

Is Promiscuity Associated with Enhanced Selection on MHC-DQ α in Mice (genus *Peromyscus*)?

Matthew D. MacManes^{*‡}, Eileen A. Lacey

Department of Integrative Biology, Museum of Vertebrate Zoology, University of California, Berkeley, California, United States of America

Abstract

Reproductive behavior may play an important role in shaping selection on Major Histocompatibility Complex (MHC) genes. For example, the number of sexual partners that an individual has may affect exposure to sexually transmitted pathogens, with more partners leading to greater exposure and, hence, potentially greater selection for variation at MHC loci. To explore this hypothesis, we examined the strength of selection on exon 2 of the MHC-DQ α locus in two species of *Peromyscus*. While the California mouse (*P. californicus*) is characterized by lifetime social and genetic monogamy, the deer mouse (*P. maniculatus*) is socially and genetically promiscuous; consistent with these differences in mating behavior, the diversity of bacteria present within the reproductive tracts of females is significantly greater for *P. maniculatus*. To test the prediction that more reproductive partners and exposure to a greater range of sexually transmitted pathogens are associated with enhanced diversifying selection on genes responsible for immune function, we compared patterns and levels of diversity at the Class II MHC-DQ α locus in sympatric populations of *P. maniculatus* and *P. californicus*. Using likelihood based analyses, we show that selection is enhanced in the promiscuous *P. maniculatus*. This study is the first to compare the strength of selection in wild sympatric rodents with known differences in pathogen milieu.

Citation: MacManes MD, Lacey EA (2012) Is Promiscuity Associated with Enhanced Selection on MHC-DQ α in Mice (genus *Peromyscus*)? PLoS ONE 7(5): e37562. doi:10.1371/journal.pone.0037562

Editor: Nadia Singh, North Carolina State University, United States of America

Received: November 23, 2011; **Accepted:** April 25, 2012; **Published:** May 23, 2012

Copyright: © 2012 MacManes and Lacey. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a NSF-DDIG (IOS0909798), as well as by awards from the Museum of Vertebrate Zoology and Department of Integrative Biology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: MDMs' current postdoctoral adviser is Michael Eisen, a co-founder of PLoS and a member of the board of directors. This does not alter the authors' adherence to all PLoS ONE policies.

* E-mail: macmanes@gmail.com

‡ Current address: California Institute for Quantitative Biosciences, University of California, Berkeley, California, United States of America

Introduction

Understanding how natural selection shapes adaptive variation in populations of free-living animals has long been a focus of evolutionary biology [1–3]. Genes of the Major Histocompatibility Complex (MHC) provide logical targets for such studies because of the relatively well-understood function of these loci, which code for proteins that participate in recognition of and response to foreign antigens [4,5]. Although the extreme diversity characteristic of MHC loci is well documented, identifying the underlying bases for diversifying selection on these genes has been more challenging. While multiple studies have attributed such selection on MHC loci to behavioral phenomenon such as mate choice [6–9], others have linked extreme allelic diversity at these genes to enhanced pathogen exposure [10–12]. Distinguishing between these alternatives provides an important opportunity to explore interactions between adaptively significant genotypic variation and the dynamic and variable selective pressures acting on free-living vertebrates.

One aspect of a species' biology that may significantly impact exposure to pathogens is its mating system, specifically the number of partners with which an individual typically mates during either a single round of reproduction or over the course of its lifetime. Relationships between sexual behavior and sexually transmitted disease have been extensively studied in humans from both empirical and theoretical perspectives [13–16]. These studies have revealed that the number of concurrent partners [17,18] and the

amount of time between sexual encounters [19] are critical in determining the overall population level of infection. For example, in populations characterized by extreme promiscuity (i.e. high concurrency), rates of sexually transmitted pathogens tend to be high [20,21]. In contrast, members of populations characterized by lifetime monogamy (i.e., low concurrency) have a much lower risk of contracting sexually transmitted diseases. Despite logically compelling links between sexual behavior, rates of pathogen transmission, and selection on immunogenes, few analyses have undertaken a comprehensive exploration of these relationships. As a result, studies that explore directly interactions among behavior, pathogen exposure, and selection on MHC genes are required.

