August 2019

Major histocompatibility complex does not affect mating patterns in bluegill sunfish (Lepomis macrochirus)

Lucas Silveira
The University of Western Ontario

Supervisor
Neff, Bryan D.
The University of Western Ontario

Graduate Program in Biology

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

Abstract

In many species, individuals prefer mates that are genetically dissimilar at the major histocompatibility complex (MHC). This is likely because it improves offspring resistance to pathogens. Here I provide the first genotypic characterization of the MHC class II peptide binding region in bluegill (*Lepomis macrochirus*), a species of sunfish, and examine its effect on mating patterns. I hypothesized females would choose to mate with MHC dissimilar males in order to increase the resulting offspring' fitness. I captured females and males during spawning and sequenced the DNA of these fish at the MHC class II putative peptide binding region. I found evidence that positive selection promotes genetic diversity at the MHC in bluegill, with a 5:2 ratio of non-synonymous to synonymous mutations. I found no evidence, however, that MHC genotypes affected mating patterns, either between females and parental males (social mates) or between females and cuckolder males (extra-pair mates). Given that parental males provide sole parental care to the eggs in their nest, the quality of male parental care may outweigh any potential benefit of MHC-based mate choice in bluegill.

Keywords: sunfish, major histocompatibility complex, mate choice, genetic quality, alternative reproductive tactics

Summary for Lay Audience

Many organisms prefer mates that are genetically dissimilar to themselves at certain genes. This produces offspring with multiple functional variants of the molecules these genes code for, which could be advantageous. The Major Histocompatibility Complex (MHC) genes are an example of a set of genes in which organisms look for mate dissimilarity. Here I provide the first genotypic characterization of the MHC class II peptide binding region in bluegill (Lepomis macrochirus), a species of sunfish, and examine its effect on mating patterns. I hypothesized females would choose to mate with MHC dissimilar males leading to an increase in the resulting offspring's fitness. I captured females and males during spawning and sequenced the DNA of these fish at one of the MHC genes. I found evidence that positive selection promotes genetic diversity at the MHC in bluegill. I found no evidence, however, that MHC genotypes affected mating patterns, either between females and parental males (social mates) or between females and cuckolder males (extra-pair mates). Given that parental males provide sole parental care to the eggs in their nest, the quality of male parental care may outweigh any potential benefit of MHC-based mate choice in bluegill.

Co-Authorship Statement

A version of this thesis has been submitted for publication to Behavioral Ecology with Bryan Neff and Shawn Garner as co-authors. I contributed to the study design, collected and analyzed the data, and drafted the manuscript. Bryan Neff and Shawn Garner contributed to the study design and provided guidance in the analysis of the data and the preparation of the manuscript.

Acknowledgements

Thank you to everyone who has helped me throughout my graduate studies. Thank you to the Neff lab and Dr. Bryan Neff for taking me in as a graduate student. I would like to thank Dr. Shawn Garner specifically for his countless hours helping me and working with me on this project, going above and beyond what could possibly be expected. Without Shawn and his seemingly limitless patience and knowledge, this project would not have been possible. Thank you to all the members of the Neff lab past and present that helped me throughout this project. In particular, thank you to Adriano, Malcolm, Mat, and Scott who all helped me greatly.

Thank you to the members of my advisory committee, Dr. Beth MacDougall-Shackleton and Dr. Graham Thompson. They played an important role in shaping my project, from proposal to defense. Thank the Queens University Biological Station where I conducted my sampling, met many great people, and learned a great deal about all kinds of interesting ecological research. I would also like to thank the wonderful people in the Biology office for all their help and support. Thank you Arzie Chant, Carol Curtis, and Diane Gauley.

Finally, thank you to my family. To my mother and father, thank you for supporting me throughout my entire education. Without extremely supporting and loving parents, I would not be where I am today. I love you both. Thank you to Courtney for putting up with me through all the stress that ~~is~~ was grad school. Your endless support, understanding, and love made all this possible in our busy lives. Thank you for being such a loving mother to our beautiful little girl. I love you both. Riley, if you ever find a passion for the sciences like I had, I look forward to sharing this research with you one day.

Table of Contents

List of Tables and Figures

1 Introduction

Biologists have long been interested in understanding mate choice and the nature of the benefits that it provides. Within a mating system, do individuals select mates essentially at random from the pool of potential mates, or do they choose specific mates based upon some criterion? Ample evidence suggests that mate choice is widespread in animals (Tregenza & Wedell 2000). For example, female vervet monkeys (*Cercopithecus aethiops*) are known to aggressively reject subordinate males in favour of dominant males (Keddy 1986). In the black grouse (*Tetrao tetrix*), males congregate within visual distance of one another in order to display to females in a behaviour known as lekking (Alatalo *et al.*, 1991). Male black grouses fight with each other while lekking and will remove feathers from the tails of their competitors; females then prefer to mate with males that have more feathers (Alatalo *et al.*, 1991). In guppies (*Poecilia reticulata*), males develop carotenoid-based orange colouration on their skin, and females prefer males whose colouration covers a greater proportion of their bodies (Houde 1997). Female guppies are generally less receptive to males after their first mating, but the likelihood is higher when second male is more orange than her first mate. Females are also able to bias paternity towards this second male (Pitcher *et al.*, 2003). Given the frequency and diverse taxa in which mate choice has been observed, one of the key questions is to understand the nature of the benefits provided by mate choice, which must outweigh the intrinsic costs of being selective and turning down potential mating opportunities.

One of the first explanations for why individuals may favour certain mates was proposed by Fisher (1915), who postulated that an organism's success was not only measured by the quantity of offspring they produced, but also by the quality of those offspring. By this he meant, offspring that are more likely to survive and produce offspring of their own. To

maximize the quality of their offspring, organisms can be selective of who they mate with based on a potential partner's phenotype (Tregenza & Wedell 2000). In many mating systems, females are the choosier sex, owing to their greater parental investment (Trivers 1972). In breeding systems where females are choosy, males may invest in traits that display their quality (Zahavi 1975). When selecting a mate, individuals may choose for so-called direct benefits, which are resources or advantages that increase fitness of the choosy sex. Examples include brood defense, territory quality, and fecundity. Alternatively, indirect benefits are fitness increases imparted to potential offspring through the chosen sex's genetic quality. In an example of direct benefits, sand goby (*Pomatoschistus minutus*), females show preference to mating with males that have more eggs in their nests. This is linked to higher survival of the female's offspring due to greater parental investment by the male and a dilution effect whereby predation risk to the individual is reduced due to the number of offspring in the nest (Forsgren *et al*., 1996). Indirect benefits are often displayed to females by males by some marker for genetic quality. In guppy, males with brighter orange coloration have increased foraging ability. Female guppy show preference for these bright orange males as the foraging ability can then be inherited by her potential offspring (Karino *et al*., 2005). In mating systems in which individuals select for indirect benefits, individuals may be consistent in their mating preferences and favour individuals that provide additive genetic benefits, termed "good genes". Here, females would choose mates based on traits that are correlated with the male's fitness (Kirkpatrick 1996). The outcome of this choice, however, should lead to fixation or near-fixation of the alleles associated with this increased fitness by directional selection and consequently the loss of any variation of those alleles in the population. This issue has become known as the lek paradox (reviewed by Kotiaho *et al*., 2008). Alternatively, individuals may choose mates for non-additive genetic benefits, which are often referred to as "compatible

genes" (Neff & Pitcher 2005). Compatible genes increase fitness through either overdominance (i.e. being paired with a specific homologue) or through epistasis (i.e. being paired with an allele at a separate locus) (Neff & Pitcher 2005). In this case, individuals choose mates based on a set of genes that will most complement their own; often because those genes are dissimilar and will increase the heterozygosity of their offspring, which can, for example, benefit offspring through increased immunity (Lenz *et al.*, 2009). Growing evidence suggests that mate choice for non-additive genetic benefits is common (Neff & Pitcher 2005).

The genes of the Major Histocompatibility Complex (MHC) may be important determinants of genetic quality. MHC genes encode cell-surface proteins that bind short peptides (antigens) from foreign pathogens and present those peptides to other immune cells, which initiate an immune response (Murphy *et al*., 2012). There are two classes of MHC molecules. Class I appears on all nucleated cells and primarily functions to present intracellular pathogens to cytotoxic T cells. Class II appears on antigen-presenting cells such as B cells and primarily functions to present extracellular pathogens to helper T cells (Murphy *et al*., 2012). The peptide binding region of the MHC protein acts as a pocket with high specificity for antigens (Murphy *et al*., 2012). An organism's immune cells, when functioning properly, can differentiate between antigens coming from it self and those of foreign bodies such as bacteria (non-self) (Murphy *et al*., 2012). The 3-dimensional shape of the peptide binding region is determined by the sequence of amino acid residues that comprise it. Brown *et al*. (1993) used X-ray crystallography to visualise the peptide binding region of human MHC and identify key amino acid positions that determine the specificity of the peptides that are bound. Differences in those key residues lead to different folding arrangements of the protein which ultimately lead to differences in the ability of the MHC molecule to bind particular pathogens (Murphy *et al*., 2012). Langefors *et al*. (2001), for example, showed that a specific allele in Atlantic

salmon (*Salmo salar*) was linked to higher resistance to infection from the bacterium *Aeromonas salmonicida*. Individuals that possess different MHC alleles, particularly at the amino acid level of the peptide biding region, are expected to recognize a greater range of pathogens. Indeed, heterozygosity at the MHC has been linked to increased survival. Penn *et al*. (2002) showed that mice (*Mus domesticus*) that were MHC heterozygous were more likely to survive bacterial infection than mice that were MHC homozygous. Worley *et al*. (2010) similarly showed that red junglefowl (*Gallus gallus*) that were MHC heterozygous survived infection longer than MHC homozygous individuals, and that the survival difference was independent of genome-wide heterozygosity. In Chinook salmon (*Oncorhynchus tshawytscha*), Evans & Neff (2009) found a positive effect of MHC heterozygosity on survival in populations with high rates of bacterial infection. Genetic variation at the MHC may thus be associated with genetic quality, particularly through non-additive effects on immunity and survival.

Given the examples of heterozygote advantages at the MHC, it follows that having three or more MHC alleles might be associated with an even greater ability to recognize pathogens. Indeed, there are multiple species where MHC genes are duplicated and an individual can express three or more unique MHC alleles. Karlsson & Westerdahl (2013) found evidence for at least 12 MHC loci in house sparrow (*Passer domesticus*). Guppies *(Poetcilia* reticulata) were found to have two MHC Class II loci (Fraser & Neff, 2009; McConnell *et al*., 1998). The question then becomes, what constrains the total number of MHC alleles that an individual has in its genome? Milinski (2006) addressed this question by postulating that too many MHC alleles would necessitate the immune system's removal of the inflated number of T-cells that would recognize self-antigens with so many different MHC alleles. This T-cell removal could be more costly than any advantage that such a large MHC repertoire would

confer. Research on fishes adds some support to this, as in three-spined sticklebacks (*Gasterosteus aculeatus*), which have up to six MHC class II loci, an optimal level of MHC dissimilarity appears to be more beneficial to survival than either extreme, (low or high dissimilarity) (Kurtz *et al*., 2004).

