattern could be observed in the lower overall number of parasites that the highly chosen males harboured in comparison with the not-chosen males (MWU-test, U = 145, p = 0.013).

Based on microsatellite typing, the observed mating pairs were not more or less related to each other than expected under random mating (all I_{xy} not significantly different from 0).

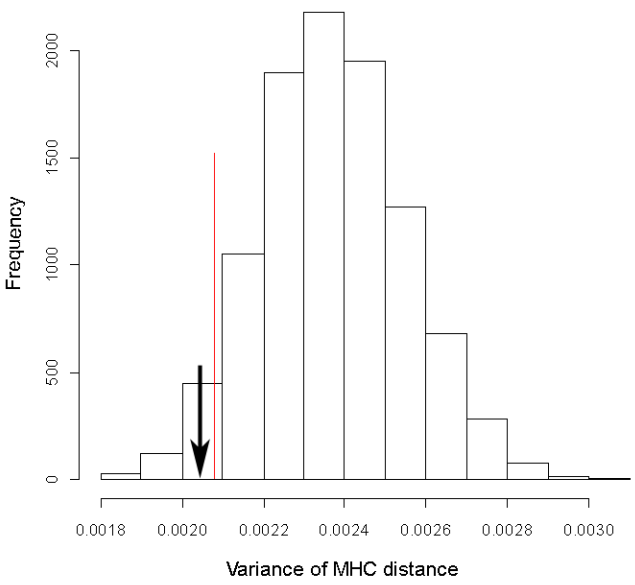


Figure III.3: Variance of MHC amino acid distances between mates in observed matings in comparison to a variance distribution under random mating. The observed variance and the random distribution are based on the set of females, which had the free choice between MHC identical, intermediate and dissimilar males. The red line indicates the one-sided 95% confidence interval and the arrow the observed variance (0.00203).

Immunological parameters

We measured the total number of lymphocytes and granulocytes in the head-kidney. The granulocyte to lymphotype ratio (G/L ratio) in combination with the total number of cells gives an idea about activation of innate and adaptive immune system respectively. The total cell counts were corrected for body size to account for differences in head-kidney size. Interestingly, the total number of granulocytes was positively correlated with the average MHC genetic distance within an individual ($F_{1,86}$ = 5.2, p = 0.026). This was also reflected in a strong positive trend between genetic distance and G/L ratio ($F_{1,86}$ = 3.9, p = 0.051). In parallel, the number of granulocytes was positively correlated with the number of parasite species within an individual ($F_{1,86}$ = 10.3, p = 0.002). In contrast, the number of lymphocytes, which stand for the adaptive immune system, was negatively correlated with parasite diversity ($F_{1,86}$ = 9.5, p = 0.003).

Discussion

Here we aimed at investigating the mechanisms that underlie MHC-mediated mate choice under semi-natural conditions. We assigned adult fish, selected for particular MHC class IIB genotypes, to six outdoor enclosures in a design which allowed most females to choose mates with identical, semi-dissimilar and fully dissimilar MHC genotypes.

This experiment is a follow up of two previous enclosure experiments, which showed an influence of MHC IIB allele diversity (Kalbe *et al.* In press) and of an MHC IIB haplotype (Eizaguirre *et al.* Submitted-b, Annex I) on reproductive success as well as mate choice for intermediate combined MHC allele diversity between mates. Here we focussed on fish with a limited range of common haplotypes and investigated their effect on mate choice and evolutionary relevant parameters with respect to reproductive success. In addition we took immunological parameters into account to better understand the mechanisms through which MHC mediates parasite resistance.

The effect of MHC diversity on initial body size, i.e. when the fish were brought out into the enclosures, shows that this immunogenetic diversity was important already at an early stage during life. Combining the facts that this holds true only for females which are also larger could indicate that females

and males follow different life history strategies. For females it is obviously important to grow fast, as this increases the final capability of laying larger clutches and by that produce more offspring. To grow, they need food and feeding influences parasitisation, for instance via feeding on intermediate hosts (Lozano 1991). Therefore, females might be more challenged during early life and require a more general immune system with a broad recognition of many different pathogens. Males could reduce food intake to a minimum to avoid the intake of parasites and would therefore not be in need for a high MHC diversity, however reproductive success is linked to body size also in males and they should therefore equally thrive to grow faster. This is still an unresolved paradox, which awaits further investigation. Unfortunately we don't have parasite data from before the reproductive period and can therefore not support the proposed male strategy by showing a lower parasite load in males. In contrast, when we measured parasite infection, i.e. after the reproductive period, we found higher parasite diversity in males and no difference in the parasite community between sexes. However, there was obviously a strong influence of parasites on the growth during this reproductive period and while females grew substantially, male size nearly stagnated. This might indicate costs that are connected to defending a territory and maintaining a nest (Guderley & Guevara 1998). We found more support for this cost in the difference in parasite communities between the highly chosen and the not chosen males. Chosen males were necessarily nest owner at the time when they were chosen and the increased intensity of actively infecting parasites might reflect the more active behaviour but also again the costs that come with maintaining a nest.

We found that smaller males were more likely to be sneakers. This might be due to the fact that they were too small to defend a territory and build a nest. Sneaking behaviour is known as an alternative strategy to enable at least fertilization of a minimum number of eggs (Wootton 1984; Blais *et al.* 2004). In contrast, larger females were more likely to be sneaked, which might be simply due to an increased number of eggs per laid clutch. The fertilization of a large number of eggs might take more time, because the sperm have to travel longer to reach all eggs. Therefore the probability of sneakers fertilizing

some of the eggs might increase. Another possibility would be that sneaker males also exert mate choice for larger females.

The higher variance of reproductive success in males indicates that sexual selection, mediated by female mate choice, was acting during our experimental period. Accordingly we found a strong sign for non-random mate choice with respect to MHC diversity. This follows a growing body of literature, which shows MHC dependent mate choice in several taxa (e.g. Potts *et al.* 1991; Eklund 1997; Reusch *et al.* 2001a; Ekblom *et al.* 2004; Milinski *et al.* 2005; Bonneaud *et al.* 2006a; Forsberg *et al.* 2007; Neff *et al.* 2008; Schwensow *et al.* 2008a, b; Eizaguirre *et al.* Submitted-b, Annex I). We could, however, not detect any haplotype specific mating patterns as tested with the ANOSIM on haplotype composition of chosen and not chosen males. Mate choice for '*good genes*', as we found earlier (Eizaguirre *et al.* Submitted-b, Annex I) could therefore not be detected in the present data. This might be due to a lower parasite pressure in 2007. We found, however, a good indication for mate choice aiming for an intermediate genetic distance between the MHC genotypes of both mates. As the variance in this parameter was significantly lower in the observed matings than in the randomly assigned pairs, we conclude that pairs with intermediate genetic distance mated more often than expected. This pattern of mate choice can be regarded as aiming for '*compatible genes*', since it involves mate evaluation based on the own MHC repertoire.

To understand these mate choice results, which in the tendency for intermediate diversity support previous findings (Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Eizaguirre *et al.* Submitted-b, Annex I), we analysed immunological parameters, which provide a more direct view on the state of the immune system. We found that the number of granulocytes, which represent an important component of the innate immune system in fish (Whyte *et al.* 1989), were positively correlated with parasite diversity. A similar observation was already made in an earlier experiment, where fish from one habitat were exposed to the more diverse parasite fauna of another habitat and showed an increase in granulocytes (Scharsack *et al.* 2007). This indicates a general activation of the innate immune system by more diverse parasite infection. The positive correlation of granulocytes with average

genetic distance could be interpreted in the context of increased T-cell depletion with high MHC diversity, where the reduced potential of the depleted T-cell repertoire is partly compensated by an increased activation of the innate immune system. While we did not find a direct link between intermediate MHC diversity and parasitisation, earlier studies support this interpretation (Wegner *et al.* 2003a; Wegner *et al.* 2003b).

Among the haploypes that we included in our system was the one that had been shown to provide an advantage against infection of the monogenean *Gyrodactylus* sp. in the previous year (Eizaguirre *et al.* Submitted-b, Annex I). This actively infecting parasite showed extremely high infection intensities in 2006 (from 1 to 244 with a mean of 13 individual parasites), which might have driven the increased mate preference by females for males that carried this haplotype. This strategy would result in an increased frequency of this haplotype in the next generation, which is in fact the generation that was sampled for our experiment in 2007. In deed, this haplotype was the second most common in the population as estimated from the initial screen of 288 randomly collected fish. Interestingly and as expected under '*negative frequency-dependent seletion*' (Takahata & Nei 1990) this year the infection intensity by *Gyrodyctylus* sp. was strongly reduced with 0 to 20 and on average 2 parasites per fish. Again as expected, we did not find a resistance effect of this haplotype on *Gyrodactylus* intensity. The selection for hosts with this haplotype in 2006 decreased the host range for all *Gyrodactylus* with these specific antigens and probably led to a reduction in frequency of these *Gyrodactylus* genotypes. This would explain both, the high frequency of this MHC haplotype in the 2007 generation, and also the fact that we did not detect any advantage of this haplotype against *Gyrodactylus* in 2007. This process of *'negative frequency-dependent selection*' in a host-parasite system has been proposed as one of the major mechanisms to maintain the high polymorphism in the MHC (Bodmer 1972; Takahata & Nei 1990). Surprisingly, we found a higher overall resistance against parasites in female sticklebacks with this haplotype. In comparison to the other haplotypes, is does however not show any particular pattern on the sequence level. Potentially, other parasite species had been slower in the adaptation to this haplotype than *Gyrodactylus* and were now overtaken in the arms race by

the high frequency of this haplotype caused by the strong selection of *Gyrodactylus*.

Besides this, we found evidence for heterozygote advantage in the positive correlation between MHC IIB diversity and body size, which is a trait that directly influences the reproductive success. This hypothesis assumes an advantage for individuals carrying more different MHC variants (Doherty & Zinkernagel 1975b) and is also expected to contribute to the high polymorphism in this system (Apanius *et al.* 1997). However, this mechanism is still not resolved (Penn *et al.* 2002; Wegner *et al.* 2003a; De Boer *et al.* 2004; Wedekind *et al.* 2005; Ilmonen *et al.* 2007; Woelfing *et al.* 2009) and needs a more thorough investigation on the sequence level (Lipsitch *et al.* 2003). We therefore also determined the average genetic distance between alleles of each individual, which might give a more precise estimate of the distinctiveness of the individual alleles and of the range of different antigens they are able to bind. However, we did not find correlations between this genetic distance and body size or parasite infection.

Overall, we added evidence to the growing body of studies that found mate choice for intermediate MHC diversity, i.e. '*compatible genes*', but found only indirect evidence for the occurrence of '*good genes*' in MHC by comparing our data with an earlier study. This haplotype, which had in 2006 been identified as '*good gene*', still led to decreased parasitisation in females, but it did not influence mate choice decisions in 2007.

Forsberg *et al.* (2007), who found mate choice for intermediate MHC diversity in brown trout, tried to explain this by referring to the phenomenon of outbreeding depression. There, females would avoid too dissimilar mates, as they might not possess locally adapted alleles. A more common explanation is depletion of the T-cell receptor repertoire by too diverse MHC genotypes. This has been modelled theoretically (Nowak *et al.* 1992; Woelfing *et al.* 2009; but see also Borghans *et al.* 2003), whereas direct evidence is still awaited. While we agree with the idea of outbreeding avoidance in MHC based mate choice under special circumstances (Eizaguirre *et al.* 2009), it seems that females in our experiment were not running the risk of mating with a locally not adapted partner, as we had provided them only with a restricted range of haplotypes, which were all common in the local population. With the observed increase of

granulocytes, our data gives a hint for an activation of the innate immune system with increased MHC diversity. In this light, the avoidance of a depleted T-cell repertoire with a mate choice strategy for intermediate diversity might appear more likely. Nevertheless, more experimental work is needed to rule out one or the other explanation.

Acknowledgements

We thank Ines Schultz for invaluable help in the lab and in the field. We are also indebted to V. Buescher, G. Augustin, D. Martens, H. Deiwick and R. Derner for technical assistance. We thank the county of Schleswig-Holstein and the city of Ploen for permitting us to run the outdoor experiments in the lake. All animal experiments described were approved by the Ministry of Nature, Environment and Country Development, Schleswig-Holstein, Germany.

Chapter IV

MHC local adaptation: a field experiment

Christophe Eizaguirre*, Tobias L. Lenz*, Chris Harrod, Martin Kalbe & Manfred Milinski

* CE and TLL contributed equally to this study

Abstract

The polymorphic genes of the major histocompatibility complex (MHC) have been proposed to enable individuals to adapt and therefore survive and reproduce in their local pathogenic environment. However, the presence of local MHC adaptation has so far not been experimentally tested. For the present study, we crossed three-spined sticklebacks from a river and a lake habitat displaying almost totally distinct MHC allele pools. Intercrossing F1 hybrids, we obtained a second fish generation segregating into pure river, pure lake and two hybrid MHC genotypes, while at the same time randomized for the genetic background. We used a double common garden experiment where fish were exposed during 10 months simultaneously to lake and to river conditions using replicated cages. We found that MHC genotype was correlated to parasite load in hybrid lines. The total parasite load harboured by the fish appears to result from an adaptive interaction between the innate and the acquired immune system of the hosts, in turn driven by different parasite strategies. Generalist parasites were more successful in river fish lines, which naturally encounter low pressure from these parasites. However, fish bearing lake MHC genotypes were more infected than fish carrying river MHC genotypes, when confronted with a stickleback-specific parasite, which is commonly present in river. In addition, we found support for additive and synergistic role of the genetic background and innate immunity in parasite resistance. In conclusion, we unravel experimentally the effects of two locally adapted genomes in general and MHC allele pools in particular and provide evidence for their interaction with contrasting parasite strategies.

Introduction

The major histocompatibility complex (MHC) represents a gene dense genomic region and displays some of the most polymorphic loci in the vertebrate genome (Klein & Figueroa 1986; Apanius *et al.* 1997). It consists of genes coding for cell surface molecules that present self and non-self peptides to T-cells and therefore play a key role in the onset and modulation of the specific immune response (Klein 1986; Janeway *et al.* 2001). Due to their tight linkage with the immune system, the evolution of MHC genes is thought to be under parasite driven selection (Klein 1986; Potts *et al.* 1994; Apanius *et al.* 1997). MHC genes generally present a high intra- and inter-individual polymorphism. Such polymorphism is thought to enable individuals to fight against the spectrum of parasites they are confronted with in their local environment (Apanius *et al.* 1997; Bernatchez & Landry 2003; Piertney & Oliver 2005).

In a local environment, ecological conditions, i.e. biotic and abiotic factors, but also community structure are likely to influence the presence and abundance of certain intermediate hosts, which are in turn necessary for the maintenance of different parasite taxa (Thompson 1994). Therefore, contrasting environments are likely to support distinct parasite communities, which would then require contrasting immunogenetic adaptation of hosts as well. This is supported by recent studies investigating MHC variation in neighbouring populations, which showed that habitat heterogeneity can shape MHC composition, probably due to contrasting parasite communities (Wegner *et al.* 2003b; Bonneaud *et al.* 2006b; Blais *et al.* 2007; Ekblom *et al.* 2007; Alcaide *et al.* 2008; Babik *et al.* 2008). However, often the parasite side of these studies has not been revealed.

It has been proposed that parasitic exploitation could have the potential to promote speciation by low hybrid fitness relative to specifically co-evolved local host genotypes (Haldane 1992; Turelli *et al.* 2001). Although a variety of theoretical models have provided support for this idea (e.g. Maynard Smith 1966; Rice 1987) it remains controversial. The '*Red Queen hypothesis*' postulates that interactions among species, e.g. hosts and parasites, lead to constant natural selection for adaptation and counter-adaptation (Van Valen 1973). Usually, in a host-parasite arms race, parasites are expected to have

an advantage over the host and to trade off virulence and host mortality (Ebert 1994). In addition, studies have reported that local adaptations resulted from parasites tracking locally common host genotypes and thus meeting a necessary condition of the Red Queen hypothesis (Ebert 1994; Lively & Dybdahl 2000). Local adaptation is therefore defined as higher parasite fitness on a sympatric than on an allopatric host (Ebert 1994). Interestingly, a model (Kawecki 1998), a meta-analysis (Lajeunesse & Forbes 2002) and an experimental test (Kalbe & Kurtz 2006) point towards a more complex theory involving different parasite strategies. If parasites and hosts are engaged in a co-evolutionary process, then specialist parasite lineages will have an advantage over generalist parasites on a specific host genotype (Lajeunesse & Forbes 2002). As specialist parasites need to counteract evolution of host resistance, because they cannot switch to another host, they have to evolve faster in response to the host reaction, compared to generalist parasites. Such arms races should result in lower parasite burden from generalist parasites and higher infection from specialist parasites in local host genotypes. Since individual hosts are generally exposed to a broad variety of pathogens, an optimal host response that counteracts both threats should be expected.

In addition, it was recently proposed that the polymorphic genes of the major histocompatibility (MHC) could act as pleiotropic genes adapting individuals to their local pathogenic environment and leading to assortative mating, and therefore could facilitate speciation (Blais *et al.* 2007; Eizaguirre *et al.* 2009). However, correlative evidence is scarce and thus the scientific community awaits empirical tests. In the present study, we tested the specific role of MHC genes in host adaptation to local parasite fauna – a pre-requisite for MHC pleiotropy. For this purpose, the three-spined stickleback (*Gasterosteus aculeatus*) represents an exceptional study organism. In Northern Germany, populations colonized lakes and rivers after the last glaciation (Reusch *et al.* 2001b). Interestingly, these populations experience practically no gene flow (Reusch *et al.* 2001b), are confronted with contrasting parasite communities (Kalbe *et al.* 2002; Wegner *et al.* 2003b) and display low overlapping MHC allele pools (Eizaguirre *et al.* Submitted-a). These characteristics led us to a specific breeding design, combining the effect of

segregation and genetic mechanisms to produce hybrid lines with randomized genetic background. This in turn enabled us to focus on the MHC genotype effect without confounding background (Rauch *et al.* 2006b). Using F_2 fish, we performed a double common garden experiment. We simultaneously and reciprocally exposed these hybrid and pure line F_2 fish to the lake and the river pathogenic environments, where their parental generation had been caught. Here, we aim at experimentally showing that host MHC genes are adapted to the sympatric local parasites, which would explain the observed olfactory based assortative mating within these particular two populations (Eizaguirre *et al.* Submitted-a) by maintaining local MHC adaptation.

Material & Methods

In this section we will describe the methods used for the full experiment with a large integrative approach. However, for the results we will keep a major focus on parasite load and MHC.

Breeding design

We sampled three-spined sticklebacks (*Gasterosteus aculeatus*) from two habitats, a small, slow-flowing river (Malenter Au, 54°12'16.19"N, 10°33'32.93"E, Germany) and a lake (Großer Plöner See, 54° 9'21.16"N, 10°25'50.14"E, Germany) in spring 2006. Fish were kept singly in 16L tanks with constantly flowing water. In June 2006, the fish were transferred to summer condition (day light 14h, T 18°C) in order to reach sexual maturity. Males were provided with artificial nesting material (sand and nylon threads). We crossed fish randomly with respect to fish identity, but not to fish origin, in order to obtain lab-bred sibships with pure lake, pure river and hybrid genomic background (Fig IV.1). For breeding, males and females were allowed to display natural behaviours and females to spawn into the male's nest. Clutches were left for one hour in the nest after fertilization before being taken them out. Individual egg clutches were incubated in aerated well water (0.04 ppm Malachite Green against fungal infections) at 18°C until hatching. We obtained 40 individual clutches with the three following backgrounds: pure lake (10 crosses between lake fish), pure river (10 crosses between river fish) and hybrids (20 crosses between a lake and a river fish). As the same

parental fish were used in both pure and hybrid crosses, to assure the same genetic material in the different lines, 11 of the hybrids had a mother of lake origin and 9 had river mothers. The F1 generation was progressively transferred to autumn, winter and spring conditions, from when on the fish were kept singly. Finally they were brought to summer condition where the males were provided with nesting material. Following the same protocol, we randomly interbred within hybrid lines and within pure lines in order to get the second fish generation (F2). From the initial 40 F_1 crosses, we therefore obtained 20 F_2 families with three different genomic backgrounds: 5 river, 5 lake, 10 hybrid. With this design, we aimed at producing hybrid families with randomized genetic background, which could then be analyzed for origin specific MHC genotype effect.

In each F2 hybrid family, four different MHC genotypes can be expected: RR, LL, RL and LR. The first letter stands for the haplotype inherited from the F1 father, the second letter for the haplotype inherited from the F1 mother. L and R represent the lake or river origin of the haplotype respectively. Original lake and river population diverging strongly in their MHC allele pools, this design should result in F2 offspring segregating within each family into four distinct MHC genotypes with a shared randomized genomic background. Using F2 fish eliminates as well any potential maternal effects.

Field exposure

On September 3rd, 2007, fish were brought to the natal place of the parental generation (the river Malenter Au, 54°12'16.19"N, 10°33'32.93"E and the lake Großer Plöner See, 54° 9'21.16"N, 10°25'50.14"E, Germany, Pic IV.1). We stocked 38 cages with the 10 hybrid families. Each cage contained 20 fish, partitioned as follows: 10 random fish from one hybrid family, plus one fish of each pure family (5 river and 5 lake). For the hybrid families we created four replicates (except for one family where we achieved only two replicates). Two (one) cages per family were located in the lake and the two (one) others were placed into the river. In February, before the fish reached sexual maturity, the sexes were separated with a fine mesh net to avoid potential breeding. The fish stayed in the cages for 10 months from September to June and collection started on the 5.06.2008. In the middle of this period a cage

Picture IV.1: Pictures showing the habitats in which the cages were located. Left: picture of the slow flowing river, the Malenter Au. Right: picture of the lake Großer Plöner See.

was destroyed in the river after a thunderstorm. Once a week, all cages were checked for dead fish and the water temperature, oxygen content and conductivity were measured.

The cages were long and flat (1m x 0.25m x 0.6m) in order to assure constant immersion even in the shallow river. They consisted of stainless steel mesh and a stable framework. The 5 mm mesh allowed only small particles and most invertebrates (food items and intermediate hosts of various parasite species) to pass through. We placed 19 cages at an interval of 3 meters in the centre of the river bed (about 3 m wide and 1.5m to 0.5 m deep). The other half of the replicates was placed in a shallow part of the lake at about 1-1.2 m depth at 2 m intervals.

MHC typing

The MHC class IIB diversity was determined by reference strand conformation analysis (RSCA) according to Lenz *et al.* (Chapter II). We amplified the exon 2 of MHC class IIB genes that encodes for the peptide-binding groove. This exon has previously been suggested, to be under parasite mediated positive selection as revealed by a high ratio of non-synonymous over synonymous substitutions (d_N/d_S, Reusch *et al.* 2004). The MHC loci in the stickleback are duplicated and the number of sequence variants varies substantially between individuals (Wegner *et al.* 2003a; Reusch & Langefors 2005; Lenz *et al.* Chapter II). For simplicity we will refer

to the different sequence variants as alleles, although they may stem from different loci.

Microsatellite typing

For each F2 fish, DNA extractions from dorsal spine were performed using the Invisorb DNA Tissue HTS 96 Kit / R (Invitek, Germany) on an automated platform (Tecan, Switzerland) following the manufacturer's protocol. All fish were typed for 9 microsatellites combined in 2 different PCR multiplex protocols before and after the experiment (Reusch *et al.* 2001a; Kalbe *et al.* In press). We used these loci to identify fish at the end of the experiment.

Parasite load

Since the number of fish was too high to be handled within three days as aimed for in previous studies (Kalbe *et al.* 2002; Kalbe *et al.* In press), dissection took place with regard to cage replicates (80 fish) to keep all MHC genotype comparisons possible. Since cages were half stocked with hybrids and half with pure lines, pure and hybrids lines were directly comparable. The dissections were spread over four weeks and the total number of fish recovered was 714. All external and internal macroparasites and ciliates were determined to the lowest taxonomic level possible using dissection microscopes (for dissection protocol see Kalbe *et al.* 2002).

In order to quantify the total parasite load of each fish, the Shannon index, which accounts for both abundance and evenness of the present species, was calculated following a fourth root transformation of parasite abundance. This allowed us to summarize different combinations of parasites for each stickleback and to make a quantitative comparison of their total parasite burden. We calculated a second Shannon index taking only generalist parasites into account. For the specialist parasites we focussed on *Gyrodactylus sp.*, since it was the only one with a sufficient abundance (more than 1 %). During dissection, fish mass (± 0.1 mg), standard length (± 1 mm) and lateral plate number were recorded.

Immune assays

For immunological assays, leucocytes were isolated from head kidney (HKL). Head kidney preparations followed Scharsack *et al.*'s protocol (2004). Total cell numbers in HKL isolates were determined with the standard cell dilution assay. Additionally, flow cytometric measurements of freshly isolated HKL were used to determine proportions of granulocytes (FSC/SSC-high) and lymphocytes (FSC/SSC-low) in individual HKL samples (all according to Scharsack *et al.* 2007).

As an important factor of the cell mediated innate immunity, the respiratory burst activity of HKL was quantified (Kurtz *et al.* 2004). As a parameter for activation of the adaptive immune system, frequencies of lymphocytes in G0-1, S and G2-M phases were acquired by DNA content analysis of red fluorescence intensity (propidium iodide labelling) of single cells from the lymphocyte gate (Scharsack *et al.* 2007).

MHC expression

As gene expression is a good diagnostic tool of the current physiological state of an individual and reflects the need for the final product encoded by the target gene, we performed quantitative real-time polymerase chain reaction (qRT-PCR) to follow MHC gene expression in the fish.