Mice of the genus *Peromyscus* provide an ideal opportunity to explore these relationships in natural populations of vertebrates. Because these animals have long been targets of study, many aspects of their behavior and ecology are known [22–28]. With regard to mating system, the genus includes promiscuous [29,30] and polygynous [31] species, as well as two of the few mammalian species demonstrated to be socially and genetically monogamous [32,33]. In central coastal California, *P. californicus*—which has been shown in multiple studies to exhibit lifetime social and genetic monogamy [34]—is sympatric across its range with the socially and genetically promiscuous *P. maniculatus* [29]. Consistent with this difference in mating system, the reproductive tracts of female *P. maniculatus* are characterized by significantly more diverse bacterial communities than are those of female *P. californi-*

cus [35]. These pronounced differences in reproductive behavior and potential pathogen exposure among otherwise ecologically similar congeners offer a rare opportunity to explore the effects of differences in mating system – specifically, the contrast between monogamy and promiscuity – on selection on MHC genes.

To characterize the relationship between mating system and selection on MHC genes in *Peromyscus*, we compared patterns of allelic and genotypic diversity at the Class II DQ α locus in *P. californicus* and *P. maniculatus*. Given that promiscuity in the latter species is associated with an increased diversity of vaginal bacteria [35], we predicted that the greater number of reproductive partners per individual in *P. maniculatus* would also be related to enhanced diversifying selection at the DQ α locus. We tested this prediction using data obtained from sympatric populations of these species whose mating systems have been previously characterized. This study is one of the first to describe the relationship between mating system and selection on MHC genes in natural populations of vertebrates. The comparative approach employed is particularly compelling in that it minimizes the effects of environmental differences on selection at these loci. As a result, our analyses provide important new insights into the effects of selection apparent relationships between mating system and selection on MHC genes in free-living vertebrates.

Materials and Methods

Study Populations and Tissue Sampling

Field research was conducted on Landels-Hill Big Creek Reserve (36.011156°, -121.518644°), which is located approximately 30 km south of Big Sur, Monterey County, California, USA, during June-August 2009. Big Creek Reserve is a part of the University of California Natural Reserve System and consists of 15.57 km² of coastal scrub, redwood forest, and oak woodlands. This locality was selected because both *P. maniculatus* and *P. californicus* are abundant and can be captured in the same habitats using a single trapping grid.

To facilitate handling, all animals captured were anesthetized with Isoflurane. After light anesthesia was induced, a 3 mm³ section of the distal portion of the right ear pinna was removed using sterile scissors. Tissue was stored in 70% ethanol and frozen at -20°C within 2 hours of collection. To avoid potential resampling of the same individuals, each animal was then individually marked using a small numbered ear tag (Monel 1005, National Band & Tag Company) placed in the remaining portion of the right pinna. All procedures were approved by the University of California, Berkeley Animal Care and Use Committee and were in accord with the guidelines of the American Society of Mammalogists [36].

DNA Extraction and General PCR Procedures

We extracted genomic DNA from tissue samples from 20 individuals per species using a salt extraction method [37]; the resultant extract was tested for purity and concentration using the Nanodrop system (Thermo Scientific). From this DNA stock solution, 50 ng/ μ L dilutions were prepared for use in all PCR reactions.

Comparisons with Neutral Markers

Neutral processes associated with differences in effective population size (e.g. genetic drift) can affect how genes respond to natural selection. To ensure that differences in effective population size were not responsible for observed differences in selection regime, we genotyped all individuals sampled at 8 previously identified microsatellite loci using PCR conditions

described by Chirhart et al., [38] and Mullen et al., [39]. Genotyping was accomplished on an ABI 3730 using the size standard LIZ500; allele sizes were scored using the program GeneMapper. The resultant dataset was screened for null alleles, linkage and Hardy-Weinberg disequilibrium using the program Genepop [40,41]. Summary statistics on variability were generated in the program Arlequin [42]. Effective population size was estimated with the program ONeSAMP [43]. The sensitivity of the results to the user-assigned upper and lower limits of N_e was assessed by analyzing the data using a variety of values; analyses were run using lower bounds of 2–20 and upper bounds of 30–100.