To examine selection on the MHC, sequence variation within a population could be examined, taking into account the open reading frame of the gene. Positive selection can be a signal of functional adaptive genetic variation. Within a population, positive selection can promote increased genetic divergence through, for example, heterozygote advantage or varying pathogen abundance (Penn *et al*., 2002; Sin *et al*., 2014). Positive selection can be shown at the genetic level by high levels of non-synonymous mutations as compared to synonymous mutations. Models have been developed that examine the rates of synonymous to non-synonymous mutations to infer selection. One such model is random sites codon-model-based approach (Yang 2007). Using this model, signals of positive Darwinian selection acting on MHC genes have been found in Grey partridge (*Perdix perdix*), frog (*Rhacophorus omeimontis* and *Polypedates megacephalus*), mummichog (*Fundulus heteroclitus*), and guppies (Promerová *et al*., 2013; Zhao *et al*., 2013; Cohen 2002, Fraser *et al*., 2010a). Population-level sequence analysis can thus provide evidence that selection promotes increased genetic variation at the MHC.

The potential survival benefits through enhanced immunity provides an opportunity for mate choice for indirect benefits from the MHC. Individuals might be predicted to choose MHC dissimilar mates in order to increase the number of unique alleles in their offspring. Early studies on mammals showed that mating preferences led to a reduction in the number of MHC homozygous individuals relative to expectations under random mating (e.g. Potts *et al*., 1991;

reviewed by Tregenza & Wedell 2000). More recently, effects of MHC on mating patterns have been observed in a number of fishes. For example, Landry *et al*. (2001) examined the effect of MHC on mating patterns in Atlantic salmon. First, the authors sampled the MHC genotypes of all adult salmon placed into an isolated section of a river. The salmon were allowed to mate naturally, and then the offspring were captured for genetic parentage analysis to determine realized mating patterns. For each potential mating pair, the authors calculated an MHC dissimilarity index that represented the expected number of amino acid differences between alleles in the pair's offspring. To compare the MHC amino acid difference between the observed mating pairs and the expectations under random mating, Monte Carlo simulations were used to randomly pair and calculate MHC dissimilarity over 50,000 replications. These simulations showed that Atlantic salmon mated with individuals that significantly increased the number of amino acid differences between MHC class II alleles in their offspring relative to expectations under random mating. Forsberg *et al*. (2007) also observed MHC-based mate choice in brown trout where it was shown that female brown trout more often mated with males of intermediate MHC dissimilarity. In this study, a similar method to that outlined by Landry *et al.* (2001) was used, where a value for each allele was calculated in order to determine dissimilarity between pairs. Forsberg *et al*. further quantified MHC dissimilarity by using the two most different alleles in each pair. Both approaches yielded evidence for female choice for dissimilarity as compared to what could be expected under random mating. In Chinook salmon, Neff *et al*. (2008) similarly showed that the observed mating patterns would lead to offspring with a greater number of amino acid differences at the MHC class II than expected under random mating. Chinook salmon were allowed to spawn in artificial channels and the offspring was collected and assigned paternity via microsatellite markers. The offsprings' MHC Class II genotypes were assigned and Monte Carlo simulations were used to compare observed patterns

to those expected under random mating. In three-spined stickleback, females prefer the odor of males with more unique MHC class II alleles relative to the odor of males with fewer (Reusch *et al*., 2001). Here, the overall result was still higher polymorphism at the MHC in offspring however, females did not choose mates based on dissimilarity to their own complement of MHC genes, instead opting for males with more distinct MHC alleles. Reusch *et al*. (2001) proposed this difference in the method by which MHC polymorphism arises in the species was due to the low probability of inbreeding occurring in this mating system. In species where inbreeding avoidance is important, mate choice for dissimilarity may be more prevalent than allele counting.

Despite reports of MHC-based mating preferences, a recent meta-analysis suggests that in many cases the effect of MHC genotypes on mating patterns is small or absent (Kamiya *et al*., 2014). This weak effect may occur due to other factors interfering with the expression of MHC-based mating preferences, leading to an overall weak effect of MHC on mating patterns. For example, in Chinook salmon, even though females may prefer MHC-dissimilar mates, male aggression towards unreceptive females can override their choice, especially when sex-ratios are male-biased (Garner *et al.*, 2010). Such opportunities for competition among males to override female mating preferences are common in many of the fishes in which MHC-based mating patterns have been tested, and it would be valuable to examine MHC-based mating patterns in a fish where female mate preferences are largely independent of competition among males. If potential mates provide important direct benefits it might also lessen the role of MHC dissimilarity in mating preferences, although in mice it appears that these competing demands are resolved by female MHC preferences primarily influencing the choice of extrapair mates (Potts *et al*., 1991).

Here I use a population bluegill (*Lepomis macrochirus*) from Lake Opinicon to investigate mate choice for MHC dissimilar partners. Bluegill are a freshwater fish native to North America, and the species has attracted considerable interest in sexual selection studies because of its mating behaviors. The bluegill mating system in Lake Opinicon was first described by Gross (1979). It typically lasts from late May to early July wherein multiple spawning bouts occur throughout the lake. In the days leading up to spawning, parental males congregate in the shallows of the lake (approximately 0.4 to 1.4 m in depth) and form colonies that can range in size from 5 to 150 nests (Gross 1982). Each parental male builds a bowl-shaped nest in the sediment through sweeping motions of his caudal fin. On the day of spawning, female bluegill swim around a colony and enter males' nests to initiate spawning. Spawning is synchronous and rarely lasts more than one day at any particular colony. Parental care, which is provided solely by the nest-tending parental male, involves fanning of the eggs and defense from nest predators, and it is essential for offspring survival (Gross 1982).

Bluegill are also characterized by an alternative reproductive tactic, wherein smaller precocious males called "cuckolders" steal fertilizations from parental males. Cuckolders provide no parental care for their offspring but leave that to the nest-defending parental males, which provide sole-parental care to the eggs and larvae in their nests (Neff & Gross 2001). During the day of spawn, females encounter colonies of several dozen males. While a female is mating with a chosen male, cuckolders may also deposit their milt within the nest. Small cuckolders use a sneaking tactic to ambush a spawning pair and release sperm near the eggs before darting out of the nest. Larger (older) cuckolders switch tactics and instead mimic females in coloration and behavior (Gross 1982). These female mimics are able to remain in the nest, often immediately between a spawning pair. Cuckolders, particularly sneakers, are more opportunistic (Stoltz & Neff 2006) and might circumvent, to some degree, potential

choice for MHC-dissimilar mates. Both tactics are successful at stealing fertilizations from the parental male (Fu *et al*., 2001). In bluegill, female choice is incongruent, albeit some males in a colony end up having tens of thousands of eggs in their nest after the day of spawn whereas other males end up having relatively few (Cargnelli & Gross 1996). The bluegill system provides an excellent opportunity to test for MHC-mediated mate choice, both between social and extra-pair mates. Additionally, male bluegill can neither coerce a female into a nest, nor prevent a female from leaving. In many situations, a female will enter a nest, swim alongside a male but leave before depositing any eggs.

Here I present the first genotypic characterization of the MHC class II locus in bluegill. I first used the random sites codon-model-based approach in PAML ver. 4 (Yang 2007) to investigate the patterns of selection at this locus by comparing observed sequence variation to that expected selection, using likelihood ratio tests. Next, to determine if mate choice is non-random with respect to MHC genotypes, I caught males and females in the act of spawning, and then analyzed their genotypes at the MHC class II. Following Landry *et al.* (2001), I used Monte Carlo simulations to compare the observed MHC amino acid differences between pairs to the expectations under random mating. I analysed MHC dissimilarity and potential offspring heterozygosity first at the entirety of the captured sequence, and then at only the amino acids that are likely to be essential for pathogen binding as determined by homology to the residues found to be key to peptide binding in human MHC (determined using x-ray crystallography by Brown *et al*., 1993). I hypothesized females would choose to mate with MHC dissimilar males in order to increase the MHC diversity—and hence fitness—of their offspring. I predicted that due to females having the opportunity to mate with numerous males in a colony, they would choose males non-randomly with respect to the MHC. I also caught a sample of cuckolder males that were in the act of stealing fertilizations from the aforementioned mating pairs so

that I could establish if cuckolders might influence any potential choice for MHC-dissimilarity. Here I again hypothesized that females would be more likely to mate with MHC dissimilar cuckolder males.

2 Methods

2.1 Study Species and Sample Collection

Sample collection occurred at the Queen's University Biological Station (QUBS) on Lake Opinicon (44.5°N, 76.3°W). In the summer of 2015, I swam a 6 km transect along the shoreline of Lake Opinicon daily to monitor spawning activity. When spawning commenced, swimmers floated motionless over a colony of nests observing the activity and watching for females entering nests. Spawning pairs were observed until a female dipped, a movement where the female turns onto her side and releases a small batch of her eggs into a parental male's nest (Gross 1991). Once dipping was observed at least five consecutive times between a pair, both parental male and female were caught with a dip net. A mesh cover was then placed over the male's nest to protect the eggs from predation by other fish. The nest was also marked with a uniquely numbered ceramic tile to allow for identification of the nest. The mating pair was then brought to a boat where total body length was recorded and a small fin clip was removed from each fish's caudal fin and stored in 95% ethanol for later genetic analysis. Both male and female fish were then returned to the water. Parental males typically returned to their nests immediately and commenced courting other females. Occasionally, a previously caught parental male (identified by nest tile and fin clip) was recaptured with a new female (identified by the absence of a fin clip), or two females were captured while simultaneously mating with a parental male. In these cases, each parental male-female pairing were treated as a novel pair. In total, five of 35 pairs in my sample were comprised of previously caught males.

In addition to spawning pairs that contained a parental male and female, I also attempted to capture spawning groups that included at least one cuckolder male. Sneakers are more challenging than other bluegill to capture during spawning, as they are small, quick, and are actively chased from nests by parental males when detected. In total, I captured only four sneakers in the act of spawning. Two of the four sneakers were caught in a single spawning group that contained one parental male and two females. Each sneaker-female pairing in this group were considered a distinct pair during analysis, resulting in a total of six sneaker and female pairs. I did not collect any female mimics.