RNA was extracted from spleens, known as an immuno-relevant organ. The spleen is a major secondary lymphoid organ of teleost fish which is mainly involved in blood filtration, antigen trapping and processing. Immediately after killing and opening the fish, the spleens were stored in RNA later and kept at -20°C until RNA extraction. The RNA was extracted with the NucleoSpin 96 RNA kit (Macherey Nagel, Düren, Germany) following the manufacturer's protocol. Organs were homogenized in 1% mercapto-ethanol and buffer in a Retsch mill by shaking samples 3 minutes at 30Hz. Elution was performed in two steps in a final volume of 100 µl. Reverse transcription was carried out with Omniscript Reverse Transcriptase kit (QIAGEN). As gene expression is a relative value, the L13A gene coding for the ribosomal binding protein was used as house keeping gene following Hibbeler *et al.*'s (2008) advices. Housekeeping genes serve as internal control for reverse transcription and allow for relative quantification of the target gene. qRT-PCR

was performed following Wegner *et al.*'s (2006) protocol developed for MHC class II genes in the three-spined stickleback. For each individual, three qRT-PCR runs were performed and values were averaged to counteract potential run errors. We calculated average ΔCT values by subtracting mean CT of the target gene to the mean CT value of the internal control for each individual.

Stable isotopes

It has been proposed that fish, originating from the same populations as the ones used in the current study, could be using habitat resources in a different manner (foraging behaviour, physiological processing of the food) and habitat use could be linked to MHC genotype and parasitisation patterns (Eizaguirre *et al.* Submitted-a). Using a common garden experiment as the one employed here, allows for directly comparing food use by the fish with different genetic backgrounds (pure lake, pure river, and hybrids) and MHC genotypes. Previous studies have used stable isotope analysis (SIA) to reveal individual migration between ecologically contrasting but connected habitats in aquatic consumer species as diverse as European eels (Harrod *et al.* 2005) and Galápagos sea lions (Wolf *et al.* 2008). SIA have been used as well to report niche use and are an exceptional tool to describe the trophic level of an organism. During dissection, a section of the caudal muscle was dissected, and oven dried for 24 hours at 65 °C. Samples were grinded, weighed in tin cups and analyzed for carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) stable isotope ratios and elemental C and N composition (for details see Harrod *et al.* 2005). Repeated analyses of internal standards showed precision of < 0.1 ‰ (carbon) and < 0.3 ‰ (nitrogen). $\delta^{13}C$ data were arithmetically lipid-normalized following Kiljunen *et al.* (2006).

Statistical analysis

Difference in parasite communities between habitats was tested with an ANOSIM (Analysis of similarity) based on a Bray-Curtis similarity matrix.

We estimated the importance of MHC genotypes in parasite resistance in the hybrid lines by nesting the MHC genotypes within the family effect in an analysis of variance (ANOVA). The dependent variable represented the residuals of a two-way ANOVA between parasite load, habitat and sex to

control for both effects. For all the tests we controlled for family effects treating it as random factor. Such analysis was not only conducted on the total parasite load but as well on the load of a specialist stickleback parasite, the trematode *Gyrodactylus sp.* and on a Shannon parasite index including only the generalist parasites (Table IV.2).

To test for local adaptation in the hybrid families, we determined the interaction between origin of MHC (LL or RR) and the habitat. We tested as well a general local adaptation by estimating the interaction between breeding origin (pure lake, pure river and hybrid) and the habitat of exposure. For these tests, we used a split plot design and controlled for family effect.

Statistical analyses were conducted using R statistical package (v. 2.5.0) for Windows XP and Primer v6 (Clarke & Gorley 2006). Normality of distributions and variance homoscedasticity were verified and tests conducted accordingly. If necessary, dependent variables were log or square root transformed to meet test assumptions.

Results

MHC genotypes

In total we detected 28 different MHC alleles. We did not observe any event of recombination. In all hybrid lines but one, we could clearly determine the population origin of the different MHC genotypes. This one family was excluded from all analyses due to origin-overlapping alleles in the parental genotypes. The following results were thus obtained from nine hybrid families, five pure lake and five pure river families. In total, we detected 19 alleles in the lake pure lines, 15 in the pure river and 22 in the hybrid lines (Fig IV.1). Descriptive information of MHC diversity is given in table IV.1. One F2 hybrid family showed MHC homozygous individuals due to the presence of the same haplotype (allele No05) in both F1 crosses.

In general, homozygous F2 fish displayed a variable number of MHC variants, ranging from one to three. This suggests either that in some haplotypes the same variant is present on different loci or that individuals display a variable number of MHC loci. This variation between haplotypes could be followed and was consistent over the three fish generation (Parents, F1, F2).

Table VI.1: Description of MHC diversity in fish exposed to natural lake and river conditions.

a) All fish				
Background	Number of individuals	Median MHC diversity	Range of MHC diversity	MHC Heterozygosity
Pure Lake	176	4	2-5	0.97
Hybrids	349	4	1-5	0.898
Pure River	173	4	1-5	0.591
b) Hybrid fish				
MHC genotype	Number of individuals	Median MHC diversity	Range of MHC diversity	MHC Heterozygosity
LL	50	4	3-4	1.00
LR	53	5	1-5	0.848
RL	96	3	3-5	1.00
RR	99	4	1-5	0.68

Parasite load

- *Comparison between habitats*

We recorded a total of 24 different parasite species (Table IV.2). 22 were found in fish exposed to the lake, whereas 14 were counted in fish exposed to the river. Among these parasite species, 10 were lake-specific, whilst 2 were found only in fish exposed to river conditions (Fig IV.3). The detected parasites differed in their infection and transmission strategies (Table IV.2).

Parasite communities found on the fish differed between both habitat of exposure (ANOSIM, $R=0.625$, $p<0.001$). A similarity percentage analysis of species contributions revealed that the average dissimilarity between parasites communities was 50.6%. Such habitat differences remained when considering only actively infecting parasites (ANOSIM, $R=0.528$, $p<0.001$). The difference was mainly driven by the intensity of the specialist *Gyrodactylus sp.* in river and the generalists *Apatemon cobitis* and *Diplostomum sp.* in the lake. This suggests that the river parasite community is driven by specialist parasites, whereas generalist parasites might dominate the lake parasite community. Trophically transmitted parasite communities were as well different between habitats (ANOSIM, $R=0.087$, $p=0.001$), and the

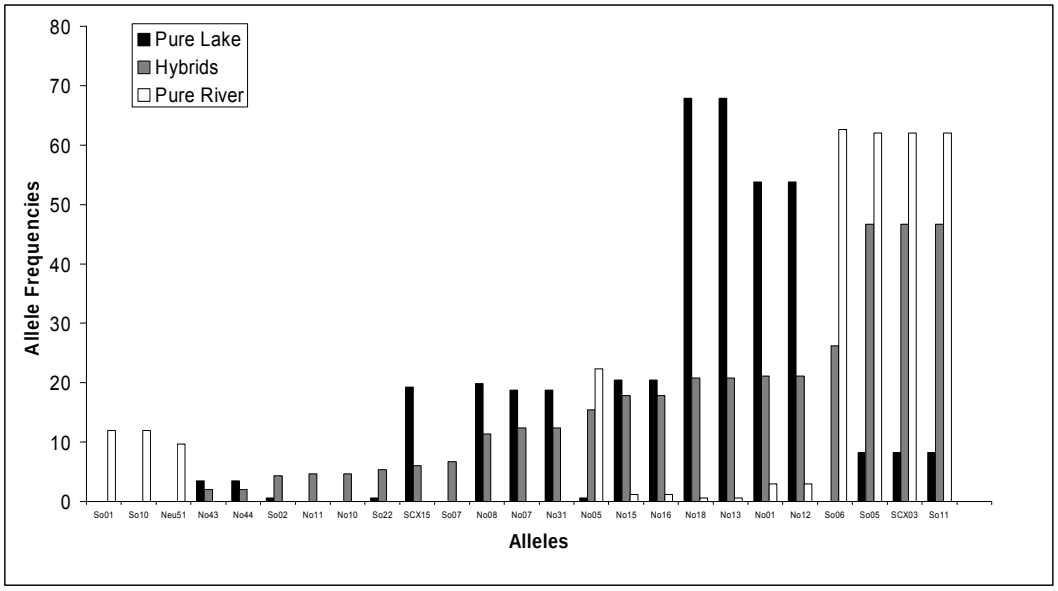


Figure IV.1: Graph presenting MHC IIB allele frequencies obtained in the F2 families with regard to breeding origin.

variation was mainly due to a higher abundance of *Raphidascaris acus* in river and *Camallanus lacustris* in lake.

- *MHC genotype*

Using the F2 hybrids for which we could determine the origin of the MHC genotype, we tested whether the parasite load was correlated with these genotypes and with the family once corrected for habitat and sex differences. We found a strong family effect ($F_{8,263}$=6.048, p<0.001. Fig IV.2a) and a significant MHC genotype effect ($F_{26,263}$=1.520, p=0.050. Fig IV.2b).

Interestingly, this effect of the MHC genotype remained when considering a specialist parasite like *Gyrodactylus sp.* ($F_{26,263}$=1.561, p=0.044), but not when focusing on the index of generalist parasites ($F_{26,263}$=0.646, p=0.909). Consistent with the results obtained with the general parasite load index, in both cases the family effect was pronounced (p<0.001)

Table VI.2: Table summarizing the prevalence (percentage of infected hosts), the mean intensity (number of parasite individuals of a given species per infected host), and the maximum number of parasite of all recorded parasites in the 714 dissected fish.

Parasite species	Parasite strategy	Trans-mission	Lake			River		
			Mean	Max.	Prevalence	Mean	Max.	Prevalence
Diplostomum sp.	Generalist	Active	18.04	68	99.18	16.88	75	97.14
Apatemon cobitis	Generalist	Active	9.22	296	86.30	0.23	6	11.14
Trichodina sp.	Generalist	Active	9.09	100	83.01	34.43	100	95.43
Glochidia	Generalist	Active	5.53	92	76.71	0.06	14	1.14
Gyrodactylus sp.	Specialist	Active	12.84	248	72.33	81.06	1073	99.43
Apiosoma sp	Generalist	Active	4.64	100	58.63	33.86	100	69.14
Cyathocotyle prussica	Generalist	Active	1.20	37	44.38	4.13	21	85.71
Echinochasmus sp.	Generalist	Active	0.73	7	40.55	0.02	2	1.14
Argulus foliaceus	Generalist	Active	0.84	16	38.63	0	0	0
Camallanus lacustris	Generalist	Trophic	0.34	4	24.93	0	0	0
Trematode X	Generalist	Trophic	0.33	8	18.36	0	0	0
Tylodelphis calvata	Generalist	Active	0.37	25	15.34	0	0	0
Raphidascaris acus	Generalist	Trophic	0.16	3	13.15	0.42	4	29.14
Contracaecum sp.	Generalist	Trophic	0.09	4	7.12	0	0	0
Paradilepis scolecina	Generalist	Trophic	0.07	2	5.48	0	0	0
Anguillicoloides crassus	Generalist	Trophic	0.05	1	4.93	0.05	2	5.14
Proteocephalus filicollis	Specialist	Trophic	0.05	2	4.38	0	0	0
Phyllodistomum folium	Generalist	Trophic	0.04	2	2.74	0	0	0
Valipora campylancristrota	Generalist	Trophic	0.03	4	1.37	0	0	0
Ichthyophthirius multifiliis	Generalist	Active	0.01	1	1.10	0.02	1	1.43
Ergasilius sp.	Generalist	Active	0.01	1	0.82	0	0	0
Nematode X	Unknown	Trophic	0.01	1	0.55	0	0	0
Cestode X	Unknown	Trophic	0.00	1	0.27	0	0	0
Piscicola	Generalist	Active	0.00	1	0.27	0.01	1	0.86
Acanthocephalus lucii	Generalist	Trophic	0	0	0	0.06	2	6
Metacercaria X	Unknown	Active	0	0	0	0.01	2	0.86
X is used for parasites currently under determination.								

- *MHC local adaptation*

Expectedly, parasite load (Shannon index) was higher in the lake than in the river habitat ($F_{1,263}$=84.226, p<0.001). We found evidence for local adaptation in the hybrids as seen in the significant interaction between habitat

of exposure and origin of MHC ($F_{3,263}$=3.226, p=0.031. Fig IV.3a). Fish with an LL MHC genotype displayed a higher parasite load than RR MHC fish in lake.

The opposite was observable in fish exposed to river conditions.

When focusing on generalist parasites, we found a habitat ($F_{1,263}$=182.678, p<0.001. Fig IV.3b) as well as a family ($F_{8,263}$=3.479, p=<0.001) effect, but no correlation with MHC genotypes ($F_{3,263}$=0.958, p=0.413) and no interaction of MHC and habitat. In contrary, investigating the specialist *Gyrodactylus sp.* revealed an MHC genotype effect ($F_{3,263}$=3.668, p=0.012. Fig IV.3c): in both habitats fish displaying LL MHC genotype harboured a higher number of *Gyrodactylus sp.*

- *Background genome local adaptation*

In the test for background genome adaptation using all the fish, we found a strong significant interaction between habitat of exposure and breeding origin (pure lake, pure river and hybrids) ($F_{2,629}$=16.166, p<0.001).

Pure lake background fish displayed a higher parasite load in lake than pure river background fish, and vice versa in the river habitat where pure river lines were more infected. Hybrid fish stand at an intermediate level (Fig IV.4). We did not find a significant interaction between origin and habitat of exposure in generalist parasites ($F_{2,629}$=0.831, p=0.4363) but instead both variables alone predicted parasitisation (Habitat: $F_{1,629}$=328.364, p<0.001; Origin: $F_{2,629}$= 17.201, p<0.001). Fish exposed to the lake habitat harboured a higher parasite load than fish exposed to river conditions. In addition, fish with a pure lake genome showed a consistently lower parasite load regardless of the habitat of exposure. Interestingly, the specialist *Gyrodactylus sp.* burden showed only a habitat effect ($F_{2,629}$= 408.931, p<0.001) being much more dominant in river.

Table IV.3 summarizes all the results for the different parasites investigated and the different hypotheses tested.

Immune assays

To obtain a parameter for granulocyte activation, we measured the oxidative burst activity in a chemoluminescence assay from the head kidney leucocytes (HKL). Sticklebacks exposed to the river showed higher zymosan-induced respiratory burst compared to fish exposed to the lake ($F_{1,629}$=67.5173, p<0.001). We could also observe an origin effect with pure lake lines displaying a higher respiratory burst than river fish ($F_{1,629}$=16.703, p<0.001. Fig IV.5). When focusing only on hybrid fish and their MHC genotype, we did not detect any relation between oxidative burst and MHC genotypes ($F_{1,263}$=0.921, p=0.431). In both cases, we found a strong inter-family variation (p<0.001).Often, immune activation has been investigated on the basis of spleen weight. We checked whether the spleenosomatic index (spleen weight relative to total weight) was related to habitat, fish origin and their interaction. We found only a significant interaction (Habitat: $F_{1,629}$=0.091, p=0.763; Origin: $F_{2,629}$=0.335, p=0.716; Habitat x Origin: $F_{2,629}$=14.574, p<0.001).

Growth rate

In order to estimate the general energy each fish needs to maintain its metabolism, we compared the fish weight (without gonads) between habitats,

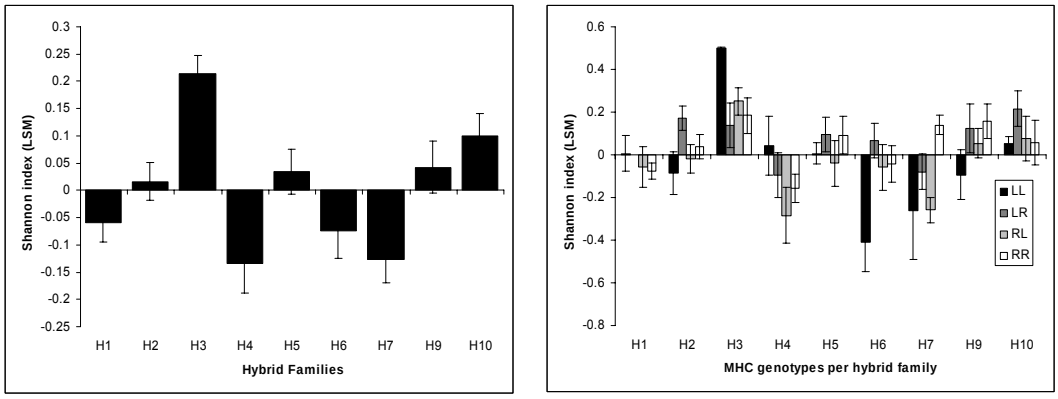
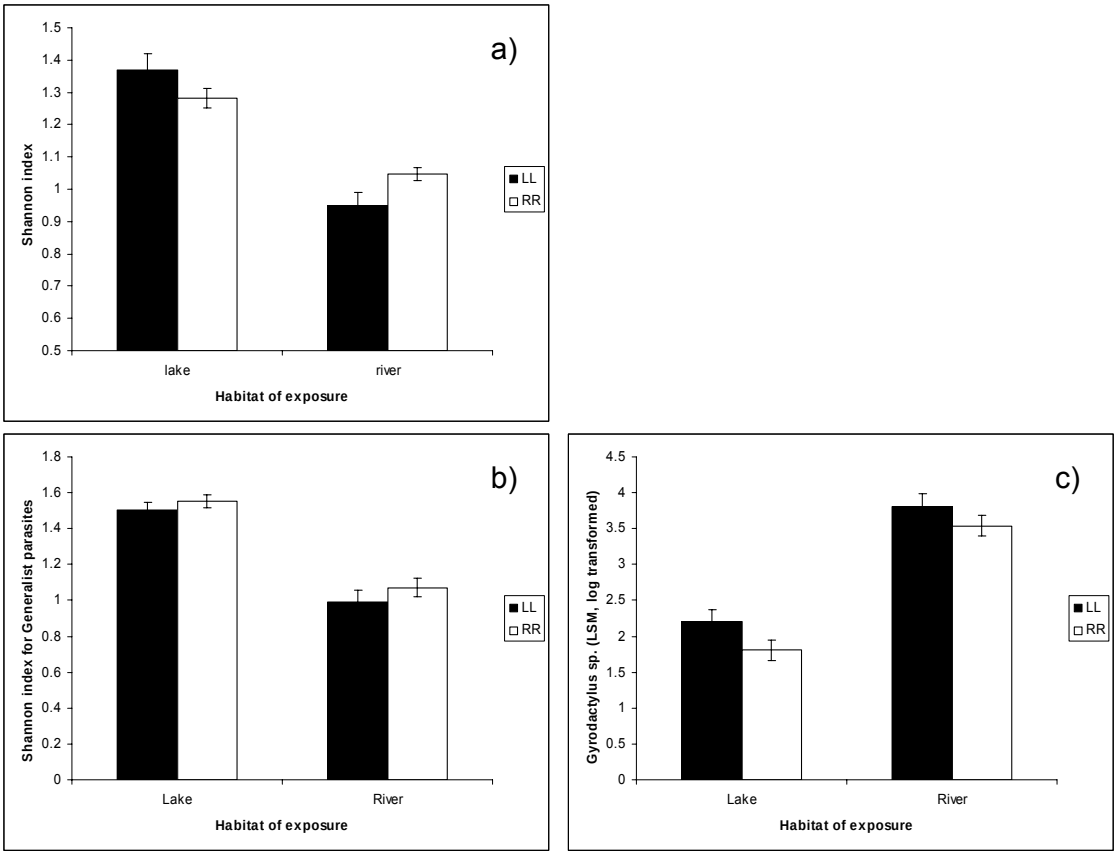


Figure IV.2: Effect of family and MHC genotype on parasite load. The left panel shows mean parasite load of stickleback in the nine hybrid families, exposed in the two habitats. The right panel shows the parasite load for the 36 MHC genotypes in the nine hybrid families. Bars show means (+/- standard errors), computed from residuals of an ANOVA including habitat effect.



MHC genotypes and breeding origin. Generally, fish exposed to the river grew better than in the lake ($F_{1,629}$=40.287, p<0.001). Fish with pure river genome grew better in both habitats than pure lake and hybrid lines ($F_{2,629}$= 23.986, p<0.001). We did not find any correlation between weight and MHC genotypes in the hybrid fish ($F_{3,263}$=0.162, p=0.851). Here as well, in both cases, we found a strong inter-family variation (p<0.001).

Discussion

It was recently proposed that MHC genes could act as magic traits by both adapting individuals to their local parasite fauna and leading assortative mating strategy to avoid disruption of adapted genotypes (Eizaguirre *et al.* 2009). Here we tested for local adaptation of MHC genes, a first pre-requisite for MHC genes acting as a magic trait. Although interest for MHC genes has increased, experimental work testing evidence for MHC adaptation under natural environments is scarce (Rauch *et al.* 2006a; Consuegra & Garcia de

Leaniz 2008). This might be due to effort of identifying the wide array of parasites a natural host population is confronted with (Thompson 1994) and to the difficulty of disentangling the role of MHC genotype from the background genome. In the present experiment we addressed the latter by using a two-lab generation breeding design, where we obtained replicated F2 fish that belonged to the same family but carried contrasting MHC genotypes. Exposing these fish to the natural habitat of their parents circumvented the limitation by the low number of parasites which can be handled in lab for experimental infections. Here we showed that in F2 hybrid fish, for which the background genome was randomized, the total parasite load was influenced by the MHC genotypes.

This clearly shows that MHC genotypes are not equal facing the parasites and that natural selection can play a crucial role in selecting best adapted MHC genotype. Although the effect was lower than the variation between families, we report the first field experimental evidence for a role of MHC genes in local adaptation to parasite resistance. The strength of the observed effect was in the expected range (Apanius *et al.* 1997). Our results

Table IV.3: Statistical details for three different parasite indices: Shannon (total parasite load), Generalists (Shannon index obtained from only generalist parasites) and Specialist (log transformed number of *Gyrodactylus sp.*). a) test for influence of the different MHC genotypes in hybrids family, b) test for local adaptation of MHC genotypes, c) test for local adaptation of fish origin.

a) Hybrids	Shannon		Generalists		Specialists	
Variables	F	P	F	P	F	P
MHC genotype	1.520	**0.050**	0.646	0.908	1.561	**0.044**
Family	6.048	**<0.001**	3.327	**0.002**	12.155	**<0.001**
b) Hybrids	Shannon		Generalists		Specialists	
Habitat	84.678	**<0.001**	182.678	**<0.001**	280.387	**<0.001**
MHC origin	1.629	0.183	0.958	0.413	4.815	**0.003**
Family	5.975	**<0.001**	3.479	**<0.001**	15.338	**<0.001**
Habitat x MHC origin	3.226	**0.031**	2.289	0.079	0.884	0.450
Family x MHC origin	1.548	0.210	0.551	0.936	1.137	0.314
Habitat x Family	1.262	0.141	0.324	0.957	10.999	**<0.001**
Habitat x Family x MHC genotype	0.926	0.540	0.8824	0.227	1.077	0.378
c) All fish	Shannon		Generalists		Specialists	
Habitat	213.862	**<0.001**	328.364	**<0.001**	408.931	**<0.001**
Origin	2.333	0.098	17.201	**<0.001**	1.372	0.254
Habitat x Origin	16.137	**<0.001**	0.831	0.436	0.718	0.488

contrast with a previous attempt of determining the role of MHC genes in parasite resistance where no effect was found (Rauch *et al.* 2006b). The discrepancy could stem from the longer exposure time in our study. Indeed, timing of exposure appears to be important for the pathways included in the adaptive immune system due to the production of antibodies (Janeway *et al.* 2001).

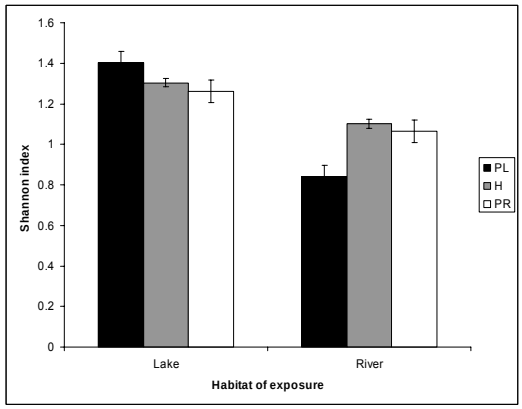


Figure IV.4: Effect of fish origin and habitat of exposure on parasite load. Bars show mean Shannon index (+/- standard error).

While some innate immune mechanisms can show specificity in their response (Schmid-Hempel & Ebert 2003) the majority are, however, rather unspecific (Parham 2003). The adaptive immune system on the other hand is characterized by high specificity due to somatic rearrangement of antigen receptors. To mount an adaptive immune response takes up to two weeks in mammals and up to several weeks in ectotherms such as fish. Interestingly splitting the parasite burden into specialist vs. generalist parasites revealed that the burden varied with MHC genotype only for a specialist parasite, the trematode *Gyrodactylus sp.* but not with generalist parasites.

A recent meta-analysis revealed a complex system where the local pattern of host-parasite co-evolution is a function of the parasite strategies (Lajeunesse & Forbes 2002). In brief, in a system driven by specialist parasites, these specialists should perform better on the local host compared to generalist parasites. Whereas in a system driven by generalist parasites, local hosts should perform better than migrants. Following this idea, we tested for MHC local adaptations in the hybrid

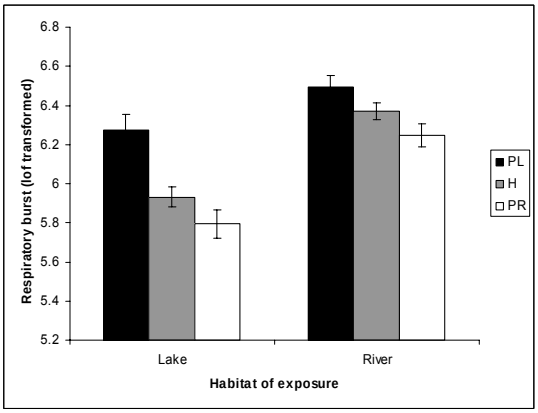


Figure IV.5: Effect of fish origin and habitat of exposure on respiratory burst. Bars show mean Shannon index (+/- standard error).

lines. First, we noted that fish exposed to the lake displayed a higher total parasite load. This result was consistent with previous studies from populations inhabiting the same drainage system (Wegner *et al.* 2003b; Kalbe & Kurtz 2006; Rauch *et al.* 2006b; Scharsack *et al.* 2007).