MHC Genotyping

Polymerase chain reaction amplification targeting a 246 base pair region of exon 2 of the DQ α locus was carried out using *Peromyscus*-specific primers [44]. Because direct sequencing of highly polymorphic nuclear genes can be confounded by heterozygosity as well as by both insertion and deletion mutations, successful MHC amplicons were TA cloned using a TOPO TA Cloning Kit for Sequencing (Invitrogen, K4575-01) following the manufacturers instructions. A minimum of 8 positive colonies per individual was PCR amplified using the vector specific primers M13R and M13F and conditions specified by the manufacturer. Successful PCR reactions were cleaned using ExoSap (USB Corporation, #78250) and cycle sequenced using ABI BigDye Terminator v3.1 Cycle Sequencing Kits (Applied Biosystems, #4337456). Clones were sequenced in both forward and reverse directions on an ABI3730 automated DNA sequencer. Sequences were edited and assembled using the program Geneious 5.53 [45]. This process included the removal of primer sequences as well as nucleotides falling below a specified quality threshold (PHRED <20).

The number of clones sequenced per individual was based on a simple probabilistic model assuming unbiased PCR amplification and recovery of alleles; under a binomial distribution, sequencing of 8 positive clones per individual should have detected both alleles in a heterozygous individual with a probability of 0.992. An individual was considered heterozygous only if each allele sequence was detected in at least 2 distinct clones. Additionally, sequences were screened for the presence of recombination using the GARD [49] and SBR [50] algorithms contained in the program HyPhy, as well as with the PERMUTE code contained within the program omegaMap [46].

To assess the potential amplification of pseudogenes, coding sequences were translated and checked for the presence of an open reading frame and absence of stop codons. In addition, we compared our sequences to those of closely related species contained in GenBank in an attempt to identify unexpectedly high levels of sequence divergence that could reflect amplification of non-coding pseudogenes. To assess the possibility of >2 alleles per individual (indicative of multiple copies of the DQ α locus), we sequenced 24 clones containing the correct insert size from each of 8 randomly chosen individuals per species. Using the same reasoning outlined above, this procedure should have detected a 3rd allele (indicative of multiple copies of the DQ α locus) with a probability of 0.9999.

Tests of Neutrality and Gene Diversity

Estimates of Tajima's D, Fu's Fs, Pi and haplotype diversity were calculated using the program Arlequin [42] and plotted in Fig. 1. The first 2 of these statistics use the site-frequency spectrum of nucleotide polymorphism to test for departures from neutrality [49,50]. Tajima's D and Fu's Fs were calculated for the whole