2.2 MHC Primer Design and Sequencing

Primers for amplification of MHC II in bluegill were designed based on sequences from a bluegill brain transcriptome (Partridge *et al*., 2015). Briefly, Partridge *et al.* (2015) used high-throughput sequencing to characterize the sequences of expressed transcripts from the brains of 20 bluegill collected in Lake Opinicon. The resulting transcriptome consisted of 235,547 transcripts. Using the transcriptome as a local database, NCBI BLAST was used to search for potential MHC class II putative peptide binding region sequences by using known MHC sequences including that of striped sea bass (*Morone saxatilis*, Genbank id: L33967) and three-spined stickleback (*Gasterosteus aculeatus*, Genbank id: DQ016429). Probing the transcriptome for exon 2 of MHC class II with these sequences yielded a single transcript in bluegill. Using the bluegill MHC II transcript sequence, Primer-Blast was used to develop a primer pair (Forward: GCATTCCTCAGTGGTCCGC and Reverse: TGTACCAGTTCCCAATGTTG) that spanned a 239 base pair region of the putative MHC II locus.

To test the MHC II primers, I first extracted DNA from bluegill fin clips via Proteinase K digestion and ethanol precipitation (Neff *et al*., 2000). Next, I PCR-amplified three parental males at the MHC locus. The PCR amplicon was cloned using a pGEM T-easy vector kit following manufacturer's instructions (Promega Corp) and used to transform *Escherichia coli*, which were then grown on lysogeny broth agar plates. Bacterial colonies containing the insert were collected and re-amplified using the standard sequencing primers SP6 and T7. Eight insert-containing colonies from three individuals were Sanger-sequenced by the London Regional Genomics Centre. The resulting sequences were analysed with the BLASTX algorithm in NCBI BLAST (Altschul *et al*., 1990), which confirmed that the bluegill MHC amplicon had high similarity to the putative peptide binding region of MHC class II in other teleost fishes (88% identity with *M. saxatilis*, 87% identity with *Epinephelus coioides*). Bluegill MHC sequences were then aligned with the human MHC class II peptide binding region to identify the specific amino acid positions likely to comprise the key residues of the pathogen peptide binding region following the X-ray crystallography determinations of Brown *et al*. (1993).

Next-generation sequencing was used to sequence 50 bluegill collected from Lake Opinicon. First, samples were PCR-amplified with of the MHC primers that included a UniA tail on the forward primer and a UniB tail on the reverse primer. After the PCR, the product was visualized on a gel to ensure amplification occurred and then cleaned using ethanol precipitation (Neff *et al*., 2000). A second PCR with primers specific to the UniA and UniB tails was then used to attach an Ion Torrent adaptor and sample-specific barcode in the forward direction, and an Ion Torrent adaptor sequence in the reverse direction (Venney *et al*., 2016). The sample-specific barcode (a unique 10-11 bp sequence) allowed multiple individuals to be pooled in a single sequencing run, with the resulting sequences assigned to individuals based

on these unique barcodes. A QIAquick PCR Purification kit (QIAGEN) was used to purify the products after the second PCR. DNA concentrations in each sample were measured using a Nanodrop (ND-3300, NanoDrop Technologies) and pooled in equal concentrations. The resulting library was sequenced on an Ion Torrent Personal Genome Machine (Life Technologies) at the University of Windsor Environmental Genomics Facility.

AmpliSAT software (Sebastian *et al.*, 2016) was used to sort and clean the resulting Ion Torrent sequence data. Briefly, AmpliSAT de-multiplexes, clusters, and filters the raw sequencing data allowing for the removal of artefacts and the assignment of alleles to the amplicon. Sequence variants that appeared as less than 1% of an individual's total reads were discarded as sequencing errors following protocols established by Galan *et al.* (2010). Chimeric sequences were identified within an individual as low-frequency sequence variants that were a combination of two common alleles possessed by that individual and were removed. After these clean-up steps, all individuals possessed either one or two unique alleles, consistent with a single unduplicated MHC II locus in bluegill.

In addition to the individuals sequenced using the Ion Torrent sequencing method, an additional 19 individuals were sequenced using a direct Sanger-sequencing technique. The MHC amplicons for these individuals were PCR-amplified as described above, and then sequenced in both directions at the London Regional Genomics Centre. The resulting chromatograms were manually examined in BioEdit software (Hall 1999) to determine the sequences and to identify heterozygous positions, which were characterized by two peaks of similar intensity at a variable site. Based on the alleles identified using the Ion Torrent sequencing, each combination of alleles would result in a unique pattern of variation on the

chromatograms. I then assigned each individual an MHC genotype based on their chromatogram sequence.

2.3 MHC Characterization

To investigate how selection is acting on the MHC in bluegill at the codon level, the random sites codon-model-based approach in PAML ver. 4 (Yang 2007) was used as described in Fraser *et al*. (2010b). Briefly, PAML uses the ratio of estimated rates of synonymous dS to non-synonymous dN substitutions (ω) to examine how codons are evolving according to several models. Four potential models were assessed for their likelihood using a nested log-likelihood ratio test. Model M1a is the 'nearly neutral' model which creates an estimate for the proportion of codons undergoing purifying selection (p_0, $0<\omega_0<1$) and the remaining proportion of codons that are neutrally evolving ($p_1=1-p_0$, $\omega_1=1$). Model M2a is the 'positive selection' model which includes model M1a with the addition of a third class of codons where positive selection is occurring ($\omega_2>1$) defined by the proportion p_2 ($=1-p_0-p_1$). Models M7 and M8 apply a less restrictive definition for ω between 0 and 1 with the use of a β distribution. The β distribution is a flexible probability density function used to capture further variation in rates across codons and is estimated from the data itself (see Yang *et al*., 2000). Model M7 is analogous to M1a and serves as the null β model ($0<\omega_0<1$). Model M8 is the "positive selection plus β" model which is analogous to M2a which again is equivalent to M7 but allows for a proportion of codons that are undergoing positive selection (p_1, $\omega_1>1$). Using a Bayes Empirical Bayes approach (Yang 2005), codons were allotted to different selection classes under a 95% posterior probability cut-off. Likelihood ratios were calculated in PAML in order to compare between models.

ML-NullFreq was used to test for an excess in observed homozygosity, which may indicate the presence of a null allele that was undetectable through sequencing (e.g. because of a mutation in the primer-binding sites; Kalinowski & Taper 2006). ML-NullFreq was further used to predict the frequency of a potential null allele and recalculate the frequencies of other alleles based on this prediction.

2.4 Mating pair analysis

I used two measures to compare the MHC similarity of mating pairs. First, expected heterozygosity of a pair's offspring was calculated. Second, following Landry *et al*. (2001), for each mating pair I calculated the number of amino acid differences between the MHC alleles of the male and female. An average of the four values was then calculated, which represents the expected number of amino acid differences between MHC alleles in the pair's offspring. These two measures were calculated using the entire length of the MHC amplicon, and again using only the key peptide binding residues identified by Brown *et al*. (1993) as being most important for determining the binding properties of the peptide binding region.

Following Landry *et al*. (2001) and Neff *et al*. (2010), Monte Carlo simulations were then used to create expected distributions for offspring heterozygosity and amino acid differences between pairs under a model of random mating. The simulation randomly paired females and males from the entire population and maintained the observed number of mates for each fish. Each simulation was repeated 10,000 times to generate a distribution under random mating. The observed values for offspring heterozygosity and amino acid differences between pairs were then compared to these random distributions to determine if the observed values differed significantly from the expectations under random mating. P-values were calculated as the proportion of the simulated values greater than the observed values (i.e. the

probability of observing offspring heterozygosity or amino acid differences between pairs at least as high as was observed under random mating).

I used a power analysis to quantify my power to detect non-random mating at the MHC if it were occurring, given a particular effect size. For this analysis I first calculated the maximum offspring heterozygosity and amino acid differences between pairs for each female based on the observed male genotypes. I then assigned each female a mate at random, varying the rate at which she would choose her optimal mate in 10% increments ranging from 0 to 100%. This process was repeated 10,000 times. Power at each increment was then calculated as the proportion of the replicates that were greater than the 95^{th} percentile when mates were assigned at random (i.e. the proportion of the trials in which a significant effect would be detected given the null distribution I generated). Because similar results were obtained when analyses were based on the full MHC amplicon sequence or only key residues, for this and subsequent analyses I present only the results based on the full MHC amplicon sequence.

I also used a Monte Carlo simulation to examine the potential effects of a null allele on my inferences about mating patterns. First, individuals with a single MHC allele were randomly assigned as homozygous for the observed allele or heterozygous for the observed allele and the null allele. The probability of being assigned the null allele was proportional to the population-wide frequency of the null allele relative to the frequency of the observed allele (e.g. individuals with a common allele were more likely to be true homozygotes than individuals with a rare allele). Next, in order to calculate offspring heterozygosity and amino acid differences between pairs, I assigned the null allele a sequence equivalent to one of the observed alleles at frequencies proportional to the observed allele frequencies. For each generation of the simulation the null allele had a single identity. I repeated this process 10,000

times, and for each generation calculated offspring heterozygosity and average amino acid differences between pairs for the observed mating pairs and for randomized mating pairs. I calculated a p-value as the proportion of the replicates for which the observed value was greater than the random value.

Finally, I calculated offspring heterozygosity and amino acid differences between pairs for each of the six pairs of cuckolder males and females. I then compared these values to the values for the observed parental male and female pairs using independent samples t-tests.

3 Results

3.1 Collection summary

Parental males were collected from eight colonies and had an average length of 189 ± 12 mm (mean ± SD; range: 165-209 mm; n=30). The females captured in the nests of the parental males were smaller, with an average length of 143 ± 26 mm (mean ± SD; range: 80-198 mm; n=35). The cuckolders had an average length of 80 ± 14 mm (mean ± SD; range: 67-100 mm; n=4). No significant correlation was observed between the length of parental males and females from observed mating pairs (P=0.11, Rho=0.28, n=35; Figure 1).