We found a pattern of local adaptation of parasites to local MHC genotypes, as fish with an MHC originating from lake (LL) were more infected in the lake habitat compared to fish bearing RR MHC genotypes. In turn, fish with an MHC originating from river (RR) were more infected in the river compared to fish with LL MHC genotype. Most surprisingly, this pattern was different when focusing only on generalist or specialist parasites respectively. Generalist parasite infection was not related to MHC genotype in any of the two habitats. In contrary, investigating parasite burden from a specialist parasite (*Gyrodactylus sp.*) revealed an advantage for fish bearing MHC genotypes originating from the river habitat (RR). This suggests that host MHC-parasite interaction is better seen with a specialist parasite. Generally, *Gyrodactylus* infection is higher in river than in lake (this study; Kalbe *et al.* 2002; Eizaguirre *et al.* Submitted-a), what could explain the advantage for river MHC genotypes, which might have evolved to confer resistance to this specific, but dominating parasite. Although carried out on a lake population, a recent study (Eizaguirre *et al.* Submitted-b, Annex I) showed an even tighter connection between the specialist *Gyrodactylus sp.* and a specific MHC haplotype, which might be the result of the high abundance of *Gyrodactylus* in that year. Fish carrying a specific haplotype were better at eliminating the parasite which in turn was reflected on the fish length. Fish length directly translated into Darwinian fitness, i.e. reproductive success. The results of the present study point in the same direction where host immune system (MHC) and specialist parasite are involved in a co-evolutionary arms race, where negative frequency-dependent selection could play a crucial role (Doherty & Zinkernagel 1975a; Takahata & Nei 1990).

The interaction observed on the total parasite load could arise from a habitat specific optimal trade-off of the host immunity between mounting a specific response against specialist parasites and activation of the innate immune system for fighting generalist invaders. This could arise from the non independence of both branches of the immune system (Dixon & Stet 2001).

This is best represented by the role of fish background genome (pure lake, pure river and hybrids) in the general local adaptation. Here we found as well that pure lake fish harboured a higher parasite load in lake than pure river fish, whilst pure river fish were more infected in river habitat, indicating that the local parasites adapted also to other components of the host immune system than MHC. Hybrid fish stand at an intermediate level of parasitisation. Again, total parasite load could reflect an optimal resistance strategy of the fish. Indeed, specialist parasite load was higher in river but was not related to fish genome origin, indicating the highly specific adaptation of the specialist parasites to only the host MHC and not to the rest of the genome. This was not the case for generalist parasites: Pure lake lines displayed a lower infection in both habitats, suggesting that the lake genome provides a variant of the immune system most adapted to cope with large numbers of generalist parasites. This is supported by the facts that most of the generalist parasites are extremely prevalent in lake and very little present in river and that the number of MHC alleles is generally higher in the lake, which could confer a broader detection of pathogens. This observation is congruent with previous lab experiments for selection of a generalist parasite, the trematode *Diplostomum pseudospathaceum* (Wegner *et al.* 2007). In this study, sticklebacks originating from susceptible parents displayed higher parasite infection than fish originating from resistant lines. This resistance was linked to the innate immunity. However, highly susceptible lines compensated lower resistance by increased MHC expression (see as well Wegner *et al.* 2006) bringing further evidence for the tight link between innate and acquired immune system. We could thus postulate a co-evolutionary process between host innate immune system and generalist parasites potentially backed-up by MHC heterozygosity. Again supported by the fact that pure lake fish showed higher MHC heterozygosity. A certain level of adapted MHC heterozygosity was proposed to decrease parasite infection as shown in different surveys and laboratory experiments (e.g. Doherty & Zinkernagel 1975a; Penn *et al.* 2002; Wegner *et al.* 2003a).

Further supporting such co-evolution between host innate immune system and parasites, we showed that fish of lake origin always displayed higher oxidative burst than river fish and this independent of the MHC

genotype. The respiratory burst reaction, which is based on the production of reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2) or superoxide anions (O_2^-) by granulocytes and macrophages is an innate immune mechanism of major relevance. This is interpreted as an adaptation to the higher generalist parasite abundance in the lake, as shown experimentally in sticklebacks infected with the generalist *Diplostomum pseudospathaceum* (Kalbe & Kurtz 2006). However, this result contrasts with findings from Scharsack *et al.* (2007) where river sticklebacks displayed a higher oxidative burst than lake fish.

Another striking difference between sticklebacks from the different origins was the relative size of their spleen, which is an important organ of the immune system. Lake sticklebacks in lake had much higher spleen weights than fish with river genome, while river fish displayed a higher spleen weight in river. Due to the reciprocal exposure design of our experiment, we can conclude that the actual size of this lymphatic organ seems to follow the pattern of local adaptation, and cannot be regarded a general characteristic of these populations. It was recently proposed that lower spleen size in the three-spined stickleback reflected best MHC adapted immunity and provided fish with higher fitness (Kalbe *et al.* In press). However, the mechanistic link awaits more experimental support.

In summary, we have conducted a field experiment where lab-bred fish were exposed for a substantial period of time to a range of natural parasites and could show that MHC genotype matters as does the genetic background. Interestingly we pointed out that MHC genes were mainly involved in interactions with specialist parasites, whereas genetic background and potentially MHC heterozygosity were important to resist generalist invasion. Thus, pre-requisites of a co-evolution between hosts and parasites have been demonstrated and met the first assumption for MHC to be considered as a magic trait.

We want to stress out that we are aware about the unfinished stage of this study and are confident to explain the observed processes in more detail with further analyses such as those including MHC expression, effects of specific MHC haplotype and trophic specialization with stable isotope signatures.

Acknowledgments

We thank G. Augustin, D. Martens and H. Deiwick for technical assistance with the cages. We also thank S. Bartel, K. Brzezek, W. Derner, G.Schmiedeskamp, I. Schultz, K. Weiss for inestimable help with the dissections and lab assistance. In addition, we thank the county of Schleswig-Holstein and the city of Ploen for permitting us to run the outdoor experiments in the lake and the river.

Chapter V

Parasites maintaining MHC polymorphism: an experimental test for negative frequency-dependent selection.

Christophe Eizaguirre, Tobias L. Lenz, Joern P. Scharsack, Martin Kalbe & Manfred Milinski.

Abstract

Genes of the Major Histocompatibility Complex are the most polymorphic genes in vertebrates and play a crucial pleiotropic function in immune defence and mate choice. Theoretical and observational work suggests that MHC polymorphism might be maintained by negative frequency-dependent selection by parasites. Due to a restricted optimal number of MHC loci, an individual cannot display the extreme MHC allelic diversity present in the population. Thus, individuals carrying specific MHC alleles that confer resistance against currently predominating parasites should perform better and should be favoured during female mating decision. We designed an experiment to test for negative frequency-dependent selection not only including parasite pressure but also competition among males and mate choice. We created 6 replicate populations of three-spined sticklebacks, formed of a number of lab-bred families in equal proportions and equal sex ratio. Three of these populations were exposed to the nematode *Camallanus lacustris*, while the three others were confronted with the nematode *Anguillicoloides crassus*. These parasites are present in the original population from which the founders of the lab-bred populations were caught. Letting the system evolve in controlled outside tanks, we predicted that females would mate with partners possessing resistance alleles against the experimental parasite and thus the frequency of such alleles would increase in the following generation. After raising the collected eggs in the lab, we confronted the offspring with both parasites again. We hypothesized that fish from parents that faced one parasite, e.g. *Camallanus lacustris*, will be less heavily infected by this parasite than those coming from parents that encountered *Anguillicoloides crassus* and vice versa.

Unpublished manuscript

Introduction

Host-parasite co-evolutionary arms races (Van Valen 1973) could be sufficient to explain the maintenance of sexual reproduction – a fundamental question in evolutionary biology (Hamilton 1980; Hamilton *et al.* 1990). Indeed, theory predicts that sexual reproduction overcomes the disadvantage of the low evolutionary rate in asexual hosts, compared to their rapidly co-evolving parasites (Maynard Smith 1978). The impact parasites have on the evolution and ecology of their hosts depends on their virulence. Virulence is usually defined as the parasite-induced loss of fitness (Ebert & Hamilton 1996) and varies with host genotypes. The principal underlying assumption is that genetic variation for host-parasite interaction exists within populations and leads to differential fitness between co-evolving organisms. Such host-parasite co-evolution maintains polymorphism in parasite virulence and host resistance genes (Frank 2002).

Genes of the major histocompatibility complex (MHC) are an important part of the vertebrate adaptive immune system (Klein 1986; Janeway *et al.* 2001). MHC genes encode MHC molecules that transport antigens from the cytoplasm and present them at the cell surface to T-cells. The T-cells recognize pathogen-derived peptides and subsequently initiate a specific immune response. Particularly, the MHC class II genes are involved in the immune answer against extra-cellular parasites by encoding MHC molecules that present exogenously derived peptides to CD 4+ T helper cells (Klein 1986; Janeway *et al.* 2001).

MHC genes are the most polymorphic genes found in jawed vertebrates (Bodmer 1972; Apanius *et al.* 1997). The relative roles of different processes in the maintenance of polymorphism at MHC loci are still under investigation, but balancing selection is thought to play an important role. High intra- and inter-individual polymorphism is thought to be parasite-mediated and maintained by heterozygote advantage (Doherty & Zinkernagel 1975a; Carrington *et al.* 1999; Suarez *et al.* 1999; McClelland *et al.* 2003) and/or by negative frequency-dependent selection (Bodmer 1972; Takahata & Nei 1990; for review see Apanius *et al.* 1997; Piertney & Oliver 2005; Milinski 2006). Theoretical works showed that heterozygote advantage alone cannot maintain MHC polymorphism (De Boer *et al.* 2004; Ding & Goudet 2005). Moreover,

theory would predict a stable polymorphism (Slade & McCallum 1992) that has so far not been reported. Under negative frequency-dependent selection, a rare allele is more likely to be resistant to common parasite genotypes, and will therefore confer selective advantages to its bearer. Consequently, such an allele should increase in frequency in the next generation. However, parasite adaptation to host resistance should lead to selection for a parasite or parasite strain, which is more virulent against this common genotype. In support of this model, many associations between MHC alleles and resistance to parasitism have been found in natural populations (e.g. Hill *et al.* 1991; Langefors *et al.* 2001; Grimholt *et al.* 2003; Harf & Sommer 2005; Bonneaud *et al.* 2006b; reviewed in Piertney & Oliver 2005; Milinski 2006). Vertebrates present a reduced optimal number of MHC loci and cannot display the entire potential MHC allele pool present in the populations (Lawlor *et al.* 1990; Janeway *et al.* 2001; Robinson *et al.* 2003). To overcome this restriction, MHC alleles might fluctuate in frequency over years.

Although theoretically straightforward, to our knowledge, only a handful of studies provided hints of a potential negative frequency-dependent scenario of resistance with MHC genes. Surveys of populations of great reed warblers (Westerdahl *et al.* 2004a), sticklebacks (Wegner 2004) and soay sheeps (Charbonnel & Pemberton 2005) represent the seldom examples of significant variation in MHC allele frequencies between successive age cohorts. Such variations were explained by potential fluctuating selective pressure from parasites. Contrary to the heterozygote advantage, negative frequency-dependent selection will maintain a dynamic polymorphism with fluctuating allele frequencies (Slade & McCallum 1992). Local parasite-driven selection favours some alleles and disfavours others in a frequency-dependent manner, and so will female mate choice.

There are two different ways for an allele to increase in frequency: either selection against individuals lacking this resistance allele, or specific mate choice for this allele. There is growing evidence for MHC-based mate choice, e.g in mice (Yamazaki *et al.* 1976; Potts *et al.* 1991; Penn & Potts 1998) in birds (Richardson *et al.* 2005; Bonneaud *et al.* 2006a), in fish (Reusch *et al.* 2001a; Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Forsberg *et al.* 2007; Consuegra & Garcia de Leaniz 2008) and in primates

(Schwensow *et al.* 2008a). For instance, female sticklebacks can assess their potential partner's MHC make up via perception of corresponding peptide ligands in odor signals (Milinski *et al.* 2005). Studies on great snipes and sticklebacks reported evidence for mate choice for males carrying specific MHC alleles (Ekblom *et al.* 2004; Eizaguirre *et al.* Submitted-b, Annex I). Thus, MHC genes possess the necessary prerequisites to investigate negative frequency-dependent selection.

Using the three-spined stickleback (*Gasterosteus aculeatus*) as a model organism, we designed an experiment to test for negative frequency-dependent selection including both parasite pressure and female mate choice. We used large outside concrete tanks in which the fish populations could evolve without any uncontrolled parasite or predatory pressures. We created 6 replicate populations: each was formed from an equal mixture of 6 different lab-bred families (F1 generation). Three experimental populations were exposed to the nematode *Camallanus lacustris*, while the three others were confronted with the nematode *Anguillicoloides crassus*. Both parasites were common in the population where the parental generation was caught. We assumed that females would mate with partners carrying alleles conferring resistance against the experimental parasite. This should lead to an increase in frequency of particular alleles in the following generation. During the reproductive period, we collected egg clutches and raised the offspring (F2 generation) under standardized conditions. After a certain growth period, we exposed the F2 generation to both parasites. We hypothesized that fish originating from parents that faced one parasite, e.g. *C. lacustris*, would be less heavily infected by this parasite than those coming from the population that encountered *A. crassus;* and vice versa. Investigating the association between parasites and specific MHC alleles, we aimed to demonstrate the role of negative frequency-dependent selection in maintaining MHC polymorphism.

Materials and Methods

This section of the manuscript describes the different steps of the experiment that have been performed so far. The fish of the F2 generation have been dissected, but analysis has not yet been performed and therefore

results are not available. In order to remain blind with respect to the parasitisation, some molecular work has been performed but not analyzed.

Study organism

Three-spined sticklebacks were caught in the lake Großer Plöner See (54° 9'21.16"N, 10°25'50.14"E, Germany) in May 2006. After catching, fish were held singly in 16l tanks under summer conditions (day light 16/8h, 18°C air and water temperature). Males were provided with artificial nesting material (sand and nylon threads). We randomly crossed fish in order to obtain lab-bred sibships, however each individual was only used once. Breeding took place naturally, i.e. males and females were allowed to display natural courtship behaviour and females spawned into the male's nest. Clutches were left for one hour in the nest after fertilization before being taken out. Individual egg clutches were incubated in aerated well water (with 0.04 ppm malachite green) at 18°C until hatching. We obtained 10 clutches and randomly kept six for the experiment. The fish stayed under summer condition (18°C, 16/24 day light) until November. Then the fry was transferred to autumn conditions for two weeks (12°C, 12/24 day light,) and to winter conditions for three weeks (6°C, 8/24 day light). Eventually the fish were brought to spring conditions (12°C, 12/24 day light,) for a 5 month period. Eleven days before release into the concrete enclosures (see below), the fish were accustomed to a longer day light and higher temperature to meet the external conditions (18°C, 16/24 day light).

DNA extraction, microsatellites and sex typing

In order to identify fish, a piece of the second dorsal spine was clipped for DNA typing. DNA extractions were performed using DNeasy extraction kit (Qiagen, UK) following the manufacturer’s protocol. All the fish were typed for 9 microsatellites combined in 2 different PCR multiplex protocols (Kalbe *et al.* In press). Moreover, fish were typed for sex using Griffiths *et al.*’s (2000) protocol adapted for high-throughput typing: One PCR primer was fluorescently labelled and the signal was subsequently detected on an automated sequencer. For this, 2 µl PCR product was mixed in 8 µl HiDi before a 2 minute denaturation step (90°C) followed by snap-cooling (4°C) for

5 min. An Applied Biosystem 3730 DNA Analyzer was used and runs were analyzed with GeneMarker (v.1.70). In the chromatograms males were represented by two peaks (370bp and 600bp) while females displayed only one peak (600bp). During DNA extraction and sex typing, fish were kept singly. Length and weight were recorded.

Laboratory populations and concrete outside tanks

By randomly assigning 7 males and 7 females per family to each of the six concrete outside tanks, we created six populations of 84 fish each and with an equal sex ratio. The outside concrete tanks had a capacity of about 46m^3 (4.8m x 4.8 m x 2m). On one side, the tanks had an additional shallower part (1m width x 0.5 m depth) in order to provide male sticklebacks with nesting areas. In each tank, the shallow area was divided into six territories of 80x60 cm using bricks (Fig V.1). Nesting material and sand were added at the beginning of the experiment. See figure V.2 for flow diagram of the full experiment.

Fish Infection

- *First generation (F1)*

Before being released into the outside tanks, the 604 sticklebacks from the F1 generation were infected twice with one of two different parasites. Three lab-populations were exposed to the nematode *Camallanus lacustris*, while the three others were exposed to the nematode *Anguillicoloides crassus*. These two parasites were chosen because they occur naturally in the population where the parental generation originated from and because resistance against them has been shown experimentally to be

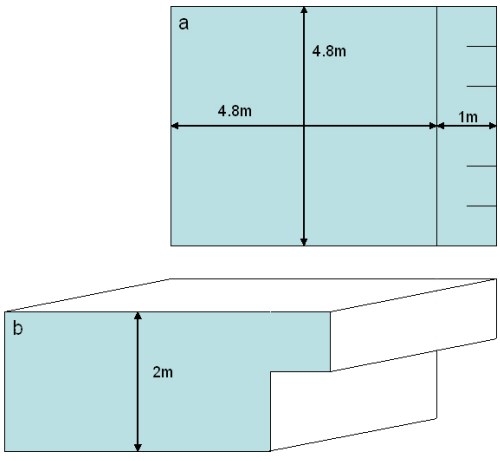


Figure V.1: Schematic representation of the concrete outside tanks. a) shows a view from above the basins, while b) shows a perspective view with the breeding area.

MHC-dependent (Wegner *et al.* 2003a).

C. lacustris ripe females were obtained from the blind sacks of perch guts. Larvae were isolated from single worms and counted. *A. crassus* larvae were collected after maternal worms, isolated from eel swim-bladders, had released eggs. These two parasites necessitate different intermediate hosts to achieve their complete life cycle, one of which is the three-spined stickleback. Stickleback infection occurs through trophic transmission. In the lab, we used *Macrocyclops sp.* copepods as intermediate hosts (se as well Wegner *et al.* 2003a; Krobbach *et al.* 2007). In the first round of fish infection, 6,000 copepods were exposed to more than 48,000 *C. lacustris* larvae. This resulted in an average of 8 larvae/copepods. We also exposed 3,000 copepods to 7,600 *A. crassus* larvae resulting in an average of 2.5 larvae per copepod. The infected copepods were kept during three weeks in 1l tanks in groups of 100 before fish infection. They were fed twice a week once with artemia and once with paramecia. This protocol was followed for each copepod infection.

To stimulate the adaptive immune system, a second round of infection took place 6 weeks later. We exposed 6,000 copepods to more than 48,000 *A. crassus* larvae and 3,000 copepods exposed to 21,000 *C. lacustris* larvae.

The infection status of the copepods was determined by microscopically counting nematode larvae inside the copepod before exposure to the fish. Variation in the intensity of infection in copepods prevented the exposure of all fish with an identical infection dosage. For the two rounds of infection, fish were kept singly in tanks overnight and only fed with the infected copepods.

- *Second generation (F2)*

Fry was raised in the lab. After 8 weeks, we isolated the fish in group of 20 in order to control for density effect in the growth. We randomly picked 30 clutches. Half originated from the outside tanks where the F1 fish had been exposed to *C. lacustris* and the other half from the outside tanks where the F1 fish had been exposed to *A. crassus*.

For each clutch, 15 randomly assigned F2 fish were exposed to *C. lacustris,* while another 15 were exposed to *A. crassus*. This resulted in a total of 900 fish to be infected, each half with one of the two parasites. Here as

well, in order to trigger the adaptive immune system of the fish, we exposed the fish twice to the parasites. For each exposure period, more than 7,000 copepods were bred in the lab and more than 40,000 parasite larvae were collected in the same way as for the F1 generation. Parasite larvae were always freshly collected from either perch guts or eel swim-bladders. Each copepod was checked individually and copepods were excluded if the exact number of larvae could not be determined.

In the first round, fish exposed to *C. lacustris* received 7 larvae, while the fish exposed to *A. crassus* received 9 larvae. In the second round of infection for the *C. lacustris* treatment, the fish were exposed to 7 parasite larvae, whilst the fish exposed to *A. crassus* received randomly from 6 to 9 larvae, because of limited availability of larvae. Before exposure to the parasites, F2 fish were weighed and length was recorded. At the same time the second dorsal spine was clipped for further molecular analysis.

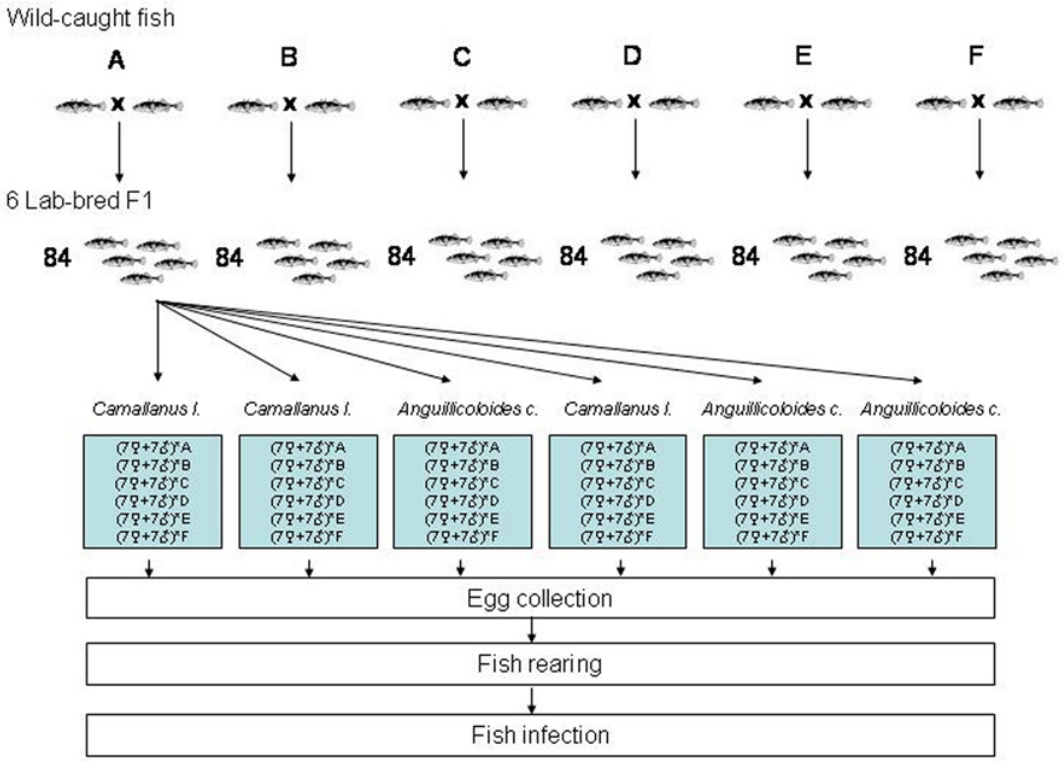


Figure V.2: Representation of the experimental design used for breeding and parasite exposure during the experiment.

Egg collection

To estimate reproductive success and to infer mate choice, once a week, all stickleback nests from the outside tanks were detected by careful observations. All egg batches were collected and the nests carefully replaced in their original location. Egg batches were brought to the laboratory. Within a nest, clutches were separated on the basis of the eggs' developmental stages. Until hatching, eggs were kept using 1l glass jars filled with well water and provided with air supply (for egg handling see Kalbe *et al.* In press).

F1 fish dissection

After 4 weeks in the outside tanks, the experiment was stopped since the number of collected clutches decreased markedly. Fish were recaptured and dissected. Dissection took place to record parasite intensity in the F1 fish. Dissections were done blind, i.e. fish were checked for both parasites. Fish length, weight and inner organs weight were recorded. A piece of tail fin was cut for further molecular analysis.

Immunological assays

For immunological assays, leucocytes were isolated from head kidney (HKL). Head kidney preparations followed Scharsack *et al.*'s protocol (2004). Total cell numbers in HKL isolates were determined with the standard cell dilution assay according to Scharsack *et al.* (2007). Additionally, flow cytometric measurements of freshly isolated HKL were used to determine proportions of granulocytes (FSC/SSChigh) and lymphocytes (FSC/SSClow) in individual HKL samples. As an important factor of the cell mediated innate immunity, the respiratory burst activity of HKL was quantified (Kurtz *et al.* 2004). As a parameter for activation of the adaptive immune system, frequencies of lymphocytes in G0-1, S and G2-M phases were acquired by DNA content analysis of red fluorescence intensity (propidium iodide labelling) of single cells from the lymphocyte gate (Scharsack *et al.* 2007) .

DNA extraction and microsatellite typing after dissection

DNA extractions from tail fin tissue were performed using the DNeasy extraction kit (Qiagen, UK) following the manufacturer's protocol. All F1 fish

were re-typed with the same 9 microsatellites as before the experiment plus an additional 6 to allow parenthood resolution. All were combined in 4 different PCR multiplex protocols (Kalbe *et al.* In press). These markers were used to reliably identify the fish at the end of the experiment and to perform parenthood analysis on the collected clutches.