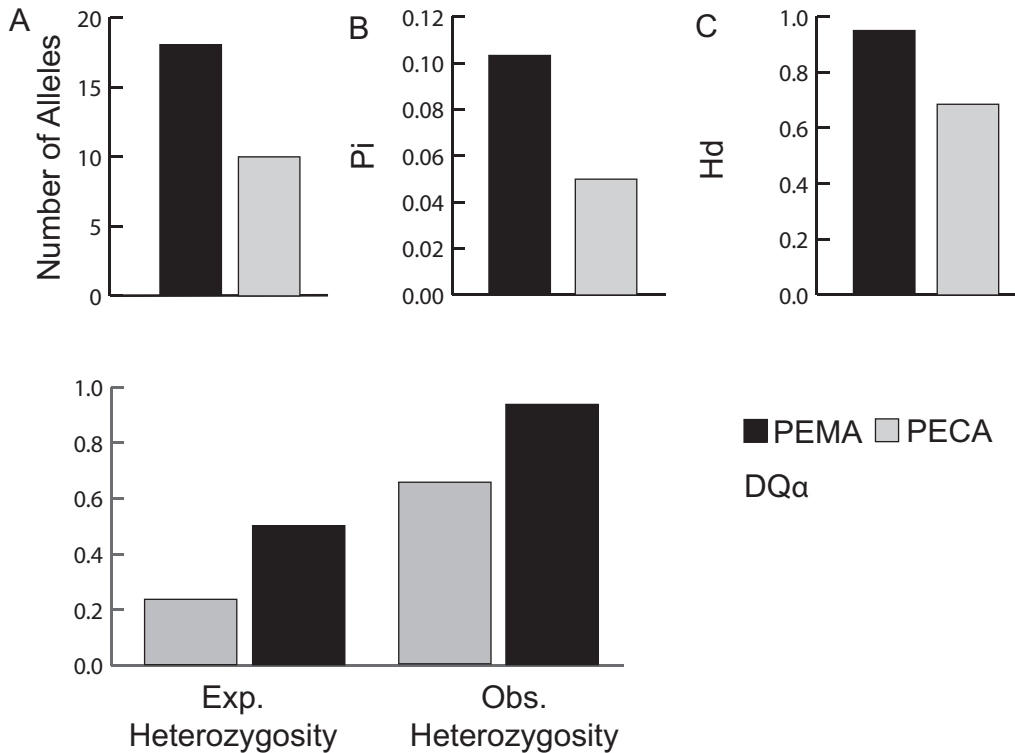


Figure 1. Comparisons of variability at the DQ α locus. A. Number of alleles detected for each species. B. Estimated values for Pi for each species. C. Estimated values for haplotype diversity for each species. D. Observed and expected values of heterozygosity in both species. For both species, Hardy-Weinberg tests resulted in values of $P=0.000$. In all cases, indices are greater for *P. maniculatus* than for *P. californicus*. doi:10.1371/journal.pone.0037562.g001

exon in an attempt to identify deviations from neutral expectations. In addition, the McDonald-Kreitman [47] test of neutrality was attempted in the program DNAsp [48], but no test statistic was calculated as the dataset lacked fixed differences between study species. Observed and expected heterozygosities were calculated and heterozygote deficiency was assessed using the Hardy-Weinberg Exact test implemented in Genepop.

Alignment and MHC Gene Tree Construction

Because many algorithms for the detection of selection are based on comparative phylogenetic methods [51,52], gene trees were constructed for DQ α . The outgroup species consisted of a single DQ α allele from 2 rodents—*Cricetulus barabensis* (FJ209306) and *Microtus arvalis* (DQ137813). Alignments were constructed using the program Muscle [53]. Sequences were partitioned by codon position and analyses were run in MrBayes [54], using the HKY85 model of molecular evolution, until model convergence was reached. A 50% consensus tree was built after discarding the first $1/3^{\text{rd}}$ of parameter values and trees as burn in.

Analyses of Selection

To examine evidence of selection on the DQ α locus, we implemented two distinct approaches. First, the subroutine codeml of the program PAML4.5 [51] was run. PAML employs a classic, phylogenetically-based maximum likelihood approach to detect the signature of natural selection using a series of likelihood ratio tests. Specifically, this procedure compares a series of models that allow omega ($=\omega$, equivalent to d_N/d_S) to vary between sites but not between lineages. Four independent models of molecular evolution were used to construct two likelihood ratio tests. M1a

(Nearly Neutral) allows 2 site classes, one with $\omega = 1$ and one with $0 < \omega < 1$. This model is compared with M2a, which is identical to M1a, but with the addition of another site class where $\omega > 1$. M7 (beta) assumes a beta distribution (limited to $0 < \omega < 1$) of codon sites, and is compared to M8, which is identical to M7, but with the addition of another site class where $\omega > 1$. For each test (M1a vs. M2a and M7 vs. M8), log likelihood values were calculated and compared using a likelihood ratio test.