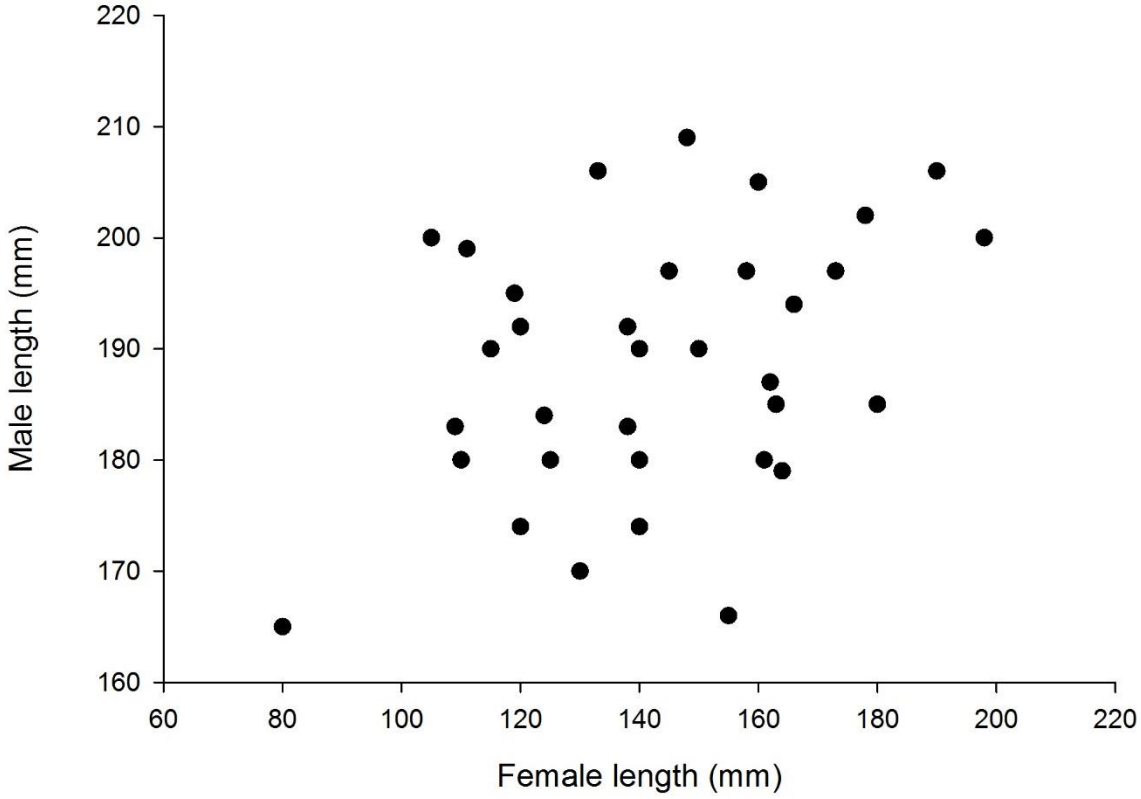

Figure 1. Plot of male and female length for observed mating pairs in bluegill (*Lepomis macrochirus*).

3.2 MHC Characterization

Ion Torrent sequencing resulted in 143,122 useable sequence reads across 50 bluegill samples. In total, I found ten different putative MHC class IIB alleles of varied frequency with five variable amino acid sites (Table 1). No insertions/deletions or stop codons were observed in these sequences, which is consistent with each sequence representing an expressed protein-coding gene. No apparent gene duplications were identified due to the absence of more than two alleles in any individual. Across all individuals the observed heterozygosity was 53%, whereas the expected heterozygosity assuming Hardy-Weinberg equilibrium was 83%. The calculated probability of observing a heterozygosity deficit of this magnitude given Hardy-Weinberg equilibrium is highly unlikely. MLNullFreq detected a significant excess of homozygosity ($p<0.001$) and predicted that there was a null allele with a frequency of 0.24 (Table 1).

When comparing selection models on the MHC sequences, the M2a positive selection model was significantly more likely than the M1a nearly neutral model (LnLRT =64.55; $P<0.0001$). The M8 positive selection model was also significantly more likely than the M7 null model (LnLRT = 64.78; $P<0.0001$). This model fitting indicates that the peptide binding region in bluegill is under positive selection and that 6.5% of codons showed signs of positive selection while the remaining 93.5% were under purifying or neutral selection. In both positive selection models, the Bayes Empirical Bayes method showed five codons in the peptide binding region of MHC class II were under positive selection (positions 15, 23, 39, 55, and 62, with posterior probabilities >0.97 for each; Figure 2).

Table 1. Unique alleles observed at the MHC class II putative peptide binding region in bluegill (*Lepomis macrochirus*). Amino acid residues are shown only for polymorphic positions. Asterisks denote analogs of sites found to be key to peptide binding in human MHC class II peptide binding region through 3-dimensional imaging (Brown *et al*., 1993). Adjusted frequencies with the inclusion of a null allele were calculated using ML-NullFreq (Kalinowski 2006). Genbank accession numbers are included for each allele.

				Polymorphic amino acid position				
Allele	Frequency	Adjusted frequency	Genbank accession #	15	23 *	39	55 *	62 *
BG1	0.31	0.24	MK620857	N	E	V	R	Q
BG2	0.14	0.1	MK620858	H	E	V	R	Q
BG3	0.17	0.1	MK620859	H	E	L	R	Q
BG4	0.05	0.03	MK620860	N	Q	V	R	I
BG5	0.09	0.07	MK620861	H	E	V	H	Q
BG6	0.05	0.05	MK620862	H	Q	L	R	Q
BG7	0.09	0.07	MK620863	N	Q	V	R	I
BG8	0.06	0.05	MK620864	N	E	V	H	I
BG9	0.01	0.01	MK620865	N	Q	L	R	I
BG10	0.02	0.01	MK620866	H	E	V	R	I
null		0.24						

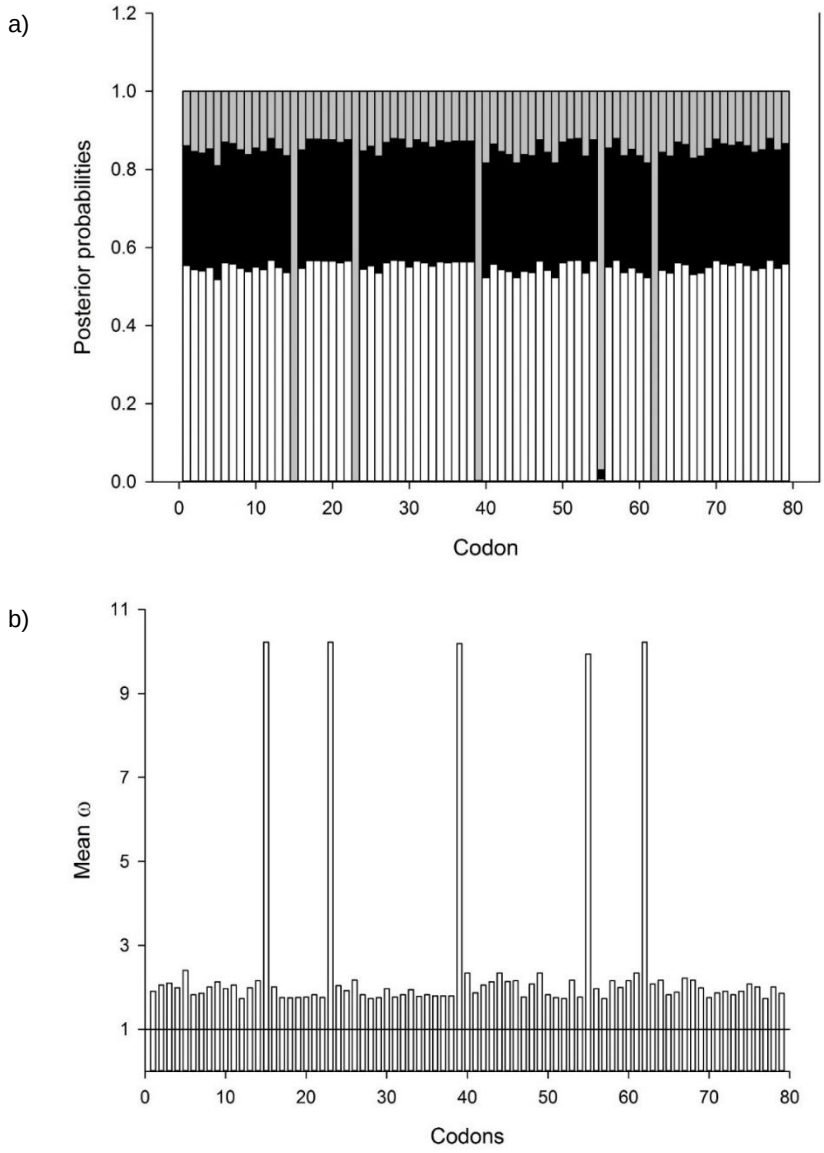

Figure 2. Sequence-level selection occurring at the MHC class II putative peptide binding region in bluegill (*Lepomis macrochirus*). a) MHC class II putative peptide region codons and their posterior probabilities of being under different selection classes: purifying selection (white), neutral selection (black), and positive selection (grey). b) Weighted mean ω values for each codon.

3.3 Mating pair analyses

For the observed mating pairs, 83% of the offspring would be expected to be heterozygous under Hardy-Weinberg equilibrium at the MHC class II locus. There was no significant difference between the observed heterozygosity and the expectations under random mating when I analyzed the complete MHC sequence (P=0.66, Figure 3a). When I instead calculated heterozygosity based only on the key residues of the peptide binding region, there was again no significant difference between the observed heterozygosity and the expectations under random mating. (P = 0.66, Figure 3b).

For the observed mating pairs, the mean number of amino acid differences between pairs was 1.86 (range = 0 to 3). There was no significant difference between the observed amino acid differences between pairs and the expectations under random mating (P=0.55, Figure 3c). When I instead calculated amino acid differences between pairs based only on the key residues of the peptide binding region, the mean amino acid differences between pairs was 1.05 (range = 0 to 2.5). There was no significant difference between the observed amino acid differences between pairs at the key peptide binding region residues and the expectations under random mating (P=0.93, Figure. 3d).

The power of my study to detect non-random mating as a function of the frequency with which individuals select an MHC-optimal mate is summarized in Figure 4. The power to detect non-random mating with respect to offspring heterozygosity was 95% when individuals selected an optimal mate 69% of the time, and 80% when individuals selected an optimal mate 59% of the time. The power to detect non-random mating with respect to amino acid differences between pairs was 95% when individuals selected an optimal mate 32% of the time, and 80% when individuals selected an optimal mate 23% of the time.