Cloning and MHC typing

The twelve wild-caught parental fish were cloned and sequenced. In order to identify all MHC class II*B* alleles of the exon 2, we cloned and sequenced a fragment running from Exon 1 to Exon 2 of the MHC class II*B* genes, employing cautious PCR conditions according to Lenz & Becker (2008, Chapter I). Setting a primer in the highly conserved Exon 1 reduces the risk of potential primer bias during amplification. For each fish, 48 clones from at least two independent reactions were sequenced.

The MHC class IIB diversity of the F1 and F2 fish was determined by reference strand conformation analysis (RSCA) according to Lenz *et al.* (Chapter II). We amplified the exon 2 of MHC class IIB genes that encodes for the peptide-binding groove of the MHC II molecule. This exon has previously been suggested to be under parasite mediated positive selection as revealed by a high ratio of non-synonymous over synonymous substitutions (d_N/d_S, e.g. Reusch *et al.* 2004).

Initial results

MHC diversity in wild-caught parental fish

Individual MHC diversity varied from 3 to 5 alleles. In total the 12 founder fish displayed 17 different alleles.

Egg collection

During the four weeks the F1 fish spent in the outside tanks, we collected 46 clutches. The total number of eggs per clutch varied from 29 to 529.

Parental parasite load

In the parental generation we observed a strong variance in parasite infection between parasites and between individuals. 78.2 % of the fish exposed to *C. lacustris* were infected, compared to 42.1 % of the fish infected with *A. crassus*. Infection with *C. lacustris* ranged from 1 to 28 with a mean infection of 8.7 worms. Fish exposed to *A. crassus* displayed a lower mean infection of 1.8 worm. The range of worms found in the swim bladder varied between 1 and 6. Such variation encourages us to continue in our quest since differential host resistance represents the necessary prerequisite for to negative frequency-dependent selection by parasites.

Discussion

Here we aimed to test the role of negative frequency-dependent selection in the maintenance of MHC polymorphism. If our hypotheses were correct, female mate choice should lead to an increase in the frequency of resistance MHC alleles. For females, to be able to choose a mate with resistance alleles, males should signal their performance against parasites. Secondary sexual coloration of the throat of male stickleback has been investigated in length. Individual redness has been shown to indicate parasitisation (Milinski & Bakker 1990), body condition (Bakker & Mundwiler 1994) and MHC make-up (Jäger *et al.* 2007). Therefore, we expect males that are most able to resist parasitisation to be best able to gain a territory, maintain a nest and attract females. During the experiment, we performed immune assays and can thus assess the potential trade-off between parasitisation and activation of the immune system (Sheldon & Verhulst 1996). Fish with and without a specific resistance allele might use different components of the immune system to maximize resistance while minimizing energy allocated to the immune system.

Numerous hints exist in the literature to hypothesize the action of such parasite mediated negative frequency-dependent selection. First, various associations between MHC alleles and resistance to pathogen have been reported (e.g. Briles *et al.* 1977; Hill *et al.* 1991; Langefors *et al.* 2001; Lohm *et al.* 2002; for review see Haldane 1992; Piertney & Oliver 2005; Sommer 2005a; Milinski 2006). These associations are a pre-requisite to negative

frequency-dependent selection. Second, over time, the patterns of polymorphism maintained by heterozygote advantage and frequency-dependence should be very different. The latter would be reflected by a strong variance in allele frequencies between generations, while heterozygote advantage predicts an even and stable distribution of alleles (Slade & McCallum 1992). This last decade, MHC genes have been the focus of very diverse studies. Despite this diversity, all studies found a common trait: a strong variation of allele frequencies (e.g. Boyce *et al.* 1997; Wegner *et al.* 2003b; Harf & Sommer 2005; Meyer-Lucht & Sommer 2005; Cutrera & Lacey 2006). Eventually, three studies, dealing with a bird (Westerdahl *et al.* 2004a), a fish (Wegner 2004) and a mammal (Charbonnel & Pemberton 2005) as model organisms, revealed a dynamic fluctuation of MHC alleles between age cohorts; another prerequisite to polymorphism maintained by frequency-dependence (Slade & McCallum 1992).

Association between MHC alleles and resistance to parasite infection, dynamic fluctuations of MHC alleles as well as MHC-based mate choice are suggestive of negative frequency-dependent selection by parasites maintaining MHC polymorphism and ultimately explaining the necessary two-fold advantage of sexually reproducing over clonally reproducing organisms.

However, negative frequency-dependent selection can become more complex. Competition of co-infecting parasites might lead to the paradox of susceptible MHC alleles.

Almost as frequently as resistance alleles, studies have reported associations between MHC alleles and infection or disease susceptibility (Carrington *et al.* 1999; Hendel *et al.* 1999; Bonneaud *et al.* 2006b; Loiseau *et al.* 2008). The persistence of susceptible alleles represents an evolutionary paradox as there may be clear costs associated with carrying these alleles.

Theoretically, alleles that have antagonistic effects on fitness can be maintained in populations (Zhang *et al.* 2002; Loiseau *et al.* 2008) because a susceptible allele towards one parasite may confer resistance against another. This would correspond to two simultaneous arms races between one host and two parasites, where virulence between parasites would drive the system. However the prediction of antagonistic pleiotropy involving two parasites awaits empirical evidence.

Alternatively to a direct effect of antagonistic (resistance/susceptibility) MHC alleles on parasite infection, within-host-competition between parasites or parasite strains could result in the same pattern of susceptible alleles. Parasites exploiting the same resources in the host are potentially involved in competitive interactions and maintain each other in equilibrium. Specific immune recognition of different parasite antigens through MHC molecules can alter the competitive interaction between parasites and break the equilibrium. For instance, if allele A confers resistance against parasite P1, and if P1 and P2 are two parasites exploiting the same host, the allele A could reduce the competitive interactions between the two parasites. Theoretically this could result in an observed increase in susceptibility of P2 (Raberg *et al.* 2006; Loiseau *et al.* 2008). Although proposed, this theory as well needs experimental support that could be brought by further investigations based on our original experiment set-up.

The multitude of hints in favour of negative frequency-dependent selection by parasites deserved consideration and was the motivation to perform our experiment. We hope to better understand MHC polymorphism and thus contribute to the knowledge about one of the ultimate questions in evolutionary biology, namely the maintenance of sexual reproduction.

Acknowledgments

We thank G. Augustin, D. Martens for help with fish caring. We also thank S. Bartel, K. Brzezek, W. Derner, G.Schmiedeskamp, I. Schultz and K. Weiss for inestimable help with the copepod checking, dissection and lab assistance. We also want to thank the fishermen from the “Reese team” who provided us with fresh and parasitised perch and eel guts.

Conclusion

The achievements from my PhD work, pursued in tight collaboration with several colleagues, provide new insights into the highly polymorphic MHC genes of the three-spined stickleback and bring advances for the understanding of the evolutionary mechanisms that maintain this exceptional variability. First we developed a new amplification protocol, which takes into account the high risk of incorporating artefacts when amplifying multi-allele templates (Chapter I). Following the proposed protocol will increase reliability and reproducibility of MHC studies in the future. Combined with an also newly developed high throughput genotyping protocol for reference strand-mediated conformation analysis of MHC loci, a new molecular resource for the three-spined stickleback is now available (Chapter II). It enables reliable identification of individual MHC alleles and therefore provides information down to the sequence level without time-consuming cloning. For the future this will enable deeper insights into the mechanisms that maintain the polymorphism in the MHC by enabling allele specific experimental designs and analysis of host genotype-parasite interactions.

The new genotyping protocol provided the basis for an experimental investigation of MHC-based mate choice decisions under semi-natural conditions (Chapter III). Our results support previous work by showing MHC-dependent mate choice for intermediate diversity and influence of MHC genotypes on reproductively relevant traits. In addition, however, we also found haplotype-specific effects, which underline the necessity of investigating haplotype differences on the sequence level. This will shed more light onto the mechanistic processes, which underlie the maintenance of MHC polymorphism and ultimately enable hosts to keep pace in the Red Queen race of host-parasite co-evolution.

In another experiment, we investigated the spatial aspect of Red Queen dynamics by testing two reproductively isolated stickleback populations for levels of local immunogenetic adaptation to their respective sympatric parasite communities (Chapter IV). The results indicate that these co-evolutionary processes are highly complex and do not only depend on host

and parasite genotypes, but also on the parasites' strategy of host exploitation. These insights pave the way for further studies on the role of parasites in population divergence and ecological speciation.

In the last and still ongoing project, we are testing experimentally, whether frequency-dependent host-parasite co-evolution can occur (Chapter V). Although this is a widely hypothesised mechanism for the maintenance of MHC polymorphism in vertebrates, it has not yet been experimentally verified. Preliminary results indicate already substantial variation in parasite resistance between individuals. We will include host MHC genotypes in the analysis to understand the extent of host allele to parasite associations, which could ultimately explain the exceptional polymorphism in the MHC. The resulting insights will open up a new level of experimental research addressing genotype by genotype interactions in vertebrates between the highly adaptive MHC and their parasites.

Overall the presented results indicate that there is not one single mechanism mediating the co-evolution between vertebrate hosts and their parasites, supporting the notion from previous studies on sticklebacks and other species. More likely a full orchestra of different selection factors acting at the gene, the individual and the population level leads to the composition of individual genetic diversity as we can observe it nowadays in nature. It becomes apparent form this and earlier work, that on the one hand the MHC plays a crucial and adaptive role in the response to parasite pressures on different levels and on the other hand that parasites are maintaining host genetic diversity both in the MHC and in the genomic background via natural and sexual selection. The initially introduced hypothesis that the co-evolution between hosts and parasites favours and drives sexual reproduction as a mechanism to maintain variation has gained new support from the herein presented work.

References

Acinas S.G., Sarma-Rupavtarm R., Klepac-Ceraj V. & Polz M.F. (2005) PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Enviromental Microbiology*, 71, 8966-8969

Aeschlimann P.B., Haberli M.A., Reusch T.B.H., Boehm T. & Milinski M. (2003) Female sticklebacks *Gasterosteus aculeatus* use self-reference to optimize MHC allele number during mate selection. *Behavioral Ecology and Sociobiology*, 54, 119-126

Agapow P.-M. & Burt A. (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, 1, 101-102

Alberts B., Johnson A., Lewis J., Raff M., Roberts K. & Walter P. (2008) *Molecular biology of the cell*. 5th edn. Garland Science, New York, USA.

Alcaide M., Edwards S.V., Negro J.J., Serrano D. & Tella J.L. (2008) Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Molecular Ecology*, 17, 2652-2665

Altuvia Y. & Margalit H. (2004) A structure-based approach for prediction of MHC-binding peptides. *Methods*, 34, 454-459

Apanius V., Penn D., Slev P.R., Ruff L.R. & Potts W.K. (1997) The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology*, 17, 179-224.

Argüello J.R., Little A.M., Pay A.L., Gallardo D., Rojas I., Marsh S.G.E., Goldman J.M. & Madrigal J.A. (1998) Mutation detection and typing of polymorphic loci through double-strand conformation analysis. *Nature Genetics*, 18, 192-194

Argüello J.R. & Madrigal J.A. (1999) HLA typing by Reference Strand Mediated Conformation Analysis. *Reviews in Immunogenetics*, 1, 209-219

Argüello J.R., Pay A.L., McDermott A., Ross J., Dunn P., Avakian H., Little A.M., Goldman J. & Madrigal J.A. (1997) Complementary strand analysis: A new approach for allelic separation in complex polyallelic genetic systems. *Nucleic Acids Research*, 25, 2236-2238

Axtner J. & 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

Ayliffe M.A., Lawrence G.J., Ellis J.G. & Pryor A.J. (1994) Heteroduplex molecules formed between allelic sequences cause nonparental RAPD bands. *Nucleic Acids Research*, 22, 1632-1636

Babik W., Pabijan M. & Radwan J. (2008) Contrasting patterns of variation in MHC loci in the Alpine newt. *Molecular Ecology*, 17, 2339-2355

Bakker T.C.M., Kunzler R. & Mazzi D. (1999) Sexual selection - Condition-related mate choice in sticklebacks. *Nature*, 401, 234-234

Bakker T.C.M. & Mundwiler B. (1994) Female mate choice and male red coloration in a natural three-spined stickleback (*Gasterosteus aculeatus*) population. *Behavioral Ecology*, 5, 74-80

Baquero J., Miranda S., Murillo O., Mateus H., Trujillo E., Suarez C., Patarroyo M. & Parra-López C. (2006) Reference strand conformational analysis (RSCA) is a valuable tool in identifying MHC-DRB sequences in three species of *Aotus* monkeys. *Immunogenetics*, 58, 590-597

Barrett R.D.H., Rogers S.M. & Schluter D. (2008) Natural selection on a major armor gene in threespine stickleback. *Science*, 322, 255-257

Barrett R.D.H. & Schluter D. (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23, 38-44

Barton N.H., Briggs D.E.G., Eisen J.A., Goldstein D.B. & Patel N.H. (2007) *Evolution*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Barton N.H. & Charlesworth B. (1998) Why Sex and Recombination? *Science*, 281, 1986-1990

Beaumont M.A. (2008) Selection and sticklebacks. *Molecular Ecology*, 17, 3425-3427

Beck S. & Trowsdale J. (2000) The human major histocompatibility complex: Lessons from the DNA sequence. *Annual Review of Genomics and Human Genetics*, 1, 117-137

Becker S., Boger P., Oehlmann R. & Ernst A. (2000) PCR Bias in Ecological Analysis: a Case Study for Quantitative Taq Nuclease Assays in Analyses of Microbial Communities. *Applied and Environmental Microbiology*, 66, 4945-4953

Bell M.A. & Foster S.A. (1994) *The Evolutionary Biology of the Threespine Stickleback*. Oxford University Press, Oxford.

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

Binz T., Reusch T.B.H., Wedekind C. & Milinski M. (2001) SSCP analysis of Mhc class IIB genes in the threespine stickleback. *Journal of Fish Biology*, 58, 887-890

Blais J., Rico C. & Bernatchez L. (2004) Nonlinear effects of female mate choice in wild threespine sticklebacks. *Evolution*, 58, 2498-2510

Blais J., Rico C., van Oosterhout C., Cable J., Turner G.F. & Bernatchez L. (2007) MHC Adaptive Divergence between Closely Related and Sympatric African Cichlids. *PLoS ONE*, 2, e734

Blasky A., Karl J., Wiseman R., Read D. & O'Connor D. (2008) Rapid high-resolution MHC class I genotyping of Chinese rhesus macaques by capillary reference strand-mediated conformational analysis. *Immunogenetics*, OnlineEarly

Bodmer W.F. (1972) Evolutionary significance of the HL-A system. *Nature*, 237, 139-183

Boehm T. (2006) Co-evolution of a primordial peptide-presentation system and cellular immunity. *Nature Reviews Immunology*, 6, 79-84

Bolnick D.I. (2004) Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution*, 58, 608-18

Bonneaud C., Chastel O., Federici P., Westerdahl H. & Sorci G. (2006a) Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society of London B: Biological Sciences*, 273, 1111-1116

Bonneaud C., Mazuc J., Chastel O., Westerdahl H. & Sorci G. (2004a) Terminal investment induced by immune challenge and fitness traits associated with major histocompatibility complex in the house sparrow. *Evolution*, 58, 2823-2830

Bonneaud C., Perez-Tris J., Federici P., Chastel O. & Sorci G. (2006b) Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, 60, 383-389

Bonneaud C., Sorci G., Morin V., Westerdahl H., Zoorob R. & Wittzell H. (2004b) Diversity of Mhc class I and IIB genes in house sparrows (*Passer domesticus*). *Immunogenetics*, 55, 855-65

Bontrop R.E., Otting N., de Groot N.G. & Doxiadis G.G.M. (1999) Major histocompatibility complex class II polymorphisms in primates. *Immunological Reviews*, 167, 339-350

Borghans J., Noest A. & De Boer R. (2003) Thymic selection does not limit the individual MHC diversity. *European Journal of Immunology*, 33, 3353-8.

Borriello F. & Krauter K.S. (1990) Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction (PCR + 1). *Nucleic Acids Research*, 18, 5481-5487

Boughman J.W. (2001) Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, 411, 944-948

Boughman J.W. (2007) Condition-dependent expression of red colour differs between stickleback species. *Journal of Evolutionary Biology*, 20, 1577 - 1590

Boughman J.W., Rundle H.D. & Schluter D. (2005) Parallel evolution of sexual isolation in sticklebacks. *Evolution*, 59, 361-373

Bowen L., Aldridge B.M., Gulland F., Bonn W., DeLong R., Melin S., Lowenstine L.J., Stott J.L. & Johnson M.L. (2004) Class II multiformity generated by variable MHC-DRB region configurations in the California sea lion (*Zalophus californianus*). *Immunogenetics*, 56, 12-27

Bower M.A., Spencer M., Matsumura S., Nisbet R.E.R. & Howe C.J. (2005) How many clones need to be sequenced from a single forensic or ancient DNA sample in order to determine a reliable consensus sequence? *Nucleic Acids Research*, 33, 2549-2556

Boyce W.M., Hedrick P.W., MuggliCockett N.E., Kalinowski S., Penedo M.C.T. & Ramey R.R. (1997) Genetic variation of major histocompatibility complex and micrcosatellite loci: A comparison in bighorn sheep. *Genetics*, 145, 421-433

Bradley R.D. & Hillis D.M. (1997) Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution*, 14, 592-593

Briles W.E., Stone H.A. & Cole R.K. (1977) Marek's Disease: Effects of B Histocompatibility Alloalleles in Resistant and Susceptible Chicken Lines. *Science*, 195, 193-195

Brown J.H., Jardetzky T.S., Gorga J.C., Stern L.J., Urban R.G., Strominger J.L. & Wiley D.C. (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, 364, 33-39

Brown J.J., Thomson W., Clegg P., Eyre S., Kennedy L.J., Matthews J., Carter S. & Ollier W.E.R. (2004) Polymorphisms of the equine major histocompatibility complex class II DRA locus. *Tissue Antigens*, 64, 173-179

Bryja J., Galan M., Charbonnel N. & Cosson J.F. (2006) Duplication, balancing selection and trans-species evolution explain the high levels of polymorphism of the DQA MHC class II gene in voles (Arvicolinae). *Immunogenetics*, 58, 191-202

Buchholz R., Dukes M.D.J., Hecht S. & Findley A.M. (2004) Investigating the turkey's 'snood' as a morphological marker of heritable disease resistance. *Journal of Animal Breeding and Genetics*, 121, 176-185

Bui H.-H., Schiewe A.J., Grafenstein H.v. & Haworth I.S. (2006) Structural prediction of peptides binding to MHC class I molecules. *Proteins: Structure, Function, and Bioinformatics*, 63, 43-52

Carrington M., Nelson G.W., Martin M.P., Kissner T., Vlahov D., Goedert J.J., Kaslow R., Buchbinder S., Hoots K. & O'Brien S.J. (1999) HLA and HIV-1: Heterozygote Advantage and B*35-Cw*04 Disadvantage. *Science*, 283, 1748-1752

Charbonnel N. & Pemberton J. (2005) A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity*, 95, 377-388

Clarke B. & Kirby D.R.S. (1966) Maintenance of histocompatibility polymorphisms. *Nature*, 211, 999-1000

Clarke K.R. & Gorley R.N. (2006) PRIMER v6: User Manual/Tutorial. In. PRIMER-E, Plymouth

Colosimo P.F., Hosemann K.E., Balabhadra S., Villarreal G., Dickson M., Grimwood J., Schmutz J., Myers R.M., Schluter D. & Kingsley D.M. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, 307, 1928-1933

Consuegra S. & Garcia de Leaniz C. (2008) MHC-mediated mate choice increases parasite resistance in salmon. *Proceedings of the Royal Society of London B: Biological Sciences*, 275, 1397-1403

Corell A., Pay A.L., Argüello J.R., Borton M., Dunne C., Madrigal J.A. & Marsh S.G.E. (1999) RSCA (reference strand mediated conformation analysis) resolves HLA-DRB1*04 typing ambiguities where other molecular techniques fail. *Blood*, 94, 318B-318B

Cresko W., McGuigan K., Phillips P. & Postlethwait J. (2007) Studies of threespine stickleback developmental evolution: progress and promise. *Genetica*, 129, 105-126

Cresko W.A. (2008) Special issue on stickleback behavior and evolution. *Behaviour*, 145, 409-411

Cutrera A.P. & Lacey E.A. (2006) Major histocompatibility complex variation in Talas Tuco-Tucos: The influence of demography on selection. *Journal of Mammalogy*, 87, 706-716

Darwin C.R. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. 1st edn. John Murray, London.

de Bakker P.I.W., McVean G., Sabeti P.C., Miretti M.M., Green T., Marchini J., Ke X., Monsuur A.J., Whittaker P., Delgado M., Morrison J., Richardson A., Walsh E.C., Gao X., Galver L., Hart J., Hafler D.A., Pericak-Vance M., Todd J.A., Daly M.J., Trowsdale J., Wijmenga C., Vyse T.J., Beck S., Murray S.S., Carrington M., Gregory S., Deloukas P. & Rioux J.D. (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nature Genetics*, 38, 1166-1172

De Boer R.J., Borghans J.A., van Boven M., Kesmir C. & Weissing F.J. (2004) Heterozygote advantage fails to explain the high degree of polymorphism of the MHC. *Immunogenetics*, 55, 725-731

de Campos-Lima P.-O., Gavioli R., Zhang Q.-J., Wallace L.E., Dolcetti R., Rowe M., Rickinson A.B. & Masucci M.G. (1993) HLA-A11 epitope loss isolates of Epstein-Barr Virus from a highly A11+ population. *Science*, 260, 98-100

Ding A. & Goudet J. (2005) Heterozygote advantage and the maintenance of polymorphism for multilocus traits. *Theoretical Population Biology*, 68, 157-166

Dixon B. & Stet R.J.M. (2001) The relationship between major histocompatibility receptors and innate immunity in teleost fish. *Developmental & Comparative Immunology*, 25, 683-699

Doherty P.C. & Zinkernagel R.M. (1975a) A biological role for the major histocompatibility antigens. *The Lancet*, 305, 1406-1409

Doherty P.C. & Zinkernagel R.M. (1975b) Enhanced immunological surveillance in mice heterozygous at H-2 gene complex. *Nature*, 256, 50-52

Doxiadis G.G.M., Otting N., de Groot N.G., Noort R. & Bontrop R.E. (2000) Unprecedented polymorphism of MHC-DRB region configurations in rhesus macaques. *Journal of Immunology*, 164, 3193-3199

Drake G.J.C., Kennedy L.J., Auty H.K., Ryvar R., Ollier W.E.R., Kitchener A.C., Freeman A.R. & Radford A.D. (2004) The use of reference strand-mediated conformational analysis for the study of cheetah (Acinonyx jubatus) feline leucocyte antigen class II DRB polymorphisms. *Molecular Ecology*, 13, 221-229

Duchesne P., Godbout M.-H. & Bernatchez L. (2002) papa (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. *Molecular Ecology Notes*, 2, 191-193

Ebert D. (1994) Virulence and local adaptation of a horizontally transmitted parasite. *Science*, 265, 1084-1086

Ebert D. & Hamilton W.D. (1996) Sex against virulence: the coevolution of parasitic diseases. *Trends in Ecology & Evolution*, 11, 79-82

Edwards S.V., Grahn M. & Potts W.K. (1995a) Dynamics of Mhc evolution in birds and crocodilians: Amplification of class II genes with degenerate primers. *Molecular Ecology*, 4, 719-729

Edwards S.V. & Hedrick P.W. (1998) Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends in Ecology and Evolution*, 13, 305-311

Edwards S.V., Wakeland E.K. & Potts W.K. (1995b) Contrasting histories of avian and mammalian Mhc genes revealed by class II B sequences from songbirds. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 12200-12204.

Eizaguirre C., Lenz T.L., Traulsen A. & Milinski M. (2009) Speciation accelerated and stabilized by pleiotropic major histocompatibility complex immunogenes. *Ecology Letters*, 12, 5-12

Eizaguirre C., Sommerfeld R.D., Lenz T.L., Harrod C., Kalbe M. & Milinski M. (Submitted-a) Ecological pleiotropy: pre- and post-zygotic reproductive isolation through MHC mediated selection.

Eizaguirre C., Yeates S., Lenz T.L., Kalbe M. & Milinski M. (Submitted-b) MHC-based mate choice combines good genes and maintenance of MHC polymorphism.

Ekblom R., Saether S.A., Grahn M., Fiske P., Kalas J.A. & Hoglund J. (2004) Major histocompatibility complex variation and mate choice in a lekking bird, the great snipe (*Gallinago media*). *Molecular Ecology*, 13, 3821-3828

Ekblom R., Saether S.A., Jacobsson P.A.R., Fiske P., Sahlman T., Grahn M., Kalas J.A. & Hoglund J. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology*, 16, 1439-1451

Eklund A. (1997) The major histocompatibility complex and mating preferences in wild house mice (*Mus domesticus*). *Behavioural Ecology*, 8, 630-634

Feichtlbauer-Huber P., Stear M.J., Fries R. & Buitkamp J. (2000) Reference-strand-mediated conformation analysis of MHC alleles: a new method for high-resolution typing of the Ovar-DQB genes. *Immunogenetics*, 51, 65-68

Forsberg L.A., Dannewitz J., Petersson E. & Grahn M. (2007) Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout - females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology*, 20, 1859-1869

Frank S.A. (2002) *Immunology and the evolution of infectious disease*. Princeton University Press, Princeton, USA.