Because we were interested in using PAML (typically employed for inter-specific comparisons) as described above on our population level dataset, we randomly selected a single allele from each focal species for analysis of natural selection using a python script that employed a random sampling with replacement strategy. We repeated this jack-knifing sampling procedure 1000 times, creating 1000 datasets each containing 4 sequences (two focal species, two outgroup species), thus ensuring that all results were robust to the effects of specific alleles or combinations of alleles. A tree was constructed as described above for each of the 1000 four-species alignments, after which this tree was used for the PAML analyses (1000 datasets x 4 PAML models). The distribution of omega, the results of the likelihood ratio tests constructed for each of the 1000 replicates, and the results of the two tests (M1a vs. M2a and M7 vs. M8) were plotted (Fig 2a, b, c).

A major shortcoming of PAML is its inability to account for the effects of recombination when identifying signatures of natural selection. Because high levels of recombination can be associated with a high false-positive rate [55], we chose to analyze our data in an additional framework, where recombination, if present, could be explicitly considered. The program omegaMap uses a population genetic approximation to the coalescent and treats uncertainty in gene genealogy as a nuisance parameter, all within

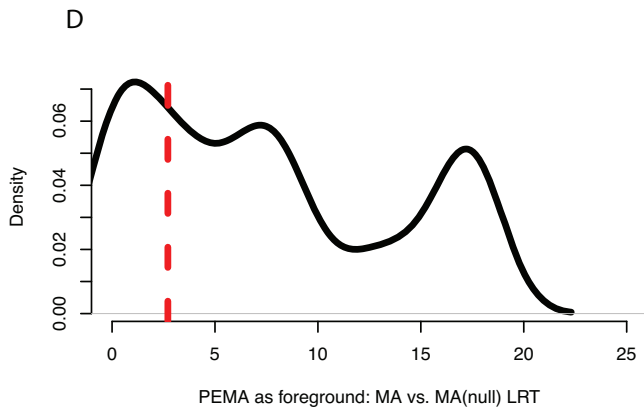
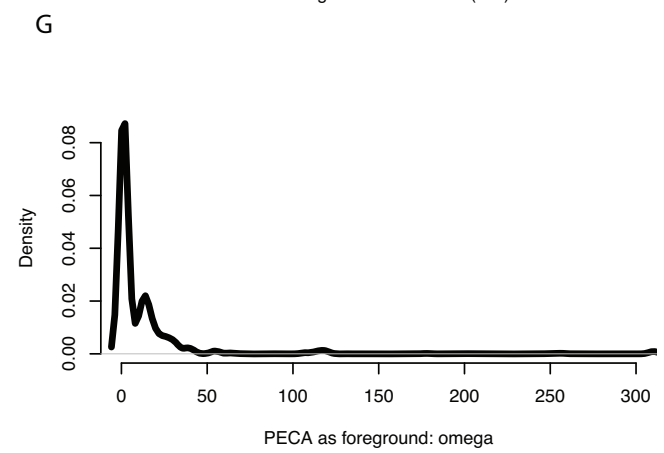
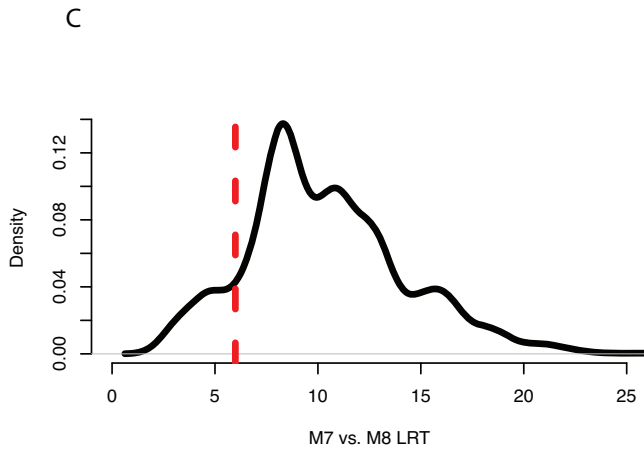
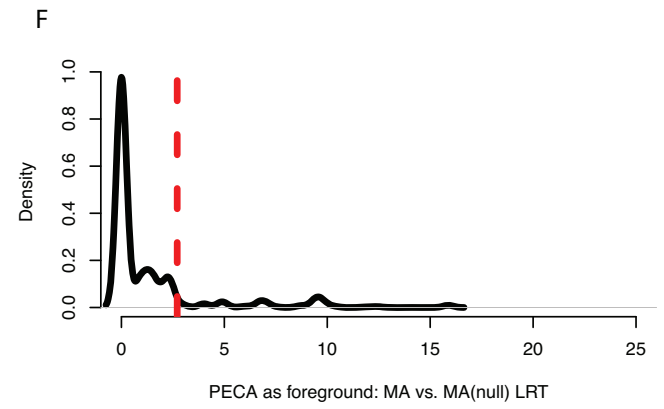
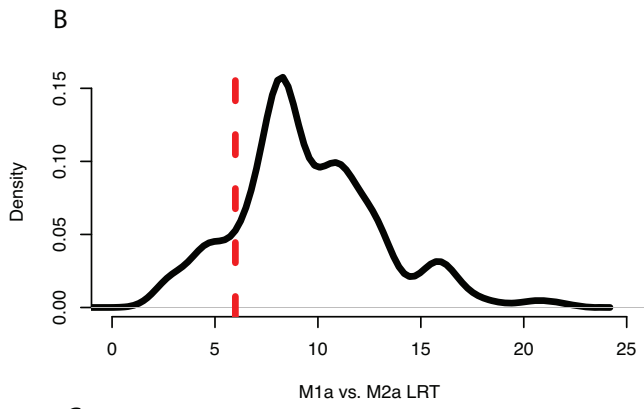
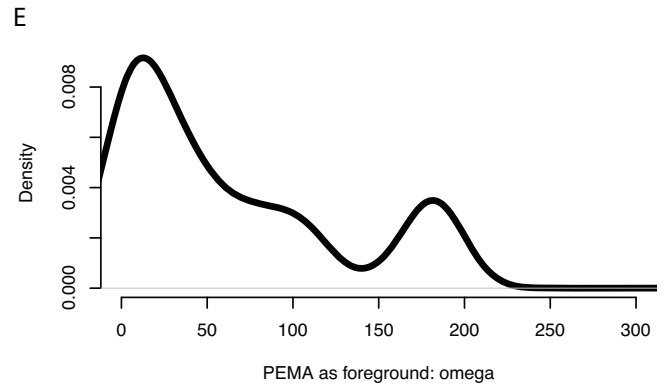
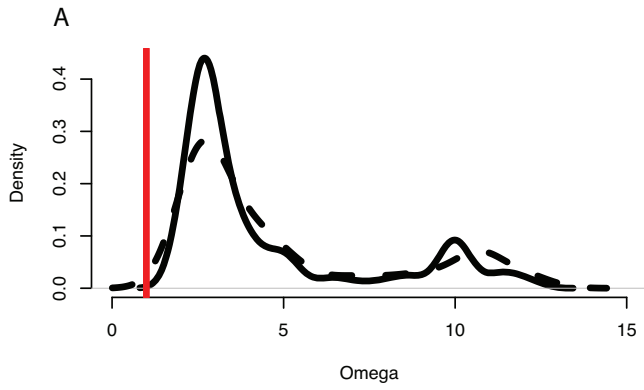


Figure 2. Plots of the distribution of selection parameters and statistical tests generated in the program PAML using the 1000 replicate datasets. Figure 2A is the distribution of omega. Dashed line is the estimate in model 2a, while the solid lines represent the estimate in model 8. 2B, 2C, 2D, 2F are the distributions of the LRT for M1a vs. M2a, M7 vs. M8, branch site tests (PEMA foreground) and branch site test (PECA foreground) respectively. The dashed red line indicates the critical value with $df=2$, while the solid red line is placed at $\omega=1$. Figure 2E and 2G plot the distribution of omega in the branch site test when PEMA or PECA are placed in the foreground position.
doi:10.1371/journal.pone.0037562.g002

a Bayesian framework that results in inferences that are robust to recombination, making this approach particularly useful for the analysis of datasets like our population level samples of DQ α variability.