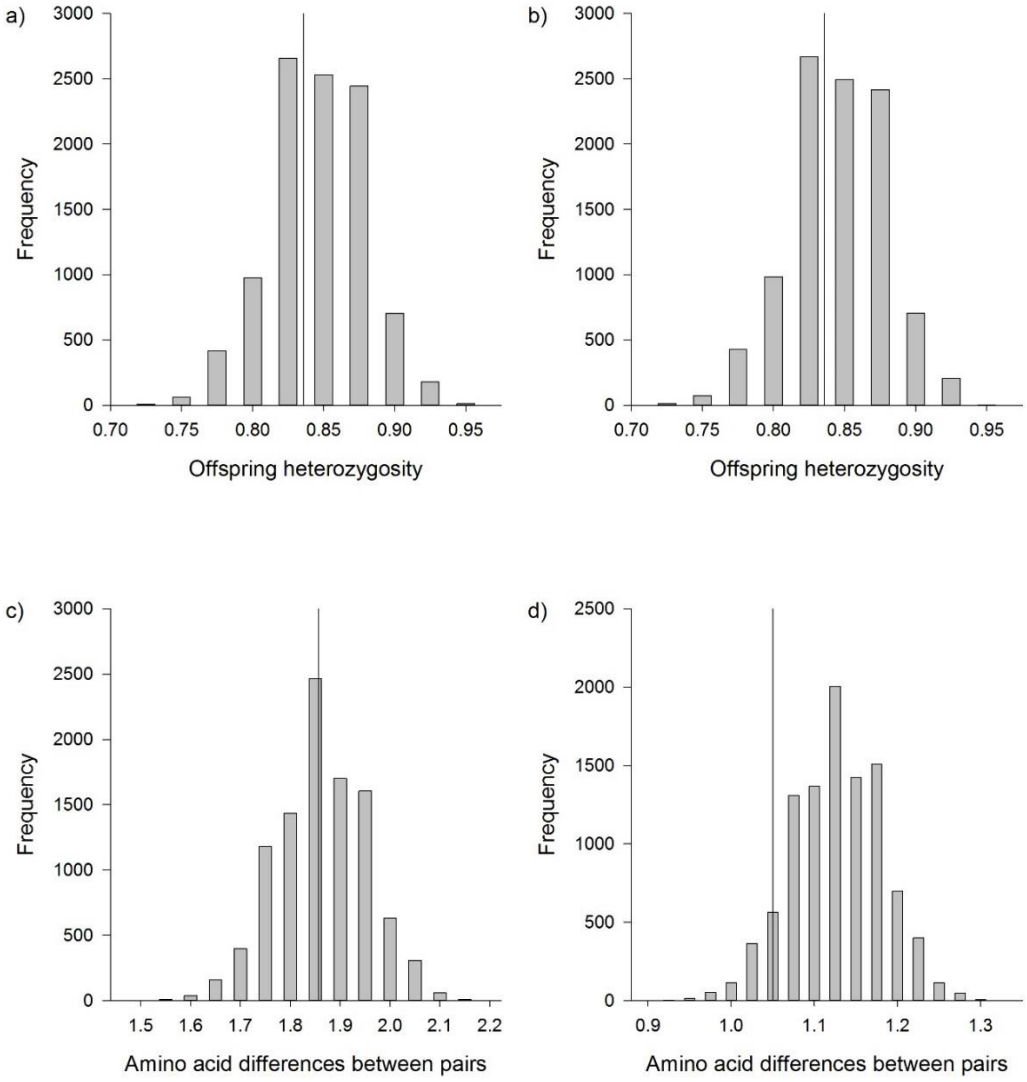

Figure 3. Expected offspring heterozygosity and amino acid differences between pairs at the MHC class II peptide binding region under random mating in bluegill (*Lepomis macrochirus*). Panel a) shows offspring heterozygosity based on the complete MHC amplicon, panel b) shows offspring heterozygosity based only on key peptide binding residues, panel c) shows amino acid differences between pairs based on the complete MHC amplicon, and panel d) shows amino acid differences between pairs based only on key peptide binding residues. Expected data are from Monte Carlo simulations and show the

results of 10,000 replicates that randomly paired males and females. Observed values are shown with a vertical line.

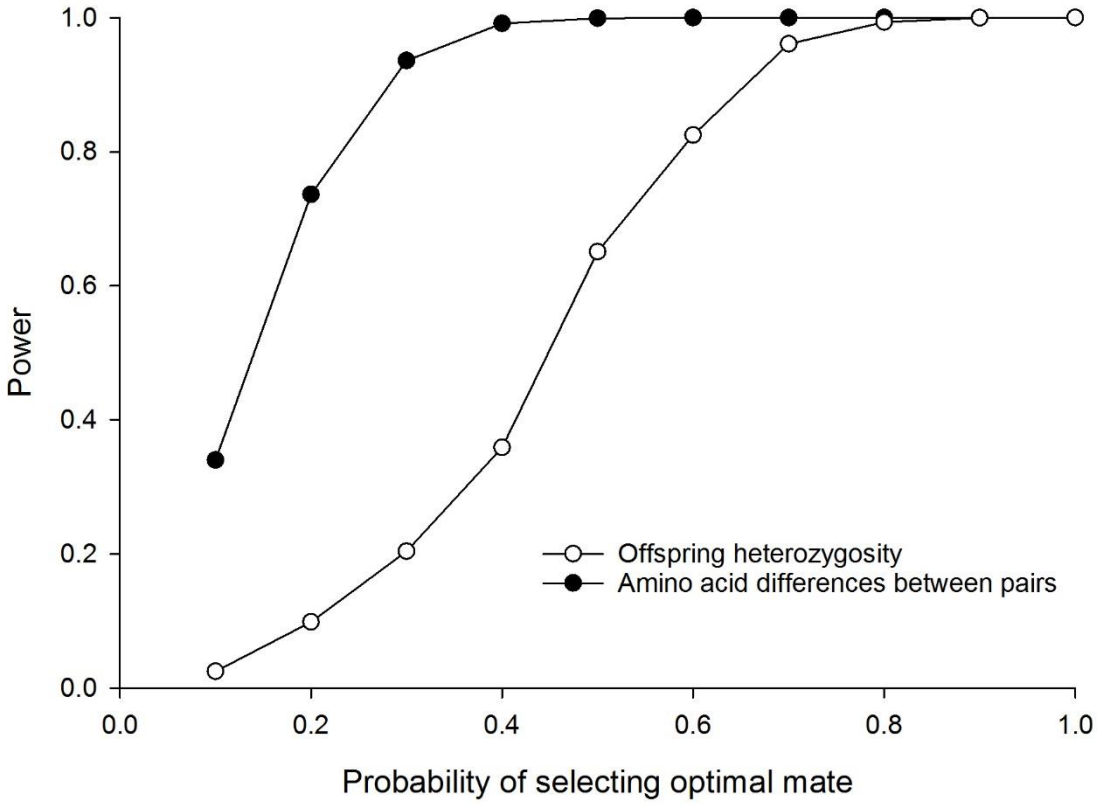

Figure 4. Power to detect non-random mating at the MHC class II peptide binding region in bluegill (*Lepomis macrochirus*) as a function of the probability that individuals select an optimal mate based on either offspring heterozygosity or amino acid differences between pairs.

When the presence of a null allele was included in simulations, offspring heterozygosity did not differ significantly between the observed mating pairs and random mating pairs (P=0.52, Figure 5a). Amino acid differences between pairs also did not differ significantly from the expectations under random mating when a null allele was included in the simulations (P=0.51, Figure5b). Expected offspring heterozygosity at the MHC did not differ between cuckolder-female and parental-female mating pairs ($t_{39} = 1.05$, P = 0.30, Figure 6a). Similarly, amino acid differences between pairs at the MHC did not differ between cuckolder-female and parental-female mating pairs ($t_{39} = 0.14$, P = 0.89, Figure 6b).

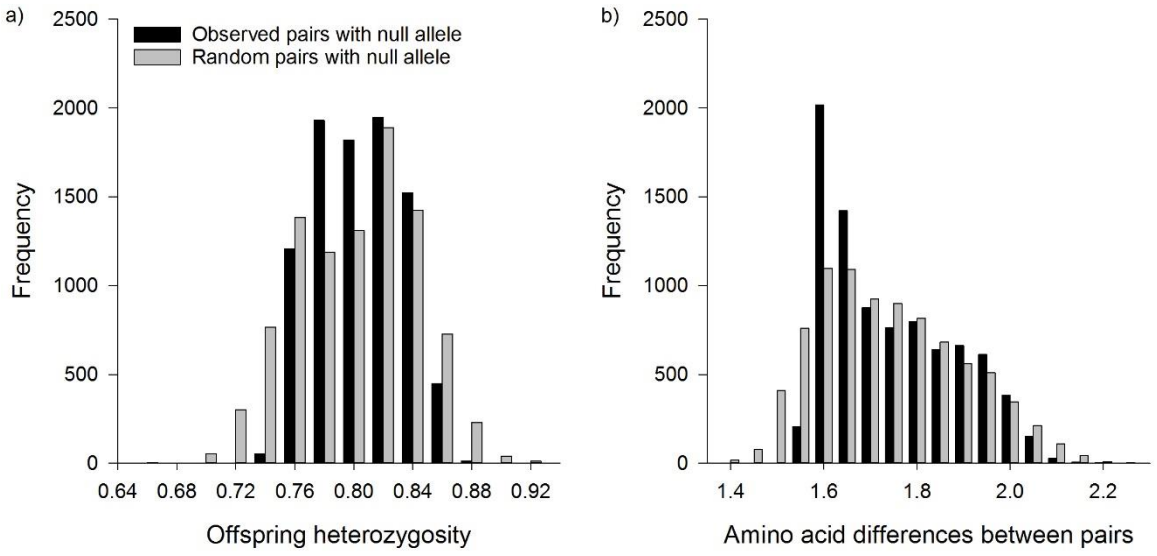

Figure 5. Effects of a null allele at the MHC class II peptide binding region on MHC dissimilarity for observed and random mating pairs in bluegill (*Lepomis macrochirus*). Panel a) shows offspring heterozygosity, and panel b) shows amino acid differences between pairs. Data are from Monte Carlo simulations and show the results of 10,000 replicates that probabilistically incorporated a null allele for individuals that were homozygous at the MHC.

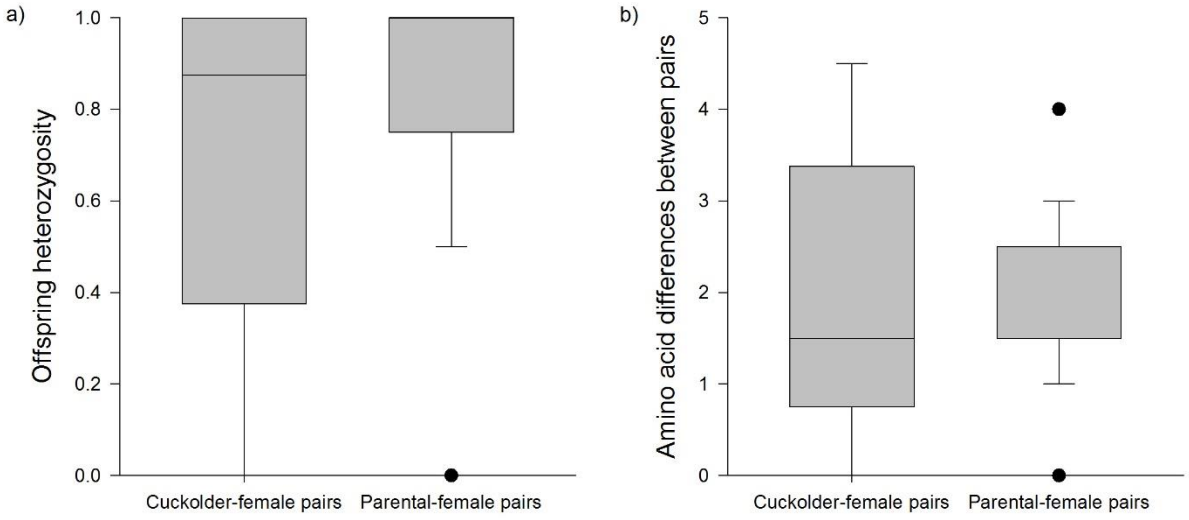

Figure 6. Box plots comparing the MHC class II peptide binding region for cuckolder-female pairs and parental-female pairs in bluegill (*Lepomis macrochirus*). Panel a) shows offspring heterozygosity, and panel b) shows amino acid differences between pairs. The horizontal line within the box represents the median, the boundaries of the box indicate the 25th and 75th percentile, the whiskers indicate the 10th and 90th percentile, and the dots indicate outliers.