Frankham R., Ballou J.D. & Briscoe D.A. (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Froeschke G. & Sommer S. (2005) MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the southern Kalahari. *Molecular Biology and Evolution*, 22, 1254-1259

Gaston K.J. (2000) Global patterns in biodiversity. *Nature*, 405, 220-227

Gibson G. (2005) Evolution: The synthesis and evolution of a supermodel. *Science*, 307, 1890-1891

Gibson G. & Dworkin I. (2004) Uncovering cryptic genetic variation. *Nature Reviews Genetics*, 5, 681-90

Gow J.L., Peichel C.L. & Taylor E.B. (2007) Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. *Journal of Evolutionary Biology*, 20, 2173-2180

Griffiths R., Orr K.L., Adam A. & Barber I. (2000) DNA sex identification in the three-spined stickleback. *Journal of Fish Biology*, 57, 1331-1334

Grimholt U., Larsen S., Nordmo R., Midtlyng P., Kjoeglum S., Storset A., Saebø 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

Guderley H. & Guevara R.C. (1998) Physiological correlates of paternal care in male threespine stickleback. *Ecoscience*, 5, 438-447

Haldane J.B.S. (1992) Disease And Evolution (Reprinted From La Ricerca Scientifica Supplemento, Vol 19, Pg 1-11, 1949). *Current Science*, 63, 599-604

Hall T. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98

Hamilton W.D. (1980) Sex Versus Non-Sex Versus Parasite. *Oikos*, 35, 282-290

Hamilton W.D., Axelrod R. & Tanese R. (1990) Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences of the United States of America*, 87, 3566-3573

Hamilton W.D. & Zuk M. (1982) Heritable True Fitness and Bright Birds - a Role for Parasites. *Science*, 218, 384-387

Harf R. & Sommer S. (2005) Association between major histocompatibility complex class II DRB alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paeba*, in the southern Kalahari. *Molecular Ecology*, 14, 85-91

Harrod C., Grey J., McCarthy T.K. & Morrissey M. (2005) Stable isotope analyses provide new insights into ecological plasticity in a mixohaline population of European eel. *Oecologia*, 144, 673-683

Hedrick P. (1999) Balancing selection and MHC. *Genetica*, 104, 207-214

Hedrick P.W. (2006) Genetic polymorphism in heterogeneous environments: The age of genomics. *Annual Review of Ecology Evolution and Systematics*, 37, 67-93

Hedrick P.W. & Thomson G. (1983) Evidence for balancing selection at HLA. *Genetics*, 104, 449-456

Hendel H., Caillat-Zucman S., Lebuanec H., Carrington M., O'Brien S., Andrieu J.-M., Schachter F., Zagury D., Rappaport J., Winkler C., Nelson G.W. & Zagury J.-F. (1999) New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS. *Journal of Immunology*, 162, 6942-6946

Hennessy L.K., Teare J. & Ko C. (1998) PCR Conditions and DNA Denaturants Affect Reproducibility of Single-Strand Conformation Polymorphism Patterns for BRCA1 Mutations. *Clinical Chemistry*, 44, 879-882

Hermel E., Yuan J. & Monaco J.J. (1995) Charactreization of polymorphism within the H-2M MHC class II loci. *Immunogenetics*, 42, 136-142

Hibbeler S., Scharsack J. & Becker S. (2008) Housekeeping genes for quantitative expression studies in the three-spined stickleback Gasterosteus aculeatus. *BMC Molecular Biology*, 9, 18

Hill A.V., Allsopp C.E., Kwiatkowski D., Anstey N.M., Twumasi P., Rowe P.A., Bennett S., Brewster D., McMichael A.J. & Greenwood B.M. (1991) Common west African HLA antigens are associated with protection from severe malaria. *Nature*, 352, 595-600.

Horton R., Gibson R., Coggill P., Miretti M., Allcock R.J., Almeida J., Forbes S., Gilbert J.G.R., Halls K., Harrow J.L., Hart E., Howe K., Jackson D.K., Palmer S., Roberts A.N., Sims S., Stewart C.A., Traherne J.A., Trevanion S., Wilming L., Rogers J., de Jong P.J., Elliott J.F., Sawcer S., Todd J.A., Trowsdale J. & Beck S. (2008) Variation analysis and gene annotation of eight MHC haplotypes: The MHC haplotype project. *Immunogenetics*, 60, 1-18

Horton R., Niblett D., Milne S., Palmer S., Tubby B., Trowsdale J. & Beck S. (1998) Large-scale sequence comparisons reveal unusually high levels of variation in the HLA-DQB1 locus in the class II region of the human MHC. *Journal of Molecular Biology*, 282, 71-97

Hubbard T.J.P., Aken B.L., Beal K., Ballester B., Caccamo M., Chen Y., Clarke L., Coates G., Cunningham F., Cutts T., Down T., Dyer S.C., Fitzgerald S., Fernandez-Banet J., Graf S., Haider S., Hammond M., Herrero J., Holland R., Howe K., Howe K., Johnson N., Kahari A., Keefe D., Kokocinski F., Kulesha E., Lawson D., Longden I., Melsopp C., Megy K., Meidl P., Ouverdin B., Parker A., Prlic A., Rice S., Rios D., Schuster M., Sealy I., Severin J., Slater G., Smedley D., Spudich G., Trevanion S., Vilella A., Vogel J., White S., Wood M., Cox T., Curwen V., Durbin R., Fernandez-Suarez X.M., Flicek P., Kasprzyk A., Proctor G., Searle S., Smith J., Ureta-Vidal A. & Birney E. (2007) Ensembl 2007. *Nucleic Acids Research*, 35, D610-617

Huchard E., Cowlishaw G., Raymond M., Weill M. & Knapp L. (2006) Molecular study of Mhc-DRB in wild chacma baboons reveals high variability and evidence for trans-species inheritance. *Immunogenetics*, 58, 805-816

Hughes A.L. & Nei M. (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335, 167-170.

Hughes A.L. & Nei M. (1989) Nucleotide substitution at major histocompatibility complex class-II loci - evidence for overdominant selection. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 958-962

Ilmonen P., Penn D.J., Damjanovich K., Morrison L., Ghotbi L. & Potts W.K. (2007) Major histocompatibility complex heterozygosity reduces fitness in experimentally infected mice. *Genetics*, 176, 2501-2508

Jäger I., Eizaguirre C., Griffiths S.W., Kalbe M., Krobbach C.K., Reusch T.B., Schaschl H. & Milinski M. (2007) Individual MHC class I and MHC class IIB diversities are associated with male and female reproductive traits in the three-spined stickleback. *Journal of Evolutionary Biology*, 20, 2005-15

Janeway C.A., Travers P., Walport M. & Shlomchik M.J. (2001) *Immunobiology: the immune system in health and disease*. 5th edn. Garland Science, London.

Janeway C.A., Travers P., Walport M. & Shlomchik M.J. (2005) *Immunobiology: the immune system in health and disease*. 6th edn. Garland Science, London.

Jansen R. & Ledley F.D. (1990) Disruption of phase during PCR amplification and cloning of heterozygous target sequences. *Nucleic Acids Research*, 18, 5153-5156

Judo M.S.B., Wedel A.B. & Wilson C. (1998) Stimulation and suppression of PCR-mediated recombination. *Nucleic Acids Research*, 26, 1819-1825

Kalbe M., Eizaguirre C., Dankert I., Reusch T.B.H., Sommerfeld R.D., Wegner K.M. & Milinski M. (In press) Lifetime reproductive success is maximised with optimal MHC diversity.

Kalbe M. & Kurtz J. (2006) Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke Diplostomum pseudospathaceum. *Parasitology*, 132, 105-116

Kalbe M., Wegner K.M. & Reusch T.B.H. (2002) Dispersion patterns of parasites in 0+ year three-spined sticklebacks: a cross population comparison. *Journal of Fish Biology*, 60, 1529-1542

Kaltz O. & Shykoff J.A. (1998) Local adaptation in host-parasite systems. *Heredity*, 81, 361-370

Kanagawa T. (2003) Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*, 96, 317-323

Karpenko O., Huang L. & Dai Y. (2008) A probabilistic meta-predictor for the MHC class II binding peptides. *Immunogenetics*, 60, 25-36

Kawecki T.J. (1998) Red queen meets Santa Rosalia: Arms races and the evolution of host specialization in organisms with parasitic lifestyles. *American Naturalist*, 152, 635-651

Keightley P.D. & Eyre-Walker A. (2000) Deleterious mutations and the evolution of sex. *Science*, 290, 331-333

Kelley J., Walter L. & Trowsdale J. (2005) Comparative genomics of major histocompatibility complexes. *Immunogenetics*, 56, 683-695

Kennedy L., Ryvar R., Gaskell R., Addie D., Willoughby K., Carter S., Thomson W., Ollier W. & Radford A. (2002) Sequence analysis of MHC DRB alleles in domestic cats from the United Kingdom. *Immunogenetics*, 54, 348-352

Kennedy L.J., Quarmby S., Fretwell N., Martin A.J., Jones P.G., Jones C.A. & Ollier W.E.R. (2005) High-Resolution Characterization of the Canine DLA-DRB1 Locus Using Reference Strand-Mediated Conformational Analysis. *Journal of Heredity*, 96, 836-842

Kennedy L.J., Ryvar R., Brown J.J., Ollier W.E.R. & Radford A.D. (2003) Resolution of complex feline leukocyte antigen DRB loci by reference strand-mediated conformational analysis (RSCA). *Tissue Antigens*, 62, 313-323

Khrapko K., Hanekamp J.S., Thilly W.G., Belenkii A., Foret F. & Karger B.L. (1994) Constant denaturant capillary electrophoresis (CDCE): a high resolution approach to mutational analysis. *Nucleic Acids Research*, 22, 364-369

Kiljunen M., Grey J., Sinisalo T., Harrod C., Immonen H. & Jones R.I. (2006) A revised model for lipid-normalizing delta C-13 values from aquatic organisms, with implications for isotope mixing models. *Journal of Applied Ecology*, 43, 1213-1222

Klein J. (1979) The major histocompatibility complex of the mouse. *Science*, 203, 516-21

Klein J. (1986) *Natural history of the major histocompatibility complex*. John Wiley &Sons, New York.

Klein J. & Figueroa F. (1986) Evolution of the Major Histocompatibility Complex. *Critical Reviews in Immunology*, 6, 295-386

Klein J., Ono H., Klein D. & O'hUigin C. (1993a) The accordion model of Mhc evolution. In: *Progress in Immunology* (eds. Gergely J & Petranyi G), pp. 137-143. Springer-Verlag, Heidelberg

Klein J., Sato A., Nagl S. & O'HUigin C. (1998) Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics*, 29, 1-21

Klein J., Sato A. & Nikolaidis N. (2007) MHC, TSP, and the Origin of Species: From Immunogenetics to Evolutionary Genetics. *Annual Review of Genetics*, 41, 281-304

Klein J., Satta Y., O'HUigin C. & Takahata N. (1993b) The molecular descent of the major histocompatibility complex. *Annual Review of Immunology*, 11, 269-295

Knapp L.A. (2005) Facts, faeces and setting standards for the study of MHC genes using noninvasive samples. *Molecular Ecology*, 14, 1597-1599

Kokko H. (2001) Fisherian and "good genes" benefits of mate choice: how (not) to distinguish between them. *Ecology Letters*, 4, 322-326

Kondrashov A.S. (1993) Classification of Hypotheses on the Advantage of Amphimixis. *Journal of Heredity*, 84, 372-387

Krobbach C.K., Kalbe M., Kurtz J. & Scharsack J.P. (2007) Infectivity of two nematode parasites, Camallanus lacustris and Anguillicola crassus, in a paratenic host, the three-spined stickleback Gasterosteus aculeatus. *Diseases of Aquatic Organisms*, 74, 119-126

Kurtz J., Kalbe M., Aeschlimann P.B., Haberli M.A., Wegner K.M., Reusch T.B.H. & Milinski M. (2004) Major histocompatibility complex diversity influences parasite resistance and innate immunity in sticklebacks. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271, 197-204

Kurtz J., Kalbe M., Langefors S., Mayer I., Milinski M. & Hasselquist D. (2007) An experimental test of the immunocompetence handicap hypothesis in a teleost fish: 11-ketotestosterone suppresses innate immunity in three-spined sticklebacks. *American Naturalist*, 170, 509-519

L'Abbé D., Belmaaza A., Décary F. & Chartrand P. (1992) Elimination of heteroduplex artifacts when sequencing HLA genes amplified by polymerase chain-reaction (PCR). *Immunogenetics*, 35, 395-397

Lajeunesse M.J. & Forbes M.R. (2002) Host range and local parasite adaptation. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 269, 703-710

Landry C., Garant D., Duchesne P. & Bernatchez L. (2001) 'Good genes as heterozygosity': The major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proceedings of the Royal Society of London B: Biological Sciences*, 268, 1279-1285

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 B: Biological Sciences*, 268, 479-485

Largiadèr C.R., Fries V., Kobler B. & Bakker T.C. (1999) Isolation and characterization of microsatellite loci from the three-spined stickleback (*Gasterosteus aculeatus* L.). *Molecular Ecology*, 8, 342-4

Lawlor D.A., Zemmour J., Ennis P.D. & Parham P. (1990) Evolution of Class-I Mhc Genes and Proteins - from Natural-Selection to Thymic Selection. *Annual Review of Immunology*, 8, 23-63

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

Lenz T.L., Eizaguirre C., Becker S. & Reusch T.B.H. (Submitted) RSCA genotyping of MHC for high-throughput evolutionary studies in the model organism three-spined stickleback Gasterosteus aculeatus.

Lessa E.P. & Applebaum G. (1993) Screening techniques for detecting allelic variation in DNA sequences. *Molecular Ecology*, 2, 119-129.

Lipsitch M., Bergstrom C.T. & Antia R. (2003) Effect of human leukocyte antigen heterozygosity on infectious disease outcome: the need for allele-specific measures. *BMC Medical Genetics*, 4, 2

Lively C.M. (1987) Evidence from a New Zealand snail for the maintenance of sex by parasitism. *Nature*, 328, 519-521

Lively C.M. (1996) Host-Parasite Coevolution and Sex. *Bioscience*, 46, 107-114

Lively C.M. (1999) Migration, virulence, and the geographic mosaic of adaptation by parasites. *American Naturalist*, 153, S34-S47

Lively C.M. & Dybdahl M.F. (2000) Parasite adaptation to locally common host genotypes. *Nature*, 405, 679-681

Lohm J., Grahn M., Langefors Å., Andersen Ø., 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 B: Biological Sciences*, 269, 2029-2033

Loiseau C., Zoorob R., Garnier S., Birard J., Federici P., Julliard R. & Sorci G. (2008) Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecology Letters*, 11, 258 - 265

Longeri M., Zanotti M. & Damiani G. (2002) Recombinant DRB sequences produced by mismatch repair of heteroduplexes during cloning in Escherichia coli. *European Journal of Immunogenetics*, 29, 517-523

Lozano G.A. (1991) Optimal Foraging Theory: A Possible Role for Parasites. *Oikos*, 60, 391-395

Madsen T. & Ujvari B. (2006) MHC class I variation associates with parasite resistance and longevity in tropical pythons. *Journal of Evolutionary Biology*, 19, 1973-1978

Mäkinen H.S. & Merilä J. (2008) Mitochondrial DNA phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe--Evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, 46, 167-182

Makino R., Yazyu H., Kishimoto Y., Sekiya T. & Hayashi K. (1992) F-SSCP: fluorescence-based polymerase chain reaction-single-strand conformation polymorphism (PCR-SSCP) analysis. *PCR Methods and Applications*, 2, 10-13

Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V., Figueroa F., Sultmann H. & Klein J. (1998) Linkage relationships and haplotype polymorphism among cichlid Mhc class II B loci. *Genetics*, 149, 1527-1537

Marsh S.G.E., Albert E.D., Bodmer W.F., Bontrop R.E., Dupont B., Erlich H.A., Geraghty D.E., Hansen J.A., Hurley C.K., Mach B., Mayr W.R., Parham P., Petersdorf E.W., Sasazuki T., Schreuder G.M.T., Strominger J.L., Svejgaard A., Terasaki P.I. & Trowsdale J. (2005) Nomenclature for factors of the HLA system, 2004. *Tissue Antigens*, 65, 301-369

Maynard Smith J. (1966) Sympatric Speciation. *The American Naturalist*, 100, 637-650

Maynard Smith J. (1978) *The evolution of sex*. Cambridge University Press, Cambridge, UK.

Maynard Smith J. (1991) Theories of Sexual Selection. *Trends in Ecology & Evolution*, 6, 146-151

Mays J.H.L. & Hill G.E. (2004) Choosing mates: good genes versus genes that are a good fit. *Trends in Ecology & Evolution*, 19, 554-559

McClelland E.E., Penn D.J. & Potts W.K. (2003) Major histocompatibility complex heterozygote superiority during coinfection. *Infection and Immunity*, 71, 2079-2086

McKinnon J.S., Mori S., Blackman B.K., David L., Kingsley D.M., Jamieson L., Chou J. & Schluter D. (2004) Evidence for ecology's role in speciation. *Nature*, 429, 294-298

McKinnon J.S. & Rundle H.D. (2002) Speciation in nature: the threespine stickleback model systems. *Trends in Ecology & Evolution*, 17, 480-488

Meyer-Lucht Y., Otten C., Puttker T. & Sommer S. (2008) Selection, diversity and evolutionary patterns of the MHC class II DAB in free-ranging Neotropical marsupials. *BMC Genetics*, 9, 39

Meyer-Lucht Y. & Sommer S. (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). *Molecular Ecology*, 14, 2233-2243

Meyerhans A., Vartanian J.-P. & Wain-Hobson S. (1990) DNA recombination during PCR. *Nucleic Acids Research*, 18, 1687-1691

Milinski M. (2006) The major histocompatibility complex, sexual selection, and mate choice. *Annual Review of Ecology, Evolution, and Systematics*, 37, 159-186

Milinski M. & Bakker T.C.M. (1990) Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. *Nature*, 344, 330-333

Milinski M., Griffiths S., Wegner K.M., Reusch T.B.H., Haas-Assenbaum A. & Boehm T. (2005) Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 4414-4418

Milinski M. & Heller R. (1978) Influence of a Predator on Optimal Foraging Behavior of Sticklebacks (Gasterosteus-Aculeatus L). *Nature*, 275, 642-644

Miller C.T., Beleza S., Pollen A.A., Schluter D., Kittles R.A., Shriver M.D. & Kingsley D.M. (2007a) cis-regulatory changes in kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*, 131, 1179-1189

Miller M.R., Dunham J.P., Amores A., Cresko W.A. & Johnson E.A. (2007b) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, 17, 240-248

Modrich P. (1987) DNA Mismatch Correction. *Annual Review of Biochemistry*, 56, 435-466

Møller A.P. (1990a) Effects of a haematophagous mite on the barn swallow (*Hirundo rustica*): A test of the Hamilton and Zuk hypothesis. *Evolution*, 44, 771-784

Møller A.P. (1990b) Parasites and sexual selection: Current status of the Hamilton and Zuk hypothesis. *Journal of Evolutionary Biology*, 3, 319-328

Møller A.P. & Alatalo R.V. (1999) Good-genes effects in sexual selection. *Proceedings of the Royal Society B: Biological Sciences*, 266, 85-91

Moore C.B., John M., James I.R., Christiansen F.T., Witt C.S. & Mallal S.A. (2002) Evidence of HIV-1 Adaptation to HLA-Restricted Immune Responses at a Population Level. *Science*, 296, 1439-1443

Muller H.J. (1964) The relation of recombination to mutational advance. *Mutation Research*, 106, 2-9

Mullis K., Faloona F., Scharf S., Saiki R., Horn G. & Erlich H. (1986) Specific Enzymatic Amplification of DNA Invitro - the Polymerase Chain-Reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, 51, 263-273

Myers R.M., Maniatis T. & Lerman L.S. (1987) Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods in Enzymology*, 155, 501-527

Nagamine C.M., Chan K. & Lau Y.F. (1989) A PCR artifact: generation of heteroduplexes. *American Journal of Human Genetics*, 45, 337-339

Neff B.D., Garner S.R., Heath J.W. & Heath D.D. (2008) The MHC and non-random mating in a captive population of Chinook salmon. *Heredity*, 101, 175-185

Neff B.D. & Pitcher T.E. (2005) Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology*, 14, 19-38

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

Nei M., Gu X. & Sitnikova T. (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 7799-7806.

Noakes M.A., Reimer T. & Phillips R.B. (2003) Genotypic Characterization of an MHC Class II Locus in Lake Trout (*Salvelinus namaycush*) from Lake Superior by Single-Stranded Conformational Polymorphism Analysis and Reference Strand-Mediated Conformational Analysis. *Marine Biotechnology*, 5, 270-278

Nowak M.A., Tarczyhornoch K. & Austyn J.M. (1992) The Optimal Number of Major Histocompatibility Complex-Molecules in an Individual. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 10896-10899

Oksanen J., Kindt R., Legendre P., O'Hara B., Simpson G.L., Henry M., Stevens H. & Wagner H. (2008) vegan: Community Ecology Package. In. R package version 1.13-2

Oleksiak M.F., Churchill G.A. & Crawford D.L. (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, 32, 261-266

Ono H., Ohuigin C., Vincek V. & Klein J. (1993) Exon-Intron Organization of Fish Major Histocompatibility Complex Class-Ii B-Genes. *Immunogenetics*, 38, 223-234

Orita M., Iwahana H., Kanazawa H., Hayashi K. & Sekiya T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 2766-2770

Orti G., Bell M.A., Reimchen T.E. & Meyer A. (1994) Global Survey of Mitochondrial DNA Sequences in the Threespine Stickleback: Evidence for Recent Migrations. *Evolution*, 48, 608-622

Parham P. (2003) Innate immunity: The unsung heroes. *Nature*, 423, 20-20

Paterson S., Wilson K. & Pemberton J.M. (1998) Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 3714-3719

Peichel C.L., Nereng K.S., Ohgi K.A., Cole B.L.E., Colosimo P.F., Buerkle C.A., Schluter D. & Kingsley D.M. (2001) The genetic architecture of divergence between threespine stickleback species. *Nature*, 414, 901-905

Penn D.J., Damjanovich K. & Potts W.K. (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11260-11264

Penn D.J. & Potts W. (1998) MHC-disassortative mating preferences reversed by cross-fostering. *Proceedings of the Royal Society of London, Series B*, 265, 1299-1306

Penn D.J. & Potts W.K. (1999) The evolution of mating preferences and major histocompatibility complex genes. *American Naturalist*, 153, 145-164

Peters B., Bui H.-H., Frankild S., Nielsen M., Lundegaard C., Kostem E., Basch D., Lamberth K., Harndahl M., Fleri W., Wilson S.S., Sidney J., Lund O., Buus S. & Sette A. (2006) A Community Resource Benchmarking Predictions of Peptide Binding to MHC-I Molecules. *PLoS Computational Biology*, 2, e65

Piertney S.B. & Oliver M.K. (2005) The evolutionary ecology of the major histocompatibility complex. *Heredity*, 96, 7-21

Potts W.K., Manning C.J. & Wakeland E.K. (1991) Mating Patterns in Seminatural Populations of Mice Influenced by Mhc Genotype. *Nature*, 352, 619-621

Potts W.K., Manning C.J., Wakeland E.K. & Hughes A.L. (1994) The Role of Infectious Disease, Inbreeding and Mating Preferences in Maintaining MHC Genetic Diversity: An Experimental Test [and Discussion]. *Philosophical Transactions: Biological Sciences*, 346, 369-378

Potts W.K. & Wakeland E.K. (1990) Evolution of diversity at the major histocompatibility complex. *Trends in Ecology and Evolution*, 5, 181-187

R Development Core Team (2007) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raberg L., de Roode J.C., Bell A.S., Stamou P., Gray D. & Read A.F. (2006) The role of immune-mediated apparent competition in genetically diverse malaria infections. *American Naturalist*, 168, 41-53

Rauch G., Kalbe M. & Reusch T.B.H. (2006a) One day is enough: rapid and specific host–parasite interactions in a stickleback-trematode system. *Biology Letters*, 2, 382-384

Rauch G., Kalbe M. & Reusch T.B.H. (2006b) Relative importance of MHC and genetic background for parasite load in a field experiment. *Evolutionary Ecology Research*, 8, 373-386

Read A.F. (1987) Comparative evidence supports the Hamilton and Zuk hypothesis on parasites and sexual selection. *Nature*, 328, 68-70

Read A.F. (1988) Sexual selection and the role of parasites. *Trends in Ecology & Evolution*, 3, 97-102

Reusch T.B., Haberli M.A., Aeschlimann P.B. & Milinski M. (2001a) Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*, 414, 300-302

Reusch T.B., Schaschl H. & Wegner K.M. (2004) Recent duplication and inter-locus gene conversion in major histocompatibility class II genes in a teleost, the three-spined stickleback. *Immunogenetics*, 56, 427-437

Reusch T.B.H. & Langefors A. (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

Reusch T.B.H., Wegner K.M. & Kalbe M. (2001b) Rapid genetic divergence in postglacial populations of threespine stickleback (Gasterosteus aculeatus): the role of habitat type, drainage and geographical proximity. *Molecular Ecology*, 10, 2435-2445

Rice W.R. (1987) Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evolutionary Ecology*, 1, 301-314

Richardson D., Komdeur J., Burke T. & von Schantz T. (2005) MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society of London B: Biological Sciences*, 272, 759-767

Ridley M. (1996) *Evolution*. 2nd edn. Blackwell Publishing, Oxford, UK.