4 Discussion

4.1 MHC Characterization

Here I present the first characterization of the MHC class II in bluegill. Using primers designed from an existing bluegill transcriptome (Partridge *et al*., 2015), I sequenced a 239 base pair section of the putative peptide binding region. Multiple indicators are consistent with these primers capturing a functional MHC locus. First, one of my sequences was identical to the consensus transcriptome sequence, and all of my sequences had high similarity to the consensus transcriptome sequence (>97% identity), which indicates that the MHC locus I sequenced is expressed. Second, the sequences did not contain stop codons or frameshift mutations that would indicate that the gene encodes a non-functional protein (i.e. is a pseudogene). Third, my sequences had high identity with MHC class II putative peptide binding regions that have previously been characterized in other teleosts (88% identity with *M. saxatilis*, 87% identity with *Epinephelus coioides*). High identity between MHC sequences in bluegill and other fishes are consistent with this region being a conserved, functional gene. My sequence data also suggest that the MHC II consists of a single locus in bluegill, with all individuals possessing either one or two unique alleles. Overall, my data suggest that the primers capture a single expressed MHC class II gene in bluegill.

Given the role of MHC sequence variation in the pathogen-binding characteristics of MHC proteins (Hedrick 2002), MHC genes have frequently been linked to positive selection and high levels of functional variation (Cohen 2002). In bluegill, I found strong evidence that the MHC is subject to positive selection. Using PAML, models of positive selection were found to explain the observed sequence data significantly better than neutral models. I also observed a high ratio of non-synonymous to synonymous mutations (5:2), which is consistent

with positive selection acting on the gene. These results are consistent with those found in other fishes, with PAML models providing support for positive selection on the MHC class II gene in guppies (Fraser *et al*., 2010b), rainbow trout (*Oncorhynchus mykiss*) (van der Aa *et al*., 2009), and Atlantic cod (*Gadus morhua*) (Fernandes *et al*., 2010). High ratios of non-synonymous to synonymous mutations have likewise been found in southern platyfish (*Xiphophorus maculatus*) (9:2, McConnell *et al.*, 1998). Overall, the evidence of positive selection affecting the MHC class II gene in bluegill appears to be in line with what has been found in other fish species and indicates that natural selection has a major role in promoting functional variation at the MHC in bluegill and other fishes.

In bluegill, tests for Hardy-Weinberg equilibrium at the MHC showed a significant excess of homozygotes, which might be explained by several processes. First, this homozygote excess may have resulted from population subdivision or inbreeding. If the homozygote excess at the MHC is caused by population subdivision or inbreeding, then I would also expect a similar homozygote excess at other genes. However, previous studies using microsatellite genetic markers in Lake Opinicon bluegill have shown no evidence of a genome-wide homozygote excess (Garner & Neff, 2013; Neff 2001), suggesting that population subdivision or inbreeding cannot explain the MHC genotype data. Second, the homozygote excess at the MHC could result from non-random mating for high MHC homozygosity in offspring. My simulations indicated that the observed mating patterns were likely random with respect to MHC heterozygosity (P=0.66) and that the observed mating patterns should lead to much higher MHC heterozygosity than was observed in the parents (the observed heterozygosity was 53%, whereas 83% offspring heterozygosity was expected for the observed mating pairs). Non-random mating is thus unlikely to explain high MHC homozygosity. Instead, the homozygote excess at the MHC II in bluegill is most consistent with the presence of a null allele, which

would not be amplified during PCR due to, for example, a mismatch in the primer-binding sequence. Null alleles have previously been indicated in a number of studies of MHC in species that include bighorn sheep (*Ovis canadensis*), Atlantic salmon (*Salmo salar*), and black-throated blue warbler (*Dendroica caerulescens*) (Dionne *et al.*, 2007; Gutierrez-Espeleta *et al.*, 2001; Smith *et al.*, 2005). In bluegill, a null allele present at an estimated frequency of 24% is the most likely explanation for the observed excess of homozygosity at the MHC.

4.2 Mate choice

Size assortative mating can be found in many mating systems. In many taxa, females prefer to mate with large males, likely as a result of positive relationships between body size and either genetic quality or the direct benefits that the male provides (Baldauf *et al*., 2009). If males similarly prefer to mate with large females, as for example when female size is linked to higher fecundity (Olsson 1993), then a pattern of positive size assortative mating will result. In bluegill I did not detect size assortative mating, as there was no correlation between male and female body size across mating pairs. The absence of assortative mating is not surprising in this system, as in my observations, parental males appear to accept any mating opportunity with a female and rarely chase females from their nests. Males do not appear to be limited by the capacity of their nests to incubate eggs, as even males that were observed to have high egg counts still accepted new mates. It is also unclear if female bluegill prefer to mate with large parental males. In smallmouth bass (*Micropterus dolomieu*), females prefer to mate with large males, likely because male smallmouth bass provide sole parental care to the developing eggs, and a preference for large body size may allow females to mate with males that have more energy available to provide high-quality parental care (Hanson & Cooke, 2009). Bluegill are closely related to smallmouth bass and share a similar division of parental care, so a preference

for large males would be predicted, but this preference has never been evaluated in bluegill. Regardless, my results suggest that assortative mating based on body size is not common in bluegill and is unlikely to constrain mate choice based on MHC genotypes.

MHC-based mate choice has been reported in many taxa, including multiple fishes (recently reviewed by Kamiya *et al*., 2014). Here I provide the first test of MHC-based mate choice in bluegill. I found that the observed mating pairs between parental males and females did not differ from the expectations based on random mating with regards to either offspring heterozygosity or amino acid differences between pairs. No evidence for non-random mating was found regardless of whether MHC differences were based on the entire amplicon sequence or only on the key residues of the peptide binding region identified by Brown *et al.* (1993). In contrast, significant effects of MHC on mating patterns have previously been identified in other fishes, including for example a preference for high amino acid differences between pairs in Atlantic salmon (Landry *et al*., 2001), and chinook salmon (Neff *et al*., 2008). These differences in mate choice for MHC could arise from species and environment-specific factors such as the nature of pathogen-mediated selection on the MHC or the opportunity for mate choice in a mating system. However, the absence of MHC-based mate choice in a fish appears to be novel, though there are examples of mammals where a similar lack of evidence for MHC-based mate choice have been found (Kuduk *et al*., 2014; Liu *et al*., 2017).

Given the frequency with which MHC-mediated mate choice has been observed in fishes, it is important to consider whether its absence in bluegill could be explained by limitations in my study design. For example, it is possible that non-random mating with respect to the MHC class II was not observed in bluegill due to low statistical power to detect an effect. Indeed, Hoover & Nevitt (2016), highlighted the importance evaluating power in MHC studies

in order to prevent misinterpreting results due to the high variability found at MHC genes. They outlined the importance of a sample size (and therefore statistical power) large enough to detect both the entire variation of MHC genes in a population, as well as the true population mean. Given my sample size, my analysis showed that I had 80% power to detect mate choice if females were able to choose MHC-optimal males at least 23% of the time. A mating preference of this strength would be associated with an increase in the average amino acid differences between mates of about 0.3 amino acids. In a population of Atlantic salmon, Landry *et al*. (2001) found that mate choice increased the number of amino acid differences between mating pairs at key residues of the MHC by about 0.2 amino acids. In chinook salmon, non-random mating was associated with an increase in amino acid differences between pairs of about 0.3 amino acids (Neff *et al*., 2008). My power analysis thus indicates that I had high power to detect an MHC-mediated mate preference of a similar magnitude to those that have been observed previously.

Another potential challenge that may have reduced my power to detect non-random mating with respect to the MHC is the presence of a null allele, which I estimated to be present at a frequency of 24% in my population. To address this concern, I used a novel simulation approach in which I incorporated the null allele into my Monte Carlo simulations to assess if its presence would alter my conclusions. I found that a null allele was unlikely to have significantly altered my results, and that after accounting for the presence of the null allele, the offspring heterozygosity and amino acid differences between pairs still did not differ significantly from expectations under random mating. Although null alleles have previously been identified in studies of the MHC (Dionne *et al.*, 2007), my study represents one of the first to incorporate a null allele at the MHC into tests of assortative mating. That the incorporation of a null allele into my mating simulations had no effect on my conclusions

suggests that the presence of a null allele does not compromise the ability to quantify mating patterns with respect to the MHC.

One reason that MHC might not influence mating patterns is that other targets of mate choice are more important. In bluegill, a likely target of female mate choice that may supersede MHC-mediated mate choice is a preference for male parental care, as male bluegill provide sole parental care to the developing larvae (Neff 2008). In other fishes where MHC-mediated mate choice has been observed, males typically provide limited parental care. For example, MHC-mediated mate choice has been observed in Atlantic salmon, in which males provide no parental care (Landry *et al*., 2001, Consuegra & Garcia de Leaniz, 2008). In brown trout, Forsberg *et al*. (2007) observed MHC-based mate choice, and male parental care is limited to defense of the eggs from cannibals for approximately 2 minutes after spawning (Tentelier *et al*., 2011). This difference in parental care may contribute to the presence of MHC-based mating in some species and the lack thereof in bluegill. Specifically, the quality of a bluegill male's parenting may outweigh any potential benefit to a female of selecting a mate based on MHC genotype.

In bluegill, with the high percentage of heterozygosity among the expected offspring I found, it is possible that female choice for dissimilarity at the MHC would be largely ineffective. If offspring are already likely to be MHC heterozygous, and thus receive any potential benefits associated, then choosing mates to maximize this likelihood may be less beneficial than choosing mates based on some other criteria. However, Landry *et al*. (2001) found similarly high levels of expected MHC heterozygosity under simulated random mating and still found evidence that Atlantic salmon choose mates based on MHC. In some populations, its possible the maximizing MHC heterozygosity of offspring is beneficial even

when MHC heterozygosity would otherwise already be likely, though this does not appear to be the case in the bluegill I sampled.

An under-studied question surrounding MHC-mediated mate choice is the role of extra-pair mating. In bluegill, extra-pair mating is driven by cuckolder males that use alternative reproductive tactics to gain paternity in the nests of parental males (Gross, 1982). These alternative reproductive tactics may enhance the opportunity for MHC-mediated mate choice by enabling females to mate with a cuckolder male for genetic benefits (MHC compatibility) and a parental male for direct benefits (parental care). Though my sample size was small, I did not find any evidence of MHC-based mate choice in bluegill between cuckolders and females, as offspring heterozygosity and amino acid dissimilarity between pairs did not differ significantly between cuckolder-female mating pairs and parental-female mating pairs. In contrast to my results, Schwensow *et al*. (2008) found evidence in primates that extra-pair mates are chosen due to dissimilarity at the MHC more so than social mates are. Promerová *et al*. (2011) found that scarlet rosefinch (*Carpodacus erythrinus*) females were more likely to seek extra-pair mates if their social mates had lower MHC diversity. Alternatively, the presence of male alternative reproductive tactics could reduce the benefit for MHC-mediated mate choice when cuckolder males mate indiscriminately with respect to MHC compatibility. By undermining female preference for a mate (parental males in the bluegill system), these cuckolder males may impede the evolution of MHC-mediated mate choice in a system. This may lend reason as to why I do not detect something similar in bluegill where females may not be actively choosing to mate with the smaller precocious cuckolders. Fu *et al*. (2001) found that female bluegill did dip more frequently in the presence of cuckolders, however this observation may be due to cuckolders actively choosing to encroach upon nests with especially fecund females are present. It is perhaps this lack of female choice in this system that has not

favoured MHC-based mate choice between females and cuckolders. However, this impediment may clearly be overcome in some cases, as for example in Atlantic salmon in which cuckolder-type males appear to mate indiscriminately with respect to the MHC, but female mating preferences still lead to an overall increase in MHC dissimilarity in their offspring (Consuegra & Garcia de Leaniz 2008). In the Atlantic salmon mating system, it also appears that the MHC is not involved in extra-pair mating with cuckolders (Lehnert *et al,*. 2018). Regardless, the role of extra-pair mating and particularly the presence of specialized cuckolder reproductive tactics in the evolution of MHC-mediated mating patterns remains a rich area for further enquiry.

4.3 Future directions

Although my analyses indicated that the presence of a null allele was unlikely to alter the conclusions of my study, identifying the null allele(s) at the MHC in bluegill would further support my conclusions and the characterization of MHC in bluegill. In the current study, multiple attempts were made to design functional primers. The final primer set used were the only functional primers that captured most of the peptide binding region. New primers would likely need to capture a longer region of the sequence in order to both detect any potential mutations within my primer sequence as well as still capture the variability within the sequence that I have seen.

Another avenue of further study would be to investigate what is maintaining the diversity of MHC class II alleles in bluegill. A positive, albeit not statistically significant, Tajima's D value indicates that balancing selection is likely occurring in this population. PAML models of positive selection were also better fits for my data than the neutral modes suggesting positive selection is occurring at the peptide binding region. As I did not detect mate choice for MHC that would promote allelic diversity in bluegill, direct selection acting

on the MHC is the most likely explanation for the strong indicators of positive selection on the MHC. Linking pathogen resistance to certain alleles from this population may also add some insight as to why these 10 alleles are maintained. In order to assess this, pathogen communities could be assessed within bluegill and associations between specific pathogen presence or total parasite load and MHC genotypes could be analyzed. Osborne *et al*. (2016) used zero-inflated negative binomial (ZINB) models to show an association between specific parasites and MHC alleles in Rio Grande silvery minnow (*Hybognathus amarus*). Infectivity trials could also be utilized to assess the effects of certain common pathogens on the survival of bluegill with different MHC alleles. Yang *et al.* (2016) challenged orange-spotted grouper with Singapore grouper iridovirus and found that a specific allele was associated with resistance to the pathogen. By using one of these methods, pathogen resistance could be linked to one or more alleles in bluegill and the effects of pathogen resistance could be linked to one or more alleles in bluegill and the effects of this on the maintenance of the observed MHC diversity could be analyzed.

Perhaps the most interesting area of further study with bluegill would be to find what, if anything, is driving mate choice in this population. For example, do female bluegill prefer to mate with larger males when presented with males of different sizes, as in other fish (Côte & Hunte 1989)? In this case, females are choosing for direct benefits as size in some fish can be correlated with parental care. While my study did not find evidence for assortative mating based on size, it is possible that female bluegill prefer to mate with larger males for the reason described above even while males show no size preference. Some females in other species of fish have also been shown to choose males based on nest quality (Sikkel 1995). While it cannot be characterized as a classical lek due to male parental care, it is possible that the bluegill mating colonies share some similarity to leks in that females are presented with a wide array

of male nests from which to choose. It may be that bluegill female prefer to mate with males with a higher nest quality as this again may be a marker of parental care quality. Alternatively, it may be that female bluegill are bet hedging and choose to mate with many males and in doing so are less choosy with respect to any of the previously mentioned markers of quality.

4.4 Conclusion

My results provide no evidence for MHC-dependent mate choice in bluegill. Neither MHC dissimilarity between mates nor maximizing offspring heterozygosity appeared to be factors that influenced mate choice. There was also no evidence of MHC-dependent mate choice in cuckolder-female pairings, albeit I captured only a small number of cuckolder males during mating so I cannot rule out MHC-mediated mate choice between cuckolders and females in bluegill. Further modelling work could address the influence of alternative reproductive tactics on the evolution and maintenance of MHC-mediated mate choice.

References

van der Aa LM, Levraud JP, Yahmi M, Lauret E, Briolat V, Herbomel P, Benmansour A, Boudinot P. 2009. A large new subset of TRIM genes highly diversified by duplication and positive selection in teleost fish. BMC Biol. 7:1–23. doi:10.1186/1741-7007-7-7.

Alatalo R V., Höglund J, Lundberg A. 1991. Lekking in the black grouse - A test of male viability. Nature. 352(6331):155–156. doi:10.1038/352155a0.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature. 364(6432):33–39. doi:10.1038/364033a0.

Cargnelli LM, Gross MR. 1996. The temporal dimension in fish recruitment: birth date, body size, and size-dependent survival in a sunfish (bluegill: *Lepomis macrochirus*). Can J Fish Aquat Sci. 53(2):360–367. doi:10.1139/cjfas-53-2-360.

Cohen S. 2002. Strong Positive Selection and Habitat-Specific Amino Acid Substitution Patterns in MHC from an Estuarine Fish Under Intense Pollution Stress. Mol Biol Evol. 19(11):1870–1880. doi:10.1093/oxfordjournals.molbev.a004011.

Consuegra S, Garcia de Leaniz C. 2008. MHC-mediated mate choice increases parasite resistance in salmon. Proc R Soc B Biol Sci. 275(1641):1397–1403. doi:10.1098/rspb.2008.0066.

Côte IM, Hunte W. 1989. Male and female mate choice in the redlip blenny: why bigger is better. Anim Behav. doi:10.1016/S0003-3472(89)80067-3.

Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L. 2007. Clinal variation in MHC diversity with temperature: Evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. Evolution (N Y). 61(9):2154–2164. doi:10.1111/j.1558-5646.2007.00178.x.

Evans ML, Neff BD. 2009. Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). Mol Ecol. 18(22):4716–4729. doi:10.1111/j.1365-294X.2009.04374.x.

Fernandes JMO, Ruangsri J, Kiron V. 2010. Atlantic cod piscidin and its diversification through positive selection. PLoS One. 5(3). doi:10.1371/journal.pone.0009501.

Fisher RA. 1915. The evolution of sexual preference. Eugen Rev. 7(3):184–92. doi:10.1109/61.25587.

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. J Evol Biol. 20(5):1859–1869. doi:10.1111/j.1420-9101.2007.01380.x.

Forsgren E, Karlsson A, Kvarnemo C. 1996. Female sand gobies gain direct benefits by choosing males with eggs in their nests. Behav Ecol Sociobiol. 39(2):91–96. doi:10.1007/s002650050270.

Fraser BA, Neff BD. 2009. MHC class IIB additive and non-additive effects on fitness measures in the guppy *Poecilia reticulata.* J Fish Biol. 75(9):2299–2312. doi:10.1111/j.1095-8649.2009.02449.x.

Fraser BA, Ramnarine IW, Neff BD. 2010a. Temporal variation at the MHC class IIb in wild populations of the guppy (*Poecilia reticulata*). Evolution (N Y). 64(7):2086–2096.

doi:10.1111/j.1558-5646.2010.00965.x.

Fraser BA, Ramnarine IW, Neff BD. 2010b. Selection at the MHC class IIB locus across guppy (*Poecilia reticulata*) populations. Heredity (Edinb). 104(2):155–167. doi:10.1038/hdy.2009.99.

Fu P, Neff BD, Gross MR. 2001. Tactic-specific success in sperm competition. Proc R Soc B Biol Sci. 268(1472):1105–1112. doi:10.1098/rspb.2001.1625.

Galan M, Guivier E, Caraux G, Charbonnel N, Cosson J-F. 2010. A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. BMC Genomics. 11(1):296. doi:10.1186/1471-2164-11-296.

Garner SR, Bortoluzzi RN, Heath DD, Neff BD. 2010. Sexual conflict inhibits female mate choice for major histocompatibility complex dissimilarity in Chinook salmon. Proc Biol Sci. 277(1683):885–894. doi:10.1098/rspb.2009.1639.

Garner SR, Neff BD. 2013. Alternative male reproductive tactics drive asymmetrical hybridization between sunfishes (*Lepomis* spp.). Biol Lett. 9:20130658. doi:10.1098/rsbl.2013.0658.

Gross MR. 1979. Cuckoldry in sunfishes (*Lepomis*: Centrarchidae). Can J Zool. 57:1507–1509.

Gross MR. 1982. Sneakers, satellites and parentals: Polymorphic mating strategies in North American sunfishes. Z Tierpsychol. 60:1–26. doi:10.1111/j.1439-0310.1982.tb01073.x.

Gross MR. 1991. Evolution of Alternative Reproductive Strategies: Frequency-Dependent Sexual Selection in Male Bluegill Sunfish. Philos Trans R Soc B Biol Sci. 332(1262):59–66.

doi:10.1098/rstb.1991.0033.

Gutierrez-Espeleta GA, Hedrick PW, Kalinowski ST, Garrigan D, Boyce WM. 2001. Is the decline of desert bighorn sheep from infectious disease the resulot of low MHC variation. Heredity (Edinb). 86(December 2000):439–450.

Hedrick PW. 2002. Pathogen resistance and genetic variation at MHC loci. Evolution. 56(10):1902–1908. doi:10.1111/j.0014-3820.2002.tb00116.x.

Hoover B, Nevitt G. 2016. Modeling the Importance of Sample Size in Relation to Error in MHC-Based Mate-Choice Studies on Natural Populations. Integr Comp Biol. 56(5):925–933. doi:10.1093/icb/icw105.

Houde AE. 1997. Sex, colour and mate choice in Guppies. Princeton University Press.

Kalinowski ST, Taper ML. 2006. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. Conserv Genet. 7(6):991–995. doi:10.1007/s10592-006-9134-9.