Robinson J., Waller M.J., Parham P., Groot N.d., Bontrop R., Kennedy L.J., Stoehr P. & Marsh S.G.E. (2003) IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Research*, 31, 311-314

Rundle H.D., Nagel L., Boughman J.W. & Schluter D. (2000) Natural Selection and Parallel Speciation in Sympatric Sticklebacks. *Science*, 287, 306-308

Ryan M.J. (2004) Sexual selection and mate choice. In: *Behavioural Ecology* (eds. Krebs JR & Davies NB). Blackwell Publishing, Oxford

Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B. & Erlich H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, 487-491

Salathé M., Kouyos R.D. & Bonhoeffer S. (2008) The state of affairs in the kingdom of the Red Queen. *Trends in Ecology & Evolution*, 23, 439-445

Sato A., Figueroa F., O'Huigin C., Steck N. & Klein J. (1998) Cloning of major histocompatibility complex (Mhc) genes from threespine stickleback, Gasterosteus aculeatus. *Molecular Marine Biology and Biotechnology*, 7, 221-231

Sawyer S. (1989) Statistical tests for detecting gene conversion. *Molecular Biology and Evolution*, 6, 526-538

Sawyer S.A. (1999) GENECONV: A computer package for the statistical detection of gene conversion

Scharsack J.P., Kalbe M., Derner R., Kurtz J. & Milinski M. (2004) Modulation of granulocyte responses in three-spined sticklebacks Gasterosteus aculeatus infected with the tapeworm Schistocephalus solidus. *Diseases of Aquatic Organisms*, 59, 141-150

Scharsack J.P., Kalbe M., Harrod C. & Rauch G. (2007) Habitat-specific adaptation of immune responses of stickleback (Gasterosteus aculeatus) lake and river ecotypes. *Proceedings of the Royal Society of London B: Biological Sciences*, 274, 1523-1532

Schlichting C.D. (2008) Hidden Reaction Norms, Cryptic Genetic Variation, and Evolvability. *Annals of the New York Academy of Sciences*, 1133, 187-203

Schluter D. (1994) Experimental evidence that competition promotes divergence in adaptive radiation. *Science*, 266, 798-801

Schmid-Hempel P. & Ebert D. (2003) On the evolutionary ecology of specific immune defence. *Trends in Ecology & Evolution*, 18

Schwensow N., Eberle M. & Sommer S. (2008a) Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proceedings of the Royal Society of London B: Biological Sciences*, 275, 555-564

Schwensow N., Fietz J., Dausmann K. & Sommer S. (2008b) MHC-associated mating strategies and the importance of overall genetic diversity in an obligate pair-living primate. *Evolutionary Ecology*, 22, 617-636

Schwensow N., Fietz J., Dausmann K.H. & Sommer S. (2007) Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity*, 99, 265-277

Shapiro M.D., Bell M.A. & Kingsley D.M. (2006) Parallel genetic origins of pelvic reduction in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 13753-13758

Shapiro M.D., Marks M.E., Peichel C.L., Blackman B.K., Nereng K.S., Jonsson B., Schluter D. & Kingsley D.M. (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, 428, 717-723

Sheldon B.C. & Verhulst S. (1996) Ecological immunology: costly parasite defences and trade-offs in evolutionary ecology. *Trends in Ecology & Evolution*, 11, 317-321

Slade R.W. & McCallum H.I. (1992) Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*, 132, 861-862

Smith M.Z., Dale C.J., De Rose R., Stratov I., Fernandez C.S., Brooks A.G., Weinfurter J., Krebs K., Riek C., Watkins D.I., O'Connor D.H. & Kent S.J. (2005) Analysis of pigtail macaque major histocompatibility complex class I molecules presenting immunodominant simian immunodeficiency virus epitopes. *Journal of Virology*, 79, 684-695

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

Sommer S. (2005b) Major histocompatibility complex and mate choice in a monogamous rodent. *Behavioral Ecology and Sociobiology*, 58, 181-189

Sorci G., Møller A.P. & Boulinier T. (1997) Genetics of host-parasite interactions. *Trends in Ecology & Evolution*, 12, 196-200

Spehr M., Kelliher K.R., Li X.H., Boehm T., Leinders-Zufall T. & Zufall F. (2006) Essential role of the main olfactory system in social recognition of major histocompatibility complex peptide ligands. *Journal of Neuroscience*, 26, 1961-1970

Suarez S., Woolley-Barker T., Jolly C.J., Disotell T.R., Phillips-Conroy J.E. & Newman T.K. (1999) A possible case of MHC heterozygote advantage in a baboon hybrid zone. *American Journal of Physical Anthropology*, 261-261

Summers K., McKeon S., Sellars J., Keusenkothen M., Morris J., Gloeckner D., Pressley C., Price B. & Snow H. (2003) Parasitic exploitation as an engine of diversity. *Biological Reviews*, 78, 639-675

Svanback R. & Bolnick D.I. (2007) Intraspecific competition drives increased resource use diversity within a natural population. *Proceedings of the Royal Society of London B: Biological Sciences*, 274, 839-44

Takahata N. & Nei M. (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*, 124, 967-978

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

Tanaka-Takahashi Y., Yasunami M., Naruse T., Hinohara K., Matano T., Mori K., Miyazawa M., Honda M., Yasutomi Y., Nagai Y. & Kimura A. (2007) Reference strand-mediated conformation analysis-based typing of multiple alleles in the rhesus macaque MHC class I *Mamu-A* and *Mamu-B* loci. *Electrophoresis*, 28, 918-924

Thompson J.N. (1994) *The coevolutionary process*. University of Chicago Press, Chicago.

Thompson J.R., Marcelino L.A. & Polz M.F. (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Research*, 30, 2083-2088

Thursz M.R., Kwiatkowski D., Allsopp C.E., Greenwood B.M., Thomas H.C. & Hill A.V. (1995) Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. *New England Journal of Medicine*, 332, 1065-1069

Tollenaere C., Bryja J., Galan M., Cadet P., Deter J., Chaval Y., Berthier K., Salvador A.R., 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

Tregenza T. & Wedell N. (2000) Genetic compatibility, mate choice and patterns of parentage: Invited review. *Molecular Ecology*, 9, 1013-1027

Trowsdale J. (1993) Genomic Structure and Function in the Mhc. *Trends in Genetics*, 9, 117-122

Trowsdale J. (1995) “Both man & bird & beast”: comparative organization of MHC genes. *Immunogenetics*, 41, 1-17

Turelli M., Barton N.H. & Coyne J.A. (2001) Theory and speciation. *Trends in Ecology & Evolution*, 16, 330-343

v. Wintzingerode F., Gobel U.B. & Stackebrandt E. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews*, 21, 213-229

Valdar W.S.J. (2002) Scoring residue conservation. *Proteins-Structure Function and Genetics*, 48, 227-241

Van Valen L. (1973) A new evolutionary law. *Evolutionary Theory*, 1, 1-30

Vines T.H. & Schluter D. (2006) Strong assortative mating between allopatric sticklebacks as a by-product of adaptation to different environments. *Proceedings of the Royal Society of London B: Biological Sciences*, 273, 911-916

von Schantz T., Wittzell H., Goransson G., Grahn M. & Persson K. (1996) MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proceedings of the Royal Society of London B: Biological Sciences*, 263, 265-71

Wagner A., Blackstone N., Cartwright P., Dick M., Misof B., Snow P., Wagner G.P., Bartels J., Murtha M. & Pendleton J. (1994) Surveys of Gene Families Using Polymerase Chain Reaction: PCR Selection and PCR Drift. *Systematic Biology*, 43, 250-261

Wakeland E.K., Boehme S., She J.X., Lu C.C., McIndoe R.A., Cheng I., Ye Y. & Potts W.K. (1990) Ancestral polymorphisms of MHC class-II genes - divergent allele advantage. *Immunologic Research*, 9, 115-122

Wedekind C. (1997) The infectivity, growth, and virulence of the cestode *Schistocephalus solidus* in its first intermediate host, the copepod *Macrocyclops albidus*. *Parasitology*, 115, 317-24

Wedekind C., Walker M. & Little T.J. (2005) The course of malaria in mice: Major histocompatibility complex (MHC) effects, but no general MHC heterozygote advantage in single-strain infections. *Genetics*, 170, 1427-1430

Wedekind C., Walker M. & Little T.J. (2006) The separate and combined effects of MHC genotype, parasite clone, and host gender on the course of malaria in mice. *BMC Genetics*, 7, 55

Wegner K.M. (2004) Major histocompatibility genes, polymorphism and balancing selection the case of parasites and sticklebacks. In. Christian-Albrechts-University, Kiel, Germany

Wegner K.M., Kalbe M., Kurtz J., Reusch T.B.H. & Milinski M. (2003a) Parasite selection for immunogenetic optimality. *Science*, 301, 1343-1343

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

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

Wegner K.M., Martin K. & Thorsten B.H.R. (2007) Innate versus adaptive immunity in sticklebacks: evidence for trade-offs from a selection experiment. *Evolutionary Ecology*

Wegner K.M., Reusch T.B.H. & Kalbe M. (2003b) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of Evolutionary Biology*, 16, 224-232

West S.A., Lively C.M. & Read A.F. (1999) A pluralist approach to sex and recombination. *Journal of Evolutionary Biology*, 12, 1003-1012

Westerdahl H. (2004) No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology*, 13, 2465-2470

Westerdahl H., Hansson B., Bensch S. & Hasselquist D. (2004a) Between-year variation of MHC allele frequencies in great reed warblers: selection or drift? *Journal of Evolutionary Biology*, 17, 485-492

Westerdahl H., Waldenstrom J., Hansson B., Hasselquist D., von Schantz T. & Bensch S. (2005) Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society of London B: Biological Sciences*, 272, 1511-1518

Westerdahl H., Wittzell H., von Schantz T. & Bensch S. (2004b) MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity*, 92, 534-42

Whitehead A. & Crawford D.L. (2006) Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology*, 15, 1197-1211

Whyte S.K., Chappell L.H. & Secombes C.J. (1989) Cytotoxic reactions of rainbow trout, *Salmo gairdneri* Richardson, macrophages for larvae of the eye fluke *Diplostomum spathaceum* (Digenea). *Journal of Fish Biology*, 35, 333-345

Wildman D.E., Uddin M., Liu G., Grossman L.I. & Goodman M. (2003) Implications of natural selection in shaping 99.4% nonsynonymous DNA

identity between humans and chimpanzees: enlarging genus Homo. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 7181-8

Wilhelm B.T., Marguerat S., Watt S., Schubert F., Wood V., Goodhead I., Penkett C.J., Rogers J. & Bahler J. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*, online early

Woelfing B., Traulsen A., Milinski M. & Boehm T. (2009) Does intra-individual major histocompatibility complex diversity keep a golden mean? *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 117-128

Wolf J.B., Harrod C., Brunner S., Salazar S., Trillmich F. & Tautz D. (2008) Tracing early stages of species differentiation: ecological, morphological and genetic divergence of Galapagos sea lion populations. *BMC Evolutionary Biology*, 8, 150

Wootton R.J. (1984) *A functional biology of sticklebacks*. Croom Helm Ltd, Kent, UK.

Worley K., Gillingham M., Jensen P., Kennedy L.J., Pizzari T., Kaufman J. & Richardson D.S. (2008) Single locus typing of MHC class I and class IIB loci in a population of red jungle fowl. *Immunogenetics*, 60, 233-247

Yamazaki K., Boyse E.A., Mike V., Thaler H.T., Mathieson B.J., Abbott J., Boyse J., Zayas Z.A. & Thomas L. (1976) Control of Mating Preferences in Mice by Genes in Major Histocompatibility Complex. *Journal of Experimental Medicine*, 144, 1324-1335

Yang Z. (2007) PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, 24, 1586-1591

Yang Z.H., Nielsen R., Goldman N. & Pedersen A.M.K. (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155, 431-449

Young K., Frodsham A., Doumbo O.K., Gupta S., Dolo A., Hu J.T., Robson K.J.H., Crisanti A., Hill A.V.S. & Gilbert S.C. (2005) Inverse associations of human leukocyte antigen and malaria parasite types in two West African populations. *Infection and Immunity*, 73, 953-955

Zahavi A. (1975) Mate selection--A selection for a handicap. *Journal of Theoretical Biology*, 53, 205-214

Zhang X.-S., Wang J. & Hill W.G. (2002) Pleiotropic model of maintenance of quantitative genetic variation at mutation-selection balance. *Genetics*, 161, 419-433

Ziegler A., Ehlers A., Forbes S., Trowsdale J., Volz A., Younger R. & Beck S. (2000) Polymorphisms in olfactory receptor genes: a cautionary note. *Human Immunology*, 61, 1281-1284

Zuk M., Thornhill R., Ligon J.D. & Johnson K. (1990) Parasites and Mate Choice in Red Jungle Fowl. *American Zoologist*, 30, 235-244

Zylstra P., Rothenfluh H.S., Weiller G.F., Blanden R.V. & Steele E.J. (1998) PCR amplification of murine immunoglobulin germline V genes: Strategies for minimization of recombination artefacts. *Immunology and Cell Biology*, 76, 395-405

Annex I

MHC-based mate choice combines good genes and maintenance of MHC polymorphism

Christophe Eizaguirre, Sarah E Yeates, Tobias L. Lenz, Martin Kalbe & Manfred Milinski

Abstract

Polymorphic genes of the major histocompatibility complex (MHC) are regarded as essential genes for individual fitness under conditions of natural and sexual selection. To test this hypothesis, we investigated the ultimate individual fitness trait- that of reproductive success. We used three-spined sticklebacks (*Gasterosteus aculeatus*) in semi natural enclosures, located in natural breeding areas where the experimental fish had been caught. During their reproductive period, fish were exposed continuously to their natural sympatric parasites. By genotyping almost 4000 eggs with 9 microsatellites, we determined parenthood and inferred female mating decision. We found that with reference to their own MHC profile, female sticklebacks preferred to mate with males sharing an intermediate MHC diversity. In addition, males with a specific MHC haplotype were bigger and better at fighting a common parasite (*Gyrodactylus* sp.). This translated directly into Darwinian fitness since fish harboring this specific MHC haplotype were more likely to be chosen and had a higher reproductive output. We conclude that females based their mating decision also on a specific MHC haplotype conferring resistance against a common parasite. This identifies and supports “good genes”. We argue that such interaction between host and parasite driving assortative mating is not only a pre-requisite for negative-frequency-dependent selection - a potential mechanism to explain the maintenance of MHC polymorphism, but also potentially speciation.

Introduction

Female mate choice commonly occurs in the animal kingdom (Andersson 1994). Females are expected to exert active mate-choice if the benefits, direct or indirect, of such choosiness outweigh the costs. For instance, costs could include increased predation risk while seeking partners, whereas benefits could stem from direct gains such as paternal care or indirect advantages such as good genes (Andersson 1994). Good genes are genes that increase individual fitness (Andersson 1994). They can be transmitted to the next generation by mating with a male that bears an advantageous allele conferring, e.g. higher body condition in specific environments. One could hypothesize that phenotypically recognizable and inheritable immune genes represent an example of good genes due to adaptation to the immediate pathogenic environment (Hamilton& Zuk 1982).

Recently, genes of the major histocompatibility complex (MHC) have received major attention across different fields (Bonneaud *et al.* 2005; Janeway *et al.* 1999; Kurtz *et al.* 2006; Lawlor *et al.* 1990; Milinski *et al.* 2005; Nowak *et al.* 1992; Parham& Ohta 1996; Schat *et al.* 1994; Westerdahl 2004; Yamazaki *et al.* 1976) as they appear to play a key role at the interface between natural and sexual selection (review in Milinski 2006).
The MHC is a large cluster of genes involved in immune response and molecular self/non-self discrimination. MHC class I genes encode MHC molecules that are expressed on nearly every cell type. They serve invigilation against intra-cellular parasites, *e.g.* viruses, mostly by presenting antigens to cytoxic CD 8+ T cells. MHC class II genes initiate defense against extra-cellular parasites: pathogen-derived peptides are presented to CD 4+ T helper cells to trigger a specific immune response (Janeway *et al.* 1999; Klein 1986). Hence, it is not surprising that several reports have shown correlations between certain MHC genotypes and the occurrence or severity of specific diseases and parasite infections (Bonneaud *et al.* 2005; Briles *et al.* 1983; Godot *et al.* 2000; Grimholt *et al.* 2003; Harf& Sommer 2005; Hill *et al.* 1991; Horton *et al.* 2004; Langefors *et al.* 2001).

Interest in MHC genes has expanded with evidence for its extreme polymorphism. The enormous allelic diversity found in natural populations is generally assumed to be a consequence of parasite-mediated balancing selection (Apanius *et al.* 1997; Bernatchez& Landry 2003; Clarke& Kirby 1966; Edwards& Hedrick 1998; Jordan& Bruford 1998; Mays& Hill 2004; Milinski 2006; Penn *et al.* 2002; Penn& Potts 1999; Takahata& Nei 1990). Pathogen driven selection could act on MHC genes through heterozygote advantage, frequency dependent selection and/or habitat heterogeneity (Apanius *et al.* 1997; Doherty& Zinkernagel 1975; Ekblom *et al.* 2007; Hedrick 2002; Penn& Potts 1999).

Furthermore, MHC polymorphism has been suggested to be also under sexual selection (Apanius *et al.* 1997; Bernatchez& Landry 2003; Edwards& Hedrick 1998; Milinski 2006). Indeed, several studies found evidence for MHC-based mate choice (Aeschlimann *et al.* 2003; Bonneaud *et al.* 2006a; Forsberg *et al.* 2007; Landry *et al.* 2001; Milinski *et al.* 2005; Olsson *et al.*

2005; Penn& Potts 1999; Potts *et al.* 1991; Reusch *et al.* 2001; Schwensow *et al.* 2008; Wedekind *et al.* 1995). Females are assumed to search for partners in order to acquire the best immunogenetic make up for their offspring, while they may also avoid mating with closely related males (Brown& Eklund 1994; Penn& Potts 1998; Potts *et al.* 1991; Yamazaki *et al.* 1976). In humans, mice and sticklebacks, olfactory based mate choice has been shown to be MHC linked (Aeschlimann *et al.* 2003; Egid& Brown 1989; Milinski *et al.* 2005; Ninomiya& Brown 1995; Reusch *et al.* 2001; Wedekind *et al.* 1995). Additionally, various secondary sexual traits, known to influence mate choice, have been correlated to MHC, e.g. the snood in turkey (Buchholz *et al.* 2004), and the red throat in stickleback (Jäger *et al.* 2007).

Combining analyses at neutral and adaptive genetic loci provides an efficient tool for evaluating the importance of various selective mechanisms, such as mate choice, on the evolution of immunogenetic competence. In contrast to MHC genotypes, which directly translate into phenotypes (Bonneaud *et al.* 2006b; Langefors *et al.* 2001; Lohm *et al.* 2002; Paterson *et al.* 1998; Wegner *et al.* 2003b), microsatellites are assumed to behave neutrally and variation is thought to be primarily driven by non-selective evolutionary factors such as mutations, gene flow or genetic drift (Nei 1987).

The three-spined stickleback (*Gasterosteus aculeatus*) is a perfect model organism to study the function of mate choice strategies. During the reproductive period, territorial males maintain a nest. The nest has been demonstrated to be central for MHC and male factor signaling (Milinski *et al.* 2005; Sommerfeld *et al.* In press). Based on olfactory as well as visual cues, females make their mating decision and spawn in the nest. Before chasing the female away, the territorial male will fertilize the eggs unless a sneaker male steals the fertilization. After fertilization, the territorial male guards the eggs and adopts a costly fanning behavior over the eggs until they hatch (Wootton 1976). Males display an intense red breeding coloration that has been shown to indicate individual parasitization status (Milinski& Bakker 1990), body condition (Bakker& Mundwiler 1994) and MHC genotype (Jäger *et al.* 2007). Olfactory-based mate choice experiments demonstrated that females choose partners providing, with reference to her own MHC make up, an optimal number of MHC variants to the offspring (Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Reusch *et al.* 2001). Such female mate choice resulted in a higher reproductive success in lab-bred fish bearing an intermediate MHC diversity (Kalbe *et al.* In Press).

In the present study, we investigated female MHC-based mate choice of wild-caught three-spined stickleback through its reproductive period in a field enclosure system. The enclosures were stocked with a representative sub-sample of the native population. These fish were exposed to their natural spectrum of sympatric parasites, but protected from any kind of predatory pressure. Every week, we collected egg clutches and genotyped a representative number of eggs per clutch with microsatellites. We could thus assign both parents and infer female mate choice (Kalbe *et al.* In Press). We predicted that individual fitness, i.e. individual reproductive success, would be correlated with MHC genotype and best parasite resistance. Under natural

breeding conditions, 1) if heterozygote advantage would be the primary force driving female MHC-based mate choice, the more diverse males should be preferred over the less diverse, regardless of the females own MHC genotype. 2) If mate choice would favor production of optimal offspring with regard to MHC, we would expect females to seek mates providing complementary and compatible set of genes for the offspring. 3) Alternatively, if mate choice would favor locally adapted genes against the most prevalent pathogens, we would expect females to seek mates carrying specific MHC alleles. Under these scenarios, we aimed to identify whether MHC genes could be regarded as good genes. Finally, because females may be choosier when the number of available mates is higher, we manipulated male density. With this setup, we aimed to simulate better the natural conditions where females might have access to a high number of potential mates, which are exposed to higher competition.

Material Methods

- Experimental fish

Young of the year three-spined sticklebacks (*Gasterosteus aculeatus*) were caught during the winter 2005 (15.12.05) in the lake Großer Plöner See (54° 9'21.16"N, 10°25'50.14"E, Germany). They were kept in the laboratory under winter light and temperature conditions (8/24 light, 6°C) until May 3, 2006 when they were brought to spring conditions (12/24 light, 12°C). On May 11, they were transferred to summer conditions (15/24 light, 18°C) before release on May 19 into enclosures located exactly in the area where the fish were caught (see for enclosure set up Kalbe *et al.* In Press). All fish were weighed, measured and spine clipped before release. The enclosures were filled up with 6 randomly allocated females and either 6 (sex ratio 1:1 in 3 enclosures) or 12 (sex-ratio 2:1, in 3 enclosures) randomly caught males. These enclosures have previously been shown to provide the conditions needed for natural mating behavior and acquiring the natural range of sympatric parasites but exclude predation pressure (Kalbe *et al.* In Press).

- Egg collection

To estimate reproductive success and mate choice, once week, all stickleback nests were detected by careful observations. All egg batches were collected and the nests carefully replaced in their original location. Egg batches were brought to the laboratory. Within a nest, clutches were separated on the basis of different developmental stages. When necessary, individual egg clutches were incubated in aerated well water (with 0.04 ppm malachite green) at 18°C, until dark eye spots and the neural tube developed, to ensure a sufficient amount of DNA for further analysis. DNA from 24 randomly picked eggs from each egg clutch was automatically extracted (Freedom evo robot, Tecan, Switzerland) using Invisorb® DNA Tissue HTS 96 kits (Invitek, Germany).

- Male breeding coloration estimation and dissection

After 4 weeks in the enclosures, the experiment was stopped since the number of clutches collected decreased markedly (protocol as in Kalbe *et al.* In press). During re-capture, males were caught above the nest they

maintained. This allowed us to determine whether they sired all, a proportion, or none of the eggs contained in the nest on that collection day. This was then used to infer egg categories (see Material & Methods, section "parenthood analysis"). The males' red throats were photographed in the field immediately after capture. The photographs were taken within a dark box, with a fixed light source, using a digital camera (Olympus E20-p) with a 36 mm macro-lens. For the camera parameters and the handling of fish, see Jäger et al. (2007). IP Lab 3.6.2 for Mac OS 9.2.2 was used to estimate the red coloration intensity from a fixed area of the throat. Jäger et al (2007) showed this method to be highly repeatable. The recaptured fish were measured, weighed and dissected within three days post re-capture. Dissection took place under a dissection microscope to determine external and internal macroparasites as well as ciliates harbored by the fish, to the lowest taxonomic level possible (Kalbe *et al.* 2002). In order to quantitatively compare parasite burden between fish, a parasite index was calculated summarizing different combinations of rare and frequently occurring parasites (Kalbe *et al.* 2002).

- DNA extraction, microsatellites and MHC typing

DNA extractions, from dorsal spines, were performed using Invisorb® DNA Tissue kits (Invitek, Germany) following the manufacturer's protocol. All the fish were typed for 9 microsatellites multiplexed in 2 different PCR protocols both before release in the enclosures, and after recapture for identification (Kalbe *et al.* In Press). The number of microsatellites was sufficient to guarantee a high parenthood resolution. The MHC class II*B* diversity was then determined by reference strand conformation analysis (RSCA) according to Lenz et al. (Submitted), targeting the exon II of MHC genes, which encodes the highly variable peptide binding groove of the beta chain of the MHC molecules.

- Parenthood analysis

Paternity assignments were performed for each representative egg with the software PAPA (Duchesne *et al.* 2002). According to parenthood analysis, each egg was assigned to one of the following categories: nest owner, sneaker and stolen (Kalbe *et al.* In Press). Males that fertilized the majority of the eggs in a nest were categorized as nest owners and these eggs were assumed to originate from females mate choice. This was verified, for the last day, with the data from recapture where the males were caught maintaining their nests. Eggs were assigned to the sneaker category when, within a nest, the same female had eggs fertilized by the nest owner and another male. Eggs were considered as stolen from other nests if they were not fertilized by the nest owner and no further eggs from the same female were found in the nest (Kalbe *et al.* In Press).

- Data analyses
    - female mate choice

In order to test whether females use MHC self–reference under natural conditions to chose their mates, MHC similarity between all females and all males from the same enclosure, was calculated as MHC variant-sharing values: the proportion of shared MHC variants in a pair is twice the number of the variants shared by two individuals, divided by the sum of the variants of

each individual [$D=2F_{ab}/(F_a+F_b)$] (Wetton *et al.* 1987). For all females, taking into account the actual number of mates they had, in computer simulations we let them choose i) randomly between the males in their enclosures, ii) the most MHC dissimilar, and iii) the most MHC similar. We first focused on males we classified as having been chosen. We compared the observed distribution to the calculated ones. The first null hypothesis was that females selected mates randomly with respect to MHC variant-sharing. The second null hypothesis tested whether females aimed at increasing MHC diversity by seeking dissimilar males (heterozygote advantage). Finally, the third null hypothesis was that females favored MHC similar males to mate with (local adaptation). The randomization was run 1000 times per enclosure to generate null distributions. A Mann-Whitney test was used to compare the observed shared variant values with the randomly obtained values. Since all the analyses are performed enclosure-wise, we combined the results from the 6 enclosures using a binomial test. Then a kolmogorov-Smirnov test was used to compare observed and expected distribution of variant sharing between mates. The same analyses were performed for males that obtained fertilizations only through sneaking behavior.

- Reproductive success, parasite load and male red breeding coloration

The statistical analyses were performed with JMP (5.0.1a, SAS Institute Inc.), R statistical package (v.2.5.0) and Primer v6 (Clarke& Gorley 2006). Normality and homoscedasticity were checked and statistics were conducted accordingly. If needed, variables were transformed to meet test assumptions. Variable co-linearities were removed using residuals. Moreover, because enclosure and male density effects were strongly correlated (Pearson Chisq=18.422, p=0.003), only the male density variable corrected for the enclosure variable was included in the analysis. Since male density was our only treatment, we included all possible double interactions with that predictor in the models. Model selections were performed stepwise and based on AIC criteria (Sakamoto *et al.* 1986).

We investigated the probabilities of breeding over the entire reproductive period using sex, male density, individual MHC diversity, standard length before the experiment, male density and all two-way interactions with male density as predictors in a binomial logistic regression. The dependent variable was coded as 1 for individual which reproduced and 0 for thus which did not. An analysis of covariance (ANCOVA) was used to test whether individual reproductive success (square-root transformed total number of eggs) was related to predictors such as sex, male density, parasite burden, standard length before experiment, individual MHC diversity and all two-way interactions with male density. If analyses were sex-specific, we added the red throat coloration as new male predictor.

In another ANCOVA, individual parasite load (log transformed) was investigated in relation to sex, male density, standard length, individual MHC diversity and all two-way interactions with male density.

For each male, we calculated the arc-sin transformed proportion of sneaked eggs over the total number of eggs in nest that were fertilized by each male. An ANCOVA was then performed with reproductive success, standard length, male breeding coloration, individual MHC diversity, male density all two-way interactions with male density as predictors.

Eventually another analysis of covariance was used to test for variation in male red throat coloration. Individual MHC diversity, male density, parasite load, initial standard length and all interaction with male density were used as predictors.

Results

- Recapture, parenthood analysis and MHC.

After four weeks, 79 individuals out of 90 were recaptured (3 were found dead during the experiment) and 162 clutches were collected. Out of 3888 analyzed eggs, 32 (less than 1%) could not be assigned unambiguously to two parents and were therefore removed from further analyses. We found no deviation from Hardy-Weinberg equilibrium at any of the microsatellites in the experimental fish (All $p>0.21$).
Transformations of the variable "individual MHC diversity" did not improve the data's approximation to normal distribution. There was no relationship between the individual MHC diversity and sex (median=4, for both sexes. Mann-Whitney test, $U= 0.992$, $p=0.321$). The number of variants ranged from 1 to 5 between individuals and the frequencies of the variants varied from 2 to 47%. There was no significant difference between enclosures, neither in MHC variant numbers (kruskall-wallis, $K=3.384$, $d.f.=5$, $p=0.641$), nor in MHC variants constitution (ANOSIM on Bray-Curtis similarity matrix, Global $R=-0.039$, $p=0.972$).

- Female mate choice

Based on 9 microsatellites, females mated randomly with regard to genetic background (coefficient of relatedness from Queller & Goodnight, $r=-0.0119$, after 1000 bootstraps, $r=-0.0131$, $p>0.05$).

The probability of breeding during the experiment was positively related to fish length (binomial logistic regression, $X^2=5.797$, $p=0.016$), but not to individual MHC diversity (binomial logistic regression, $X^2=0.718$, $p=0.397$).

In each enclosure, females did not mate randomly with regard to MHC variant-sharing (Median observed versus random distributions, Mann-Withney tests for 6 enclosures, all $p<0.0001$. Binomial test for 6 enclosures, $p=0.0313$. Kolmorogov-Smirnov, observed vs. random distributions, all $D>0.228$, all $p<0.0001$).The observed mating distribution was significantly different from the expected distribution if females would search for an MHC dissimilar partner (Median observed versus maximum distributions, Mann-Withney tests, all $p<0.0001$, binomial test for 6 enclosures, $p=0.0313$. Kolmorogov-Smirnov, Observed vs. maximum distributions, all $D>0.167$, all $p<0.0001$). Finally, females did not aim either to minimize MHC variant-sharing since the observed distribution was significantly different from the expected distribution

in the case of seeking for MHC similar males (Observed versus minimum distributions, Mann-Whitney test, all $p<0.02$, binomial test for 6 enclosures, $p=0.0313$. Kolmorogov-Smirnov, Observed vs. minimum distributions, all $D>0.113$, all $p<0.0001$).

When investigating the number of MHC variants shared between females and males that gained fertilization by sneaking behavior, we could not detect any difference from random (Mann Withney tests, 3 enclosures $p>0.05$, binomial test $p=1$).

- Individual reproductive success and fitness related traits

  - reproductive success

An ANCOVA revealed that larger fish, when corrected for sex variation, had a higher reproductive success than smaller ones ($F_{1,78}= 19.462$, $p<0.0001$. Fig 1). Moreover, due to higher number of males reproductive success was strongly biased towards females ($F_{1,78}=12.238$, $p<0.0001$). Therefore we conducted the same analysis for each sex separately (Table 1). Female reproductive success was higher in enclosures with more males ($F_{1,30}=4.336$, $p=0.047$) and marginally with increased standard length ($F_{1,30}=3.168$, $p=0.086$). However, as parasite load increased, reproductive success decreased ($F_{1,30}=7.458$, $p=0.011$. Fig 2) emphasizing the costly role of parasites (Table 1).
Contrary to females, male reproductive success was negatively affected by male density ($F_{1,46}=18.105$, $p=0.0001$). Nevertheless, male reproductive success positively correlated with standard length ($F_{1,46}=24.224$, $p<0.0001$) and negatively with parasite burden ($F_{1,46}=4.489$, $p=0.0399$). Surprizingly, the reproductive success was also negatively correlated with the male red coloration ($F_{1,46}=6.172$, $p=0.0170$; Table 1)

  - parasite load

Females harbored a higher individual parasite load than males ($F_{1,78}=10.747$, $p=0.0016$) and for both sexes, the parasite burden was negatively correlated to individual MHC diversity ($F_{1,78}=4.730$, $p=0.0328$. Table 2a, Fig. 3).

Since reproductive success was related to both standard length and parasite load, we investigated whether bigger and smaller fish harbored the same parasite communities. To assess this, we arbitrarily created three size groups for each sex: small, medium and large. An analysis of similarity on Bray-Curtis similarity matrix (ANOSIM) revealed that large and small males harbored different parasites (Males Global $R= 0.082$, $p=0.026$. Females Global $R=0.007$, $p=0.32$). A similarity percentage-species contributions test (SIMPER) revealed that the tested fish groups were dissimilar for 43.14% and that 35.41% of that difference was explained by a lower parasitization by the monogenan *Gyrodactulys sp.* and the cestode *Valipora campylancristrota* in the bigger fish. Such relationship was detected only in males.

We used the same statistical approach to investigate whether larger males had different MHC variants than smaller ones. And, indeed, the MHC genotypes of big males were significantly different from the MHC genotypes of

smaller males (Global R=0.091, p=0.014). The average dissimilarity was 73.10% and was mainly due to the presence of the haplotype F10 (43.84%) in the bigger fish.

The standard length was higher for males with the MHC haplotype F10, and larger males harbored a lower burden of *Gyrodactylus* sp. and *Valipora campylancristrota*. We subsequently investigated whether these two parasites decreased when fish carried F10. Males with the MHC haplotype F10 (corrected for length) had a lower parasite burden in *Gyrodactylus sp.* than those without F10 (ANOSIM, Global R=0.084, p=0.047. Fig. 4). However, there was no difference for the cestode *Valipora campylancristrota* (ANOSIM, Global R=0.021, p=0.329). Note, the use of a permutation based test allows for no post-hoc correction.

Gyrodactylus sp. was one of the most prominent parasites. In total, its prevalence reached 67% and was consistent over all enclosures ($F_{5,79}$=1.024, p=0.409) and between sexes ($F_{1,79}$=2.884, p=0.093). Parasitized fish harbored from 1 to 244 *Gyrodactylus* sp. with a mean infection intensity of 13 individual parasites.

We investigated whether the haplotype F10 was advantageous over a broader array of parasites by replacing individual MHC diversity of our previous model, by the presence/absence of F10 (Table 2b). Here as well, male density and sex affected significantly the parasite load, but we did not detect any effect of F10 ($F_{1,78}$=0.003, p=0.955).

- re –analyses of female mate choice and reproductive success

Since we found a correlation with a specific MHC haplotype, we investigated whether the probability of mating was a function of the presence or absence of that haplotype. We re-performed the same binomial logistic regression using F10 as a new variable (presence or absence of the haplotype). Again, we found that for both sexes the likelihood of mating during the reproductive period was positively correlated with fish length (X^2=5.797, p=0.016). Furthermore, fish with the haplotype F10 had a significantly higher likelihood of reproducing (X^2=5.550, p=0.019). Splitting the data set for sex showed that all females bared the haplotype F10, whereas 42.9% of the males harbored it.

Parasite burden, standard length of the fish and the probability of mating were dependent on whether the fish harbored the haplotype F10. These results led us to re-conduct the initial test investigating the reproductive success. Since all females displayed the haplotype F10, a female specific analysis was not possible and therefore our tests included all the fish (Table 3a) and only males (Table 3b). Interestingly, we found a positive role for both fish size and haplotype F10 on the reproductive output of the fish (ANCOVA, Size, $F_{1,78}$= 24.259, p<0.0001 ; Haplotype F10, $F_{1,78}$= 4.596, p=0.035, Fig. 5).

- Male mating behavior and breeding coloration

Performing the parenthood analysis for every typed egg, we assigned 9.8% of them to male sneaking behavior which is comparable to findings in natural populations (Blais *et al.* 2004; Largiadèr *et al.* 2001; Rico *et al.* 1992). The percentage of eggs fertilized in the nest of another male was negatively correlated with reproductive success (ANCOVA, $F_{1,46}$= 5.558, p= 0.022). Moreover, small fish achieved most of their fertilizations by sneaking behavior relative to bigger fish (ANCOVA, $F_{1,46}$=4.692, p=0.035). We also found that male density enhanced sneaking behavior (ANCOVA, $F_{1,46}$=-2.587, p=0.013). Although not significant at 5%, males lacking the MHC haplotype F10 showed an increased tendency to gain fertilizations through sneaking (Welch t test, t=1.901, p=0.065).

Finally, we investigated whether the male breeding coloration was related to MHC individual diversity, male density, parasite load, standard length and all double interaction with male density. Only MHC individual diversity was significantly negatively correlated to redness ($F_{1,46}$=4.353, p=0.043. Table 4). Including F10 into the model did not reveal any significant effect (Table 4).

Discussion

We investigated, under semi-natural conditions, female stickleback mate choice based on MHC genes in relation to individual reproductive success, motivated by a previous experiment which used lab-bred fish (Kalbe *et al.* In Press). Here, the fish were wild-caught and exposed to their natural sympatric parasites. Using 9 microsatellites we were able to determine both parents for almost 4000 stickleback eggs. We found that female mate choice was not random with regard to MHC. Females significantly chose males that were neither MHC similar nor MHC dissimilar. We therefore conclude that females preferred mates with which they shared an intermediate individual MHC diversity. This result is consistent with previous olfactory (flow-channel) mate choice experiments where female sticklebacks preferred males offering a complementary set of MHC variants to potentially achieve an intermediate individual MHC diversity in the offspring (Aeschlimann *et al.* 2003; Milinski *et al.* 2005; Reusch *et al.* 2001). Here, we show that when females are confronted with a mating decision under natural conditions, with a number of males available, they follow our prediction for optimal immunogenetic make up for their offspring.

Various studies have suggested female mate choice for an intermediate individual MHC diversity (Aeschlimann *et al.* 2003; Bonneaud *et al.* 2006a; Forsberg *et al.* 2007; Jacob *et al.* 2002; Milinski *et al.* 2005; Olsson *et al.* 2003). Such mate choice strategy is relevant when individuals with an intermediate MHC diversity are optimal with regard to parasite resistance, individual body condition, reproductive success and survival. Recently, there has been increasing evidence revealing the advantages for individuals with such MHC diversity (Aeschlimann *et al.* 2003; Bonneaud *et al.* 2005; Forsberg *et al.* 2007; Kalbe *et al.* In Press; Madsen& Ujvari 2006; Milinski *et al.* 2005; Reusch *et al.* 2001; Wegner *et al.* 2003a; Wegner *et al.* 2008; Wegner *et al.* 2003b).

Mate choice for good genes could arise if a specific MHC variant or haplotype confers a strong benefit against a common and virulent pathogen. Accordingly, our results showed that the probability of a male to be chosen by females was dependent on whether or not they bared a specific MHC haplotype. Consequently, males with this haplotype achieved a higher reproductive success. We show here that this specific MHC haplotype was conferring resistance against one of the most common parasistes- *Gyrodactylus* sp. All the surviving females carried this specific haplotype. Females displayed a higher parasite load than males. This could be a consequence of the need for high nutrient input for egg production achieved through intense feeding (Kalbe *et al.* In Press). Foraging behavior may drive exposure to an increased range of parasites, both trophically, and actively transmitted and lead to selection against weaker females. We have not determined *Gyrodactylus sp.* to the species level. However, a previous survey revealed that the majority of *Gyrodactylus* sp. found in sticklebacks from lakes in our area are *G. gasterostei* (Raeymaekers *et al.* 2006b), and it is known that *Gyrodactylus* species are relatively species-specific compared to other parasites (Raeymaekers *et al.* 2006a). Therefore, this parasite is one of the parasites of our system for which a host-parasite co-evolution based on genotype-genotype interactions can be expected. The three-spined stickleback is the only host of this parasite on which it feeds and reproduces. Several studies have reported the impact of *Gyrodactylus* species on host immune system activation (Buchmann& Lindenstrøm 2002; Collins *et al.* 2007; Lindenstrom *et al.* 2004). *Gyrodactylus* infection has been correlated with carotenoids input- the source of various male secondary sexual trait colorations (Kolluru *et al.* 2006). Such correlation between a specific MHC haplotype and a virulent parasite is a pre-requisite to negative-frequency-dependent selection: one of the mechanisms invoked to explain the maintenance of MHC polymorphism.

Standard length of the fish was a strong determinant of reproductive success. This result supports the idea of adaptation to environmental selection regimes in sticklebacks (McKinnon *et al.* 2004; McPhail 1969). The size of an individual stickleback correlated with the possession of a specific MHC haplotype. Thus, we predict the following scenario for mate choice in stickleback; based on olfactory cues females approach males' nests, and based on visual cues inspect their condition (size, redness) before deciding whether or not to invest in reproduction with these males. Consistent with our prediction, where parasites affect the ability of individuals to reproduce, females should look for males carrying a specific resistant MHC haplotype.

Mainly two propositions have been put forward to explain why females should base their mating decision on MHC genes: to avoid inbreeding and/or to complement a compatible set of alleles (reviewed in Kempenaers 2007; reviewed in Milinski 2006). Here, we can eliminate the possibility that our observations resulted from inbreeding avoidance since females mated randomly with regard to genetic background, potentially reflecting the lack of inbreeding risks in this large population. Interestingly, results suggest both, the compatible genes hypothesis as seen in the intermediate number of shared alleles between mates, and the good genes hypothesis as seen in the

advantages of the males that carried the haplotypes F10. In fact, these two hypotheses might not be mutually exclusive (Colegrave *et al.* 2002). With this strategy, females achieve a two fold immunogenetic advantage: resistance against a wide range of parasites while keeping the T-Cell repertoire depletion to a minimum level (Nowak *et al.* 1992; Woelfing *et al.* 2009) and resistance against a currently common and costly parasite.

Such mating strategy is relevant since both additive effects and non additive effects of MHC genes are not mutually exclusive and here both provided increased resistance to parasites. Conceivably, as the prevalence of various parasites and pathogens changes in the environment, so will the survivorship value of the associated MHC variant. This could lead to negative-frequency-dependent selection and the fluctuation of allele frequencies through evolutionary time - mechanism thought to maintain MHC polymorphism. However, female mating strategy is not the only source of maintenance of MHC polymorphism.

As expected, male density increased female choosiness and increased the proportion of sneaking events. Sneaking behavior enabled lower quality males to circumvent mate choice and transmit genes to the next the generation maintaining MHC polymorphism by creating sexual conflicts. Male density affected as well positively female reproductive success. During, their paternal cares, males are less receptive to females (Wootton 1976). Increasing male density increased the likelihood that at least a male was receptive when females were ripe and ready to spawn. This was reflected by 25% more nests present in high density enclosures compared to low density enclosure (76 vs. 57 nests detected within 4 weeks).

Males with a higher MHC diversity also displayed a duller breeding coloration and redness was associated with a lower reproductive success. We have previously demonstrated that red breeding coloration, which is estimated at the end of the experiment, lacks a timing component (Kalbe *et al.* In Press). We show here that when we estimated male's breeding coloration, the redder males were the low quality males: those with low reproductive success, high parasitization and low MHC diversity. We proposed that these males followed a final investment strategy. To test our assumption, we compared the proportion of egg fertilized during the last week of the experiment over the entire reproductive period between the redder and the duller males. Redder males obtained 23% of their total reproductive success during that period, whilst duller males obtained only 14%. Although it is an honest signal, the red breeding coloration might lead to misinterpretations of results if considered representative of a constant state of the fish.

The idea that mate choice might be driven by genetic compatibility and/or good genes has been suggested (Byers& Waits 2006; Grob *et al.* 1998; Landry *et al.* 2001; Mays& Hill 2004; Milinski 2006; Neff& Pitcher 2005; Penn& Potts 1999). Previous studies often have not focused on fitness related genes and/or mate choice translated into Darwinian fitness. Our results represent the first evidence for a MHC haplotype, conferring resistance

against a virulent parasite, that is directly translated into fish size and ultimately into Darwinian fitness.

As a perspective for further studies we propose that parasite/host MHC interaction, which directly translates into Darwinian fitness, could have implications on speciation. Females favored males with which they shared an intermediate MHC diversity. Such strategy could help to avoid disruption of local adaptations of co-adapted gene complexes (Hendry *et al.* 2000) and production of low quality offspring (Eizaguirre *et al.* 2009). Often assortative mating is thought to lead to purifying selection (Kirkpatrick & Nuismer 2004). However, this might not apply to MHC genes since allele frequencies vary rapidly between years responding to fluctuating parasite pressures as negative-frequency-dependent selection would envisage it (e.g. Charbonnel& Pemberton 2005; Westerdahl *et al.* 2004).

Thus, in this semi-natural field experiment, fish were exposed to the natural spectrum of parasites in their natal lake. We showed that female sticklebacks used an MHC-based mating strategy to transmit good genes. Using MHC signaling they can look for males in good condition providing the best paternal care and providing their offspring with a higher relative immunity in their pathogenic environment. Interestingly, assortative mating for MHC-adapted partners could influence our thinking in terms of understanding the rapid evolution of species.

Acknowledgments

We thank G. Augustin, D. Martens and H. Deiwick for technical assistance with the enclosure system. Furthermore, we are grateful to I. Schultz for inestimable lab-assistance. We also thank Kerstin Brzezek for daily help and Ralf D. Sommerfeld and Joel Meunier for fruitful discussions. In addition, we thank the county of Schleswig-Holstein and the city of Ploen for permitting us to run the outdoor experiments in the lake. All animal experiments described were approved by the Ministry of Nature, Environment and Country Development, Schleswig-Holstein, Germany.

References

Aeschlimann PB, Häberli MA, Reusch TBH, Boehm T, Milinski M (2003) Female sticklebacks Gasterosteus aculeatus use self-reference to optimize MHC allele number during mate selection. *Behavioral Ecology and Sociobiology* **54**, 119-126.

Andersson M (1994) *Sexual selection* Princeton University Press, Princeton.

Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997) The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology* **17**, 179-224.

Bakker TCM, Mundwiler B (1994) Female mate choice and male red coloration in a natural three-spined stickleback (Gasterosteus aculeatus) population. *Behavioral Ecology* **5**, 74-80.

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

Blais J, Rico C, Bernatchez L (2004) Nonlinear effects of female mate choice in wild threespine sticklebacks. *Evolution* **58**, 2498-2510.

Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G (2006a) Complex *MHC*-based mate choice in a wild passerine. *Proceedings of the Royal Society of London Series B: Biological Sciences* **273**, 1111-1116.

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006b) Major histocompatibilty alleles associated with local resistance to malaria in a passerine. *Evolution* **60**, 383-389.

Bonneaud C, Richard M, Faivre B, Westerdahl H, Sorci G (2005) An *MHC* Class I allele associated to the expression of T-dependent immune response in the house sparrow. *Immunogenetics* **57**, 782-789.

Briles WE, Briles RW, Taffs RE, Stone HA (1983) Resistance to a malignant lymphoma in chickens is mapped to subregion of major histocompatibility (B) complex. *Science* **219**, 977-979.

Brown JL, Eklund A (1994) Kin Recognition and the Major Histocompatibility Complex: An Integrative Review. *The American Naturalist* **143**, 435-461.

Buchholz R, Jones Dukes MD, Hecht S, Findley AM (2004) Investigating the turkey's 'snood' as a morphological marker of heritable disease resistance. *Journal of Animal Breeding and Genetics* **121**, 176-185.

Buchmann K, Lindenstrøm T (2002) Interactions between monogenean parasites and their fish hosts. *International Journal for Parasitology* **32**, 309-319.

Byers JA, Waits L (2006) Good genes sexual selection in nature. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16343-16345.

Charbonnel N, Pemberton J (2005) A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity* **95**, 377-388.

Clarke B, Kirby DRS (1966) Maintenance of Histocompatibility Polymorphisms. *Nature* **211**, 999-1000.

Clarke KR, Gorley RN (2006) *Primer v6: User Manual/Tutorial* PRIMER-E, Plymouth, UK.

Colegrave N, Kotiaho JS, Tomkins JL (2002) Mate choice or polyandry: reconciling genetic compatibility and good genes sexual selection. *Evolutionary Ecology Research* **4**, 911-917.

Collins CM, Olstad K, Sterud E*, et al.* (2007) Isolation of a novel fish thymidylate kinase gene, upregulated in Atlantic salmon (Salmo salar

L.) following infection with the monogenean parasite Gyrodactylus salaris. *Fish & Shellfish Immunology* **23**, 793-807.

Doherty PC, Zinkernagel RM (1975) A biological role for the major histocompatibility antigens. *Lancet.* **305**, 1406-1409.

Duchesne P, Godbout MH, Bernatchez L (2002) PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. *Molecular Ecology Notes* **2**, 191-193.

Edwards SV, Hedrick PW (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends in Ecology & Evolution* **13**, 305-311.

Egid K, Brown JL (1989) The major histocompatibility complex and female mating preferences in mice. *Animal Behaviour* **38**, 548.

Eizaguirre C, Lenz TL, Traulsen A, Milinski M (2009) Speciation accelerated and stabilized by pleiotropic major histocompatibility complex immunogenes. *Ecology Letters* **12**, 5-12.

Ekblom R, Saether SA, Jacobsson PAR*, et al.* (2007) Spatial pattern of MHC class II variation in the great snipe (Gallinago media). *Molecular Ecology* **16**, 1439-1451.

Forsberg LA, Dannewitz J, Petersson E, Grahn M (2007) Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout - females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology* **20**, 1859-1869.

Godot V, Harraga S, Beurton I*, et al.* (2000) Resistance/susceptibility to *Echinococcus multilocularis* infection and cytokine profile in humans. II. Influence of the HLA B8, DR3, DQ2 haplotype. *Clinical and Experimental Immunology* **121**, 491-498.

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

Grob B, Knapp LA, Martin RD, Anzenberger G (1998) The Major Histocompatibility Complex and Mate Choice: Inbreeding Avoidance and Selection of Good Genes. *Experimental and Clinical Immunogenetics* **15**, 119.

Hamilton WD, Zuk M (1982) Heritable true fitness and bright birds: a role for parasites? *Science* **218**, 384-387.

Harf R, Sommer S (2005) Association between MHC classII DRB alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paebe*, in the southern Kalahari. *Molecular Ecology* **14**, 85-91.

Hedrick PW (2002) Pathogen Resistance and Genetic Variation at MHC loci. *Evolution* **56**, 1902-1908.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: Evidence from introduced salmon. *Science* **290**, 516-518.

Hill AVS, Allsopp CEM, Kwiatkowski D*, et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature* **352**, 595-600.

Horton R, Wilming L, Rand V*, et al.* (2004) Gene map of the extended human MHC. *Nature Reviews Genetics* **5**, 889-899.

Jacob S, McClintock MK, Zelano B, Ober C (2002) Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature genetics* **30**, 175-179.

Jäger I, Eizaguirre C, Griffiths SW*, et al.* (2007) Individual MHC class I and MHC class IIB diversities are associated with male and female reproductive traits in the three-spined stickleback. *Journal of Evolutionary Biology* **20**, 2005-2015.