Kamiya T, O'Dwyer K, Westerdahl H, Senior a., Nakagawa S. 2014. A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. Mol Ecol. 23(21):5151–5163. doi:10.1111/mec.12934.

Karino K, Utagawa T, Shinjo S. 2005. Heritability of the algal-foraging ability: An indirect benefit of female mate preference for males' carotenoid-based coloration in the guppy, *Poecilia reticulata*. Behav Ecol Sociobiol. 59(1):1–5. doi:10.1007/s00265-005-0001-3.

Karlsson M, Westerdahl H. 2013. Characteristics of MHC class i genes in house sparrows passer domesticus as revealed by long cDNA transcripts and amplicon sequencing. J Mol Evol. 77(1–2):8–21. doi:10.1007/s00239-013-9575-y.

Keddy AC. 1986. Female mate choice in vervet monkeys (*Cercopithecus aethiops sabaeus*). Am J Primatol. 10(2):125–134. doi:10.1002/ajp.1350100204.

Kirkpatrick M. 1996. Good Genes and Direct Selection in the Evolution of Mating Preferences. Evolution (N Y). 50(6):2125. doi:10.2307/2410684.

Kotiaho JS, LeBas NR, Puurtinen M, Tomkins JL. 2008. On the resolution of the lek paradox. Trends Ecol Evol. 23(1):1–3. doi:10.1016/j.tree.2007.09.012.

Kuduk K, Babik W, Bellemain E, Valentini A, Zedrosser A, Taberlet P, Kindberg J, Swenson JE, Radwan J. 2014. No evidence for the effect of MHC on Male mating success in the brown bear. PLoS One. 9(12):2–4. doi:10.1371/journal.pone.0113414.

Kurtz J, Kalbe M, Aeschlimann PB, Häberli MA, Wegner KM, Reusch TBH, Milinski M. 2004. Major histocompatibility complex diversity influences parasite resistance and innate immunity in sticklebacks. Proc R Soc B Biol Sci. 271(1535):197–204. doi:10.1098/rspb.2003.2567.

Landry C, Garant D, Duchesne P, Bernatchez L. 2001. “Good genes as heterozygosity”: the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). Proc Biol Sci. 268(1473):1279–1285. doi:10.1098/rspb.2001.1659.

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. Proc Biol Sci. 268(1466):479–485. doi:10.1098/rspb.2000.1378.

Lehnert SJ, Helou L, Pitcher TE, Heath JW, Heath DD. 2018. Sperm competition, but not major histocompatibility divergence, drives differential fertilization success between alternative reproductive tactics in Chinook salmon. J Evol Biol. 31(1):88–97.

doi:10.1111/jeb.13199.

Lenz TL, Eizaguirre C, Scharsack JP, Kalbe M, Milinski M. 2009. Disentangling the role of MHC-dependent “good genes” and “compatible genes” in mate-choice decisions of three-spined sticklebacks Gasterosteus aculeatus under semi-natural conditions. J Fish Biol. 75(8):2122–2142. doi:10.1111/j.1095-8649.2009.02410.x.

Liu H-Y, Xue F, Gong J, Wan Q-H, Fang S-G. 2017. Limited polymorphism of the functional MHC class II B gene in the black-spotted frog (*Pelophylax nigromaculatus*) identified by locus-specific genotyping. Ecol Evol.(September 2016):1–9. doi:10.1002/ece3.3408.

McConnell TJ, Godwin UB, Norton SF, Nairn RS, Kazianis S, Morizot DC. 1998. Identification and mapping of two divergent, unlinked major histocompatibility complex class II B genes in Xiphophorus fishes. Genetics. 149(4):1921–34.

Milinski M. 2006. The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. Annu Rev Ecol Evol Syst. 37(1):159–186. doi:10.1146/annurev.ecolsys.37.091305.110242.

Murphy K. 2012. Janesway’s Immunobiology.

Neff BD. 2001. Genetic paternity analysis and breeding success in bluegill sunfish (*Lepomis macrochirus*). J Hered. 92(2):111–119. doi:10.1016/j.ejvs.2006.06.019.

Neff BD, Fu P, Gross MR. 2000. Microsatellite Multiplexing in Fish. Trans Am Fish Soc. 129(2):584–593. doi:10.1577/1548-8659(2000)129<0584:MMIF>2.0.CO;2.

Neff BD, Garner SR, Heath JW, Heath DD. 2008. The MHC and non-random mating in a

captive population of Chinook salmon. Heredity (Edinb). 101(2):175–185. doi:10.1038/hdy.2008.43.

Neff BD, Gross MR. 2001. Dynamic adjustment of parental care in response to perceived paternity. Proc R Soc B Biol Sci. 268(1476):1559–1565. doi:10.1098/rspb.2001.1678.

Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: An integrated framework for good genes and compatible genes. Mol Ecol. 14(1):19–38. doi:10.1111/j.1365-294X.2004.02395.x.

Osborne MJ, Pilger TJ, Lusk JD, Turner TF. 2016 Nov. Spatio-temporal variation in parasite communities maintains diversity at the major histocompatibility complex class IIβ in the endangered Rio Grande Silvery Minnow. Mol Ecol. doi:10.1111/mec.13936.

Partridge C, MacManes M, Knapp R, Neff B. 2015. Brain Transcriptional Profiles of Male Alternative Reproductive Tactics and Females in Bluegill Sunfish. bioRxiv.:025916. doi:10.1101/025916.

Penn DJ, Damjanovich K, Potts WK. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. Proc Natl Acad Sci U S A. 99(17):11260–11264. doi:10.1073/pnas.162006499.

Pitcher TE, Neff BD, Rodd FH, Rowe L. 2003. Multiple mating and sequential mate choice in guppies: Females trade up. Proc R Soc B Biol Sci. 270(1524):1623–1629. doi:10.1098/rspb.2002.2280.

Potts WK, Manning CJ, Wakeland EK. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype. Nature. 352(6336):619–621. doi:10.1038/352619a0.

Promerová M, Králová T, Bryjová A, Albrecht T, Bryja J. 2013. MHC Class IIB Exon 2 Polymorphism in the Grey Partridge (*Perdix perdix*) Is Shaped by Selection, Recombination and Gene Conversion. PLoS One. 8(7):1–10. doi:10.1371/journal.pone.0069135.

Promerová M, Vinkler M, Bryja J, Poláková R, Schnitzer J, Munclinger P, Albrecht T. 2011. Occurrence of extra-pair paternity is connected to social male's MHC-variability in the scarlet rosefinch *Carpodacus erythrinus*. J Avian Biol. 42(1):5–10. doi:10.1111/j.1600-048X.2010.05221.x.

Reusch TB, Häberli M a, Aeschlimann PB, Milinski M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. Nature. 414(6861):300–302. doi:10.1038/35104547.

Schwensow N, Fietz J, Dausmann K, Sommer S. 2008. MHC-associated mating strategies and the importance of overall genetic diversity in an obligate pair-living primate. Evol Ecol. 22(5):617–636. doi:10.1007/s10682-007-9186-4.

Sebastian A, Herdegen M, Migalska M, Radwan J. 2016. Amplisas: A web server for multilocus genotyping using next-generation amplicon sequencing data. Mol Ecol Resour. 16(2):498–510. doi:10.1111/1755-0998.12453.

Sikkel P. 1995. Effects of nest quality on male courtship and female spawning-site choice in an algal-nesting damselfish. Bull Mar Sci. 57(3):682–689.

Sin YW, Annavi G, Dugdale HL, Newman C, Burke T, MacDonald DW. 2014. Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (*Meles meles*). Mol Ecol. 23(20):5072–5088. doi:10.1111/mec.12917.

Smith SB, Webster MS, Holmes RT. 2005. The heterozygosity theory of extra-pair mate

choice in birds: A test and a cautionary note. J Avian Biol. 36(2):146–154. doi:10.1111/j.0908-8857.2005.03417.x.

Stoltz JA, Neff BD. 2006. Male size and mating tactic influence proximity to females during sperm competition in bluegill sunfish. Behav Ecol Sociobiol. 59(6):811–818. doi:10.1007/s00265-005-0127-3.

Tentelier C, Larrieu M, Aymes JC, Labonne J. 2011. Male antagonistic behaviour after spawning suggests paternal care in brown trout, Salmo trutta. Ecol Freshw Fish. 20(4):580–587. doi:10.1111/j.1600-0633.2011.00507.x.

Tregenza T, Wedell N. 2000. Genetic compatibility, mate choice and patterns of parentage: Invited Review. Mol Ecol. 9(8):1013–1027. doi:10.1046/j.1365-294x.2000.00964.x.

Trivers RL. 1972. Parental investment and sexual selection. Sex Sel descent man. 12(2):136–179. doi:10.1002/ajpa.1330400226.

Venney CJ, Johansson ML, Heath DD. 2016. Inbreeding effects on gene-specific DNA methylation among tissues of Chinook salmon. Mol Ecol. 25(18):4521–4533. doi:10.1111/mec.13777.

Worley K, Collet J, Spurgin LG, Cornwallis C, Pizzari T, Richardson DS. 2010. MHC heterozygosity and survival in red junglefowl. Mol Ecol. 19(15):3064–3075. doi:10.1111/j.1365-294X.2010.04724.x.

Yang M, Wei J, Li P, Wei S, Huang Y, Qin Q. 2016 Apr 4. MHC polymorphism and disease resistance to Singapore grouper iridovirus (SGIV) in the orange-spotted grouper, *Epinephelus coioides*. Sci Bull. doi:10.1007/s11434-016-1055-5.

Yang Z. 2005. Bayes Empirical Bayes Inference of Amino Acid Sites Under Positive Selection. Mol Biol Evol. 22(4):1107–1118. doi:10.1093/molbev/msi097.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol. 24(8):1586–1591. doi:10.1093/molbev/msm088.

Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 155(1):431–49. doi:10.1093/oxfordjournals.molbev.a003981.

Zahavi A. 1975. Mate selection-a selection for a handicap. J Theor Biol. 53(1):205–14. doi:10.1016/0022-5193(75)90111-3.

Zhao M, Wang Y, Shen H, Li C, Chen C, Luo Z, Wu H. 2013. Evolution by selection, recombination, and gene duplication in MHC class I genes of two *Rhacophoridae* species. BMC Evol Biol. 13(1):113. doi:10.1186/1471-2148-13-113.

Appendices

Ethics Statement

All experiments followed ethical guidelines from the Canadian Council on Animal Care as reviewed and approved by the Animal Use Subcommittees at the University of Western Ontario (protocol # 2010-214).

Curriculum Vitae

Name:	Lucas Thome Silveira
Post Secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 2009-2014 BSc Hon
	The University of Western Ontario London, Ontario, Canada 2014-2019 MSc
Related Work Experience:	Teaching Assistant The University of Western Ontario 2014-2017