Janeway CA, Travers P, Walport M, Shlomchik MJ (1999) *Immunobiology: The immune system in health and diseases.* Garland Science publishing, New York.

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

Kalbe M, Eizaguirre C, Dankert I*, et al.* (In Press) Lifetime reproductive success is maximized with optimal MHC diversity. *Proceedings of the Royal Society of London. Series B, Biological Sciences*.

Kalbe M, Wegner KM, Reusch TBH (2002) Dispersion patterns of parasites in 0+ year three-spined sticklebacks: a cross population comparison. *Journal of Fish Biology* **60**, 1529-1542.

Kempenaers B (2007) Mate choice and genetic quality: A review of the heterozygosity theory. In: *Advances In The Study Of Behavior, Vol 37*, pp. 189-278. Elsevier Academic Press Inc, San Diego.

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

Kolluru GR, Grether GF, South SH*, et al.* (2006) The effects of carotenoid and food availability on resistance to a naturally occurring parasite (Gyrodactylus turnbulli) in guppies (Poecilia reticulata). *Biological Journal of the Linnean Society* **89**, 301-309.

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

Landry C, Garant D, Duchesne P, Bernatchez L (2001) 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proceedings of the Royal Society of London Series B: Biological Sciences* **268**, 1279.

Langefors Å, Lohm J, Grahn M, Andersen Ø, 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.

Largiadèr CR, Fries V, Bakker TCM (2001) Genetic analysis of sneaking and egg-thievery in a natural population of the three-spined stickleback (*Gasterosteus aculeatus L.*). *Heredity* **86**, 459-468.

Lawlor DA, Zemmour J, Ennis PD, Parham P (1990) Evolution of class-I MHC genes and proteins - from natural-selection to thymic Selection. *Annual Review of Immunology* **8**, 23-63.

Lindenstrom T, Secombes CJ, Buchmann K (2004) Expression of immune response genes in rainbow trout skin induced by Gyrodactylus derjavini infections. *Veterinary Immunology and Immunopathology* **97**, 137-148.

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

Madsen T, Ujvari B (2006) MHC class I variation associates with parasite resistance and longevity in tropical pythons. *Journal of Evolutionary Biology* **19**, 1973-1978.

Mays HLJ, Hill GE (2004) Choosing mates: good genes versus genes that are a good fit. *Trends in Ecology and Evolution* **29**, 564-570.

McKinnon JS, Mori S, Blackman BK*, et al.* (2004) Evidence for ecology's role in speciation. *Nature* **429**, 294-298.

McPhail JD (1969) Predation and Evolution of a Stickleback (Gasterosteus). *Journal Of The Fisheries Research Board Of Canada* **26**, 3183-3208.

Milinski M (2006) The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. *Annual Review of Ecology, Evolution, and Systematic* **37**, 159-186.

Milinski M, Bakker TCM (1990) Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. *Nature* **344**, 330-333.

Milinski M, Griffiths S, Wegner KM*, et al.* (2005) Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4414-4418.

Neff BD, Pitcher TE (2005) Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology* **14**, 19-38.

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York.

Ninomiya K, Brown RE (1995) Removal of the preputial glands alters the individual odors of male MHC-congenic mice and the preferences of females for these odors. *Physiology & Behavior* **58**, 191-194.

Nowak MA, Tarczyhornoch K, Austyn JM (1992) The optimal number of major histocompatibility complex-molecules in an individual. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10896-10899.

Olsson M, Madsen T, Nordby J*, et al.* (2003) Major histocompatibility complex and mate choice in sand lizards. *Proceedings of the Royal Society of London B: Biological Sciences* **270**, S254-S256.

Olsson M, Madsen T, Wapstra E*, et al.* (2005) MHC, health, color, and reproductive success in sand lizards. *Behavioral Ecology and Sociobiology* **V58**, 289-294.

Parham P, Ohta T (1996) Population Biology of Antigen Presentation by MHC Class I Molecules

10.1126/science.272.5258.67. *Science* **272**, 67-74.

Paterson S, Wilson K, Pemberton JM (1998) Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries L.*). *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3714-3719.

Penn D, Potts WK (1998) MHC-disassortative mating preferences reversed by cross-fostering. *Proceedings of the Royal Society of London Series B: Biological Sciences* **265**, 1299-1306.

Penn DJ, Damjanovich K, Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 11260-11264.

Penn DJ, Potts WK (1999) The evolution of mating preferences and major histocompatibility complex genes. *American Naturalist* **153**, 145-164.

Potts WK, Manning CJ, Wakeland EK (1991) Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature* **352**, 619-621.

Raeymaekers JAM, Huyse T, Maelfait H, Hellemans B, Volckaert FAM (2006a) *The population and community structure of Gyrodactylus ectoparasites on coexisting stickleback species*, Leuven Katholieke Universiteit.

Raeymaekers JAM, Wegner KM, Huyse T, Volckaert FAM (2006b) *Infection dynamics of the monogenean parasite Gyrodactylus gasterostei on sympatric and allopatric host populations*, K.U.Leuven.

Reusch TBH, Haberli MA, Aeschlimann PB, Milinski M (2001) Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature* **414**, 300-302.

Rico C, Kuhnlein U, Fitzgerald GJ (1992) Male reproductive tactics in the threespine stickleback - an evaluation by DNA fingerprinting. *Molecular Ecology* **1**, 79-87.

Sakamoto Y, Ishiguro M, Kitagawa G (1986) *Akaike Information Criterion Statistics* D. Reidel Publing Company.

Schat KA, Taylor RLJ, Briles WE (1994) Resistance to Marek's disease in Chickens with Recombinant Haplotypes of the Major Histocompatibility (B) Complex. *Poultry Science* **73**, 502-508.

Schwensow N, Eberle M, Sommer S (2008) Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proceedings of the Royal Society of London Series B: Biological Sciences* **275**, 555-564.

Sommerfeld RD, Boehm T, Milinski M (In press) Desynchronising male and female reproductive seasonality: dynamics of male olfactory attractiveness in sticklebacks. *Ethology, Ecology and Evolution* **In press**.

Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**, 967-978.

Wedekind C, Seebeck T, Bettens F, Paepke AJ (1995) MHC-Dependent Mate Preferences in Humans. *Proceedings of the Royal Society of London Series B: Biological Sciences* **260**, 245-249.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M (2003a) Parasite Selection for Immunogenetic Optimality. *Science* **301**, 1343.

Wegner KM, Kalbe M, Milinski M, Reusch T (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 KM, Reusch TBH, Kalbe M (2003b) Multiple parasites are driving major histocompatibility comple polymorphism in the wild. *Journal of Evolutionary Biology* **16**, 224-232.

Westerdahl H (2004) No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology* **13**, 2465-2470.

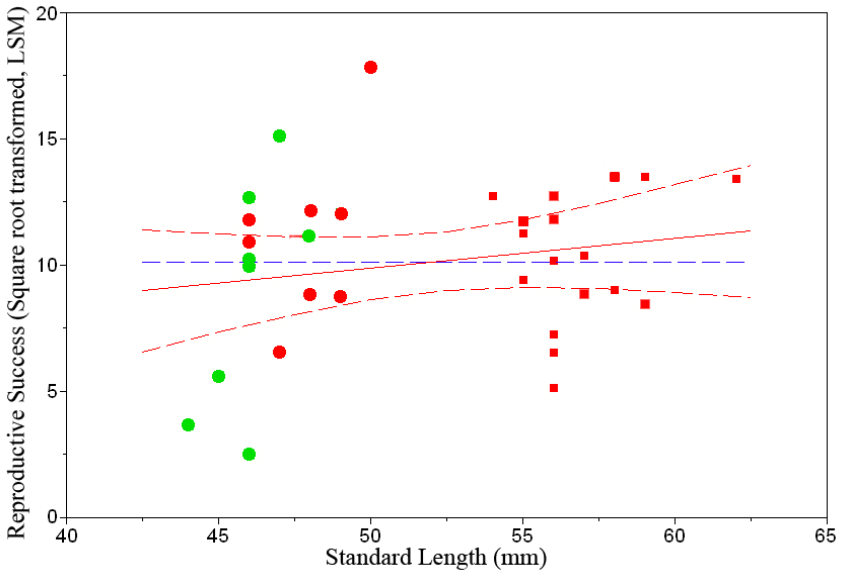
Westerdahl H, Hansson B, Bensch S, Hasselquist D (2004) Between-year variation of MHC allele frequencies in great reed warblers: selection or drift? *Journal of Evolutionary Biology* **17**, 485-492.

Wetton JH, Carter RE, Parkin DT, Walters D (1987) Demographic study of a wild house sparrow population by DNA fingerprinting. *Nature* **327**, 147-149.

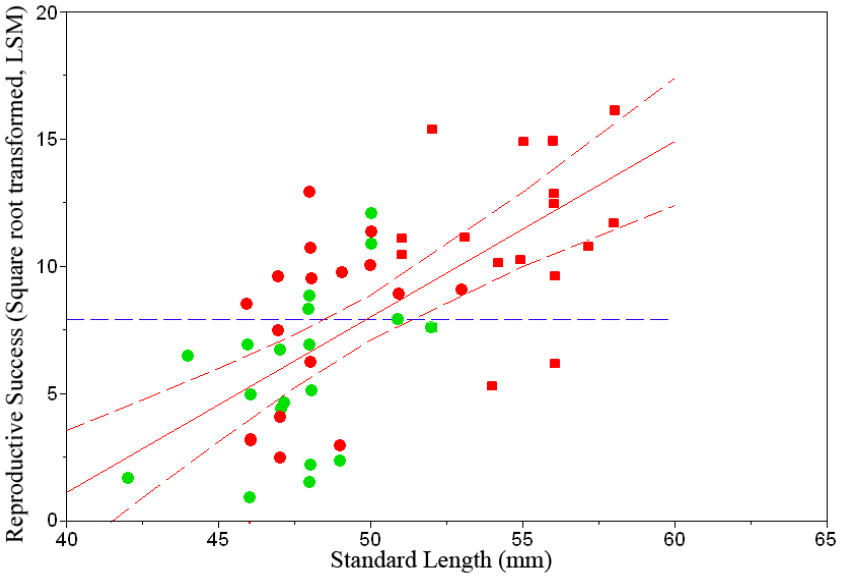
Woelfing B, Traulsen A, Milinski M, Boehm T (2009) Does intra-individual MHC diversity keep a golden mean? *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**.

Wootton RJ (1976) *The biology of the sticklebacks.* Academic Press, London.

Yamazaki K, Boyse EA, Mike V*, et al.* (1976) Control of mating preferences in mice by genes in the major histocompatibility complex. *Journal of Experimental Medecine* **144**, 1324-1335.



a)



b)

Figure 1: Figure showing the relationship between individual reproductive success as a function of standard length (Corrected for sex differences, $F=16.50$, $p<0.0001$). Blue dashed line show the mean, while the red dashed lines indicate the 95% confidence interval. The red line represents the regression line. Squares symbolize female, cycles symbolize males, red is used for individuals with F10, green for individuals without F10. a) In low density enclosures, b) in high density enclosures.

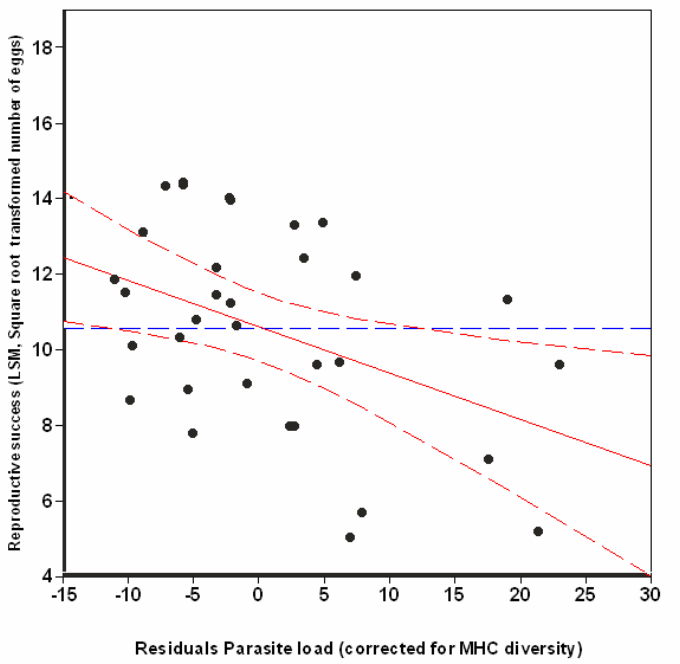


Figure 2: Figure representing individual reproductive success as a function of the parasite burden harbored by females (F=7.458, p=0.011). Blue dashed line shows the mean, the red dashed lines indicate the 95% confidence interval. The red line represents the regression line.

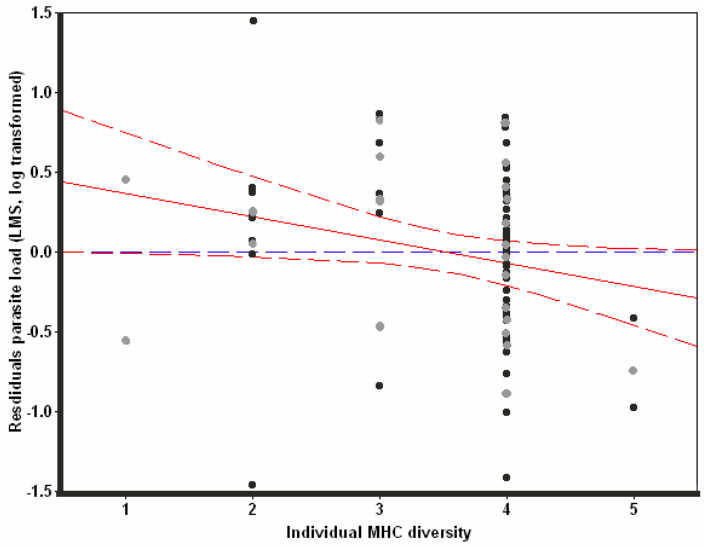


Figure 3: Graph showing the relationship between individual parasite burden and MHC intra-individual diversity (F=4.73, p=0.033). Blue dashed line shows the mean, while the red dashed lines indicate the 95% confidence interval. The red line represents the regression line. Black dots correspond to fish with F10, while grey dot symbolize fish lacking F10.

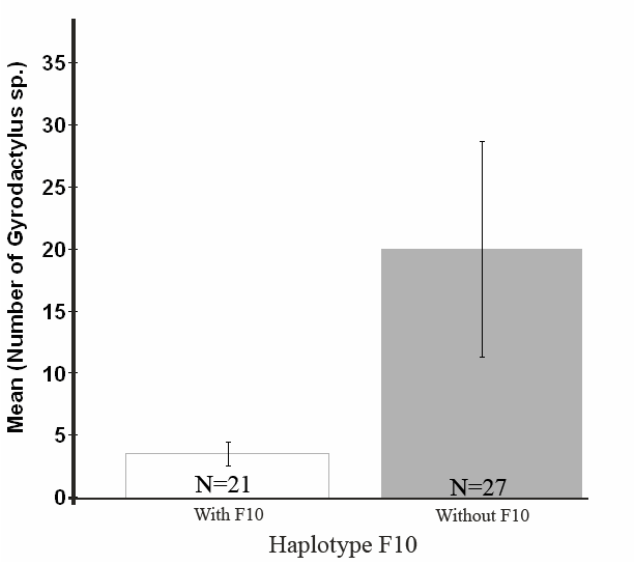


Figure 4: Relationship between the mean number of the parasite *Gyrodactylus sp.* and the occurrence of the haplotype F10 (standard error bars are depicted)

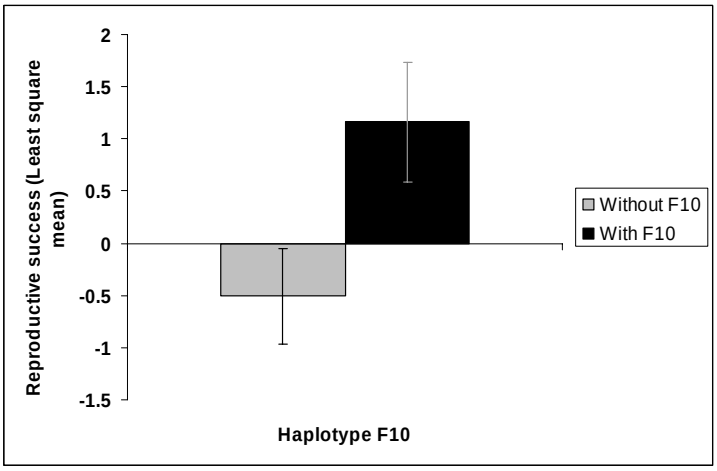


Figure 5: Individual reproductive success as least square means as a function of presence/absence of the MHC haplotype F10.

Table 1: Reproductive success of the three-spined stickleback over one reproductive period. Reproductive success is given as number of eggs square-root transformed to meet normality. a) gives the best predictors for females, and b) gives the best predictors for males. NA stands for not applicable and concerns males' secondary sexual characteristics only.

	Females			**Males**		
a) Variables	d.f	F ratio	p-value	d.f	F ratio	p-value
Male density	1	4.336	**0.0469**	1	18.105	**0.0001**
Standard length	1	3.168	0.0863	1	24.224	**<0.0001**
Parasite load [a]	1	7.458	**0.0110**	1	4.489	0.0399
MHC diversity	1	0.122	0.7293	1	0.092	0.7637
Redness [a]	NA	NA	NA	1	6.172	**0.0170**
Standard length *density	1	0.008	0.9290	1	1.006	0.3216
Parasite load [a]*density	1	0.031	0.8607	1	0.045	0.9563
Redness [a] * density	NA	NA	NA	1	0.477	0.4934

[a] corrected for MHC individual diversity

Table 2: Table summarizing results obtained from an analysis of covariance on log transformed individual parasite load a) with individual MHC diversity and b) with F10 as variable. NA stands for not applicable variable within the model.

	All fish/ Parasite load			**All fish/ Parasite load**		
a) Variables	d.f	F ratio	p-value	d.f	F ratio	p-value
Male density	1	4.883	**0.0301**	1	4.015	0.0487
Standard length [a,b]	1	0.521	0.4728	1	0.785	0.3786
Individual MHC diversity	1	4.730	**0.0328**	NA	NA	**NA**
F10	NA	NA	**NA**	1	0.003	0.9545
Sex	1	10.324	**0.0019**	1	9.364	0.0031
Standard length [a,b] *density	1	1.379	0.2582	1	1.135	0.3268
Individual MHC diversity *density	1	2.009	0.1412	NA	NA	NA
F10 *density	NA	NA	NA	1	0.279	0.7573
Sex*density	1	0.140	0.1043	1	0.289	0.5927

[a] corrected for F10; [b] corrected for sex

Table 3: Reproductive success of three-spined stickleback over one reproductive period. We re-conducted the analysis of covariance replacing MHC diversity with the variable "haplotype F10". Since all females displayed the haplotype F10, a female specific analysis was not possible to test for the effect of F10. All the fish were first included then only males. NA stands for not applicable.

	a) All fish			**b) Males**		
Variables	d.f	F ratio	p-value	d.f	F ratio	p-value
Male density	1	9.425	**0.0002**	1	4.047	**0.0504**
Standard length [a,b]	1	24.259	**<0.0001**	1	23.267	**<0.0001**
Parasite load [a,b]	1	0.132	0.7170	1	3.719	0.0606
Haplotype F10	1	4.596	**0.0353**	1	6.013	**0.0184**
Sex	1	17.251	**<0.0001**	NA	NA	NA
Redness	NA	NA	NA	1	5.542	**0.0233**
Standard length [a,b] *density	1	0.012	0.9144	1	0.526	0.4723
Parasite load [a,b]*density	1	1.760	0.1794	1	0.951	0.3352
Haplotype F10*density	2	0.003	0.9532	2	0.091	0.7640
Sex*density	2	17.463	**0.0001**	NA	NA	NA
Redness * density	NA	NA	NA	1	2.425	0.1271

[a] corrected for F10; [b] corrected for sex

Table 4: Table summarizing results obtained from an analysis of covariance on red intensity of the breeding colouration a) with individual MHC diversity and b) with F10 as variable. NA stands for not applicable variable within the model.

	All males/ Redness			**All males/ Redness**		
a) Variables	d.f	F ratio	p-value	d.f	F ratio	p-value
Male density	1	0.033	0.8572	1	0.068	0.7955
Standard length [a,b]	1	0.053	0.8191	1	0.332	0.5672
Parasite load [a,b,c]	1	0.650	0.4242	1	2.305	0.1358
Individual MHC diversity	1	4.353	**0.0425**	NA	NA	NA
F10	NA	NA	NA	1	0.120	
Standard length [a,b] *density	1	1.100	0.3593	1	1.095	0.3613
Parasite load [a,b,c]*density	1	0.247	0.8631	1	0.760	0.5527
Individual MHC diversity *density	1	0.058	0.9436	NA		
F10 *density	NA	NA	NA	1	0.346	0.7925

[a] corrected for F10; [b] corrected for sex; [c] corrected for individual MHC diversity

Acknowledgements

During the three years of my PhD work, I got to know many interesting people. I am very grateful for all the things I have experienced and learned from them. I would like to thank Manfred Milinski for giving me the opportunity to work in his group and to profit from the amazing possibilities connected to working at a Max Planck Institute. Due to his profound experience and supervision I greatly improved my methodological skills und broadened my scientific horizon.

Sven Becker led me through the first half of the PhD period and I thank him for discussions and constant support. I strongly enjoyed the collaboration with other temporary and permanent members of the group. I am grateful to Martin Kalbe, whose passion finally managed to raise my interest in the parasite point of view, for his patient and continuous support. I would also like to thank Dan Benesh, Sascha Hibbeler, Irene Samonte, Jörn Scharsack, Benno Wölfing and Sara Yeates for scientific and non-scientific discussions. The help of Gerd Augustin, Kerstin Brzezek, Regina Leipnitz, Sybille Liedtke, Daniel Martens, Ines Schultz, Gisella Schmiedeskamp and Monika Wulf in the lab and in the field enabled me to run projects that I would never have been able to handle alone. My thoughts go also to other current and former members of the institute, with whom I share one or the other memory: Anna, Arne, Chaitanya, Chris, Cristina, Dirk, Elke, Gisep, Ilka, Ilonka, Jan, Jochen, Kerstin, Katrin, Kathryn, Leslie, Philipp and Weini. Thank you all for the great time! The fluctuating basketball group made my Wednesday evenings!

The person and friend I spent most time with and probably also learnt most from during this period in scientific and non-scientific terms is Christophe Eizaguirre. I thank you for the unique experience and to work with you is simply *awesome*! I will always remember the time with Christophe and Ralf in and outside the institute, with a Döner at the lake or in front of the Wii.

I admire my parents for always supporting me and believing in me. You have a big part in where I am now.

The most important person however is Anja, who is always there for me and supports me whenever I need someone to hold on to.

Curriculum Vitae

PERSONAL DATA:

Name:	Tobias L. Lenz
Date and place of birth:	28.05.1977 in Hamburg
Nationality:	German

ADDRESS:
Bahrenfelderstraße 77
22765 Hamburg
Germany
Tel.: +49-40-39 36 25
Email: mail@tobiaslenz.de

EDUCATION:

2005 – 2008:	Graduate studies with the aim of achieving a PhD in Evolutionary Biology at the Max Planck Institute for Evolutionary Biology.
2005:	Research assistant at the University of Berne, Switzerland.
2005:	Research semester at the Harvard University in Boston, USA.
2001 – 2005:	Diplom in Zoology, Minor in Hydrobiology & Fishery Science, Nature Conservation and Computer Science, at the University of Hamburg.
1998 – 2001:	Vordiplom in Biology, University of Hamburg.
1997 – 1998:	Civil service, Schutzstation Wattenmeer e.V. on Amrum / Germany.
1990 – 1997:	Abitur, Emilie Wüstenfeld Gymnasium, Hamburg.
1989 – 1990:	Elementary school, Escuela de Pestalozzi in Ecuador / S. America.
1983 – 1989:	Elementary school, Freie Kinderschule Hamburg.

PUBLICATIONS:

Eizaguirre C., Lenz T. L., Traulsen A., Milinski M. 2009. Speciation accelerated and stabilized by pleiotropic major histocompatibility complex immunogenes. Ecology Letters 12: 5-12.

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.

Urbach, D., Bittner D., Lenz T. L., Bernet D., Wahli T., Wedekind C.. 2007. Sperm velocity in an Alpine whitefish: effects of age, size, condition, fluctuating asymmetry and gonad abnormalities. Journal of Fish Biology 71:672-683.

Lenz, T. L., Jacob A., Wedekind C. 2007. Manipulating sex ratio to increase population growth: the example of the Lesser Kestrel. Animal Conservation 10:236-244.

COMPETITIVE FUNDS:

2008:	Travel grant from Boehringer Ingelheim Fonds.
2005:	Fellowship for one semester at the Harvard University, USA from Studienstiftung des deutschen Volkes.
2002 – 2005:	Full scholarship from Studienstiftung des deutschen Volkes.

Erklärung

Hiermit versichere ich, dass diese Abhandlung, abgesehen von der Beratung durch meine adademischen Lehrer, nach Inhalt und Form meine eigene Arbeit ist, und dass ich keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Die Arbeit hat bisher weder ganz noch zum Teil an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegen.

Außerdem versichere ich, dass ich bisher noch kein anderes Promotionsverfahren begonnen habe.

Teile dieser Dissertation wurden bereits veröffentlicht oder zur Veröffentlichung eingereicht. Mein Beitrag und der Beitrag meiner Ko-Autoren ist am Anfang jedes Kapitels und in Tabelle O.1 detailliert aufgeführt.

(Tobias Lenz)