Within omegaMap, we calculated selection parameters and identified specific codons thought to be subject to enhanced selection. All analyses were comprised of two separate MCMC chains of 1 million generations in length, with sampling from the likelihood surface every 100th generation. Settings recommended in the software documentation were used, except for omega, which was modeled as either a constant variable when calculating a global estimate of omega, or as an independent variable when estimates of selection in specific codons were sought. Separate analyses were run for *P. maniculatus* and *P. californicus*. These two independent runs were compared to each other to assure that model convergence had been reached.

Comparing Selection Across Study Species

To determine if selection has acted differently in *P. maniculatus* relative to *P. californicus*, the branch site model (model A) of diversifying selection [56,57] was implemented within PAML. This model allows the user to partition the action of selection between a foreground lineage, where selection is hypothesized to be enhanced relative to a set of background lineages. The model is tested in the LRT framework against the null model in which omega is constrained to acting with equal strength in foreground and background lineages. For this part of the study, as part of separate program executions, we set either the monogamous *P. californicus* or the promiscuous *P. maniculatus* as foreground lineages. Because the dataset consisted of multiple alleles per species, we used an approach identical to that described above, randomly selecting an allele from each species (as well as the two outgroup species) and repeating this procedure to create 1000 datasets per species which were then run in the program PAML. Results of the LRTs for each test as well as the distribution of foreground omega values were plotted (Fig. 2d, e, f, g).

Results

Diversity of Microsatellite Markers

Variation at the 8 microsatellite markers examined was consistently greater for *P. maniculatus*. Mean number of alleles per locus for *P. californicus* was 7.7 (range 6–11) while the mean for *P. maniculatus* was 10.2 (range 7–14). Mean heterozygosity was 0.80 and 0.88, respectively.

Estimates of effective population size based on the microsatellite data revealed that N_e did not differ markedly between the study populations. The median effective population size for *P. maniculatus* was 24.2 (95% CI 21.5–27.5); for *P. californicus*, this value was 21.5 (95% CI 19.1–24.1). A comparable magnitude of difference was obtained regardless of the specific upper and lower bounds of N_e used for estimation.

Diversity at MHC-DQ α

As with the neutral markers examined, interspecific differences in diversity were evident for MHC-DQ α . In total, 31 distinct DQ α alleles were detected (Genbank JN703316–JN703377). Specifically,

18 alleles were detected for *P. maniculatus* and 14 alleles were detected for *P. californicus*; one of these alleles was shared between the study species. In no case did we find evidence of more than 2 alleles per individual for the DQ α locus. Effectively no evidence of recombination was found in our dataset. In the program HyPhy, 96 and 97 potential breakpoints (points of recombination) were analyzed (GARD and SBR tests, respectively). No significant improvement in cAIC scores was observed when comparing a model that included a recombination event at any breakpoint over a model lacking recombination. Additionally, the PERMUTE code within omegaMap revealed no evidence of recombination in *P. maniculatus*, although a single test – the negative correlation between LD (measured as r^2 [58]) and physical distance – did suggest a recombination event ($p = 2.5 \times 10^{-5}$) in *P. californicus*.

The number of alleles, P_i , haplotype diversity, and expected and observed heterozygosity were greater for *P. maniculatus* (Fig. 1). Thus, similar to the neutral microsatellite markers examined, diversity at DQ α was generally greater for the promiscuous *P. maniculatus*. For both study species, Hardy-Weinberg tests revealed significant heterozygote excess at exon 2 of the DQ α locus (both $p < 0.001$). However, neither Tajima's D nor Fu's F_s provided evidence of significant departures from neutral expectations in either species (PECA: $D = -0.79$, $F_s = 2.8$; PEMA: $D = 0.31$, $F_s = 1.6$, all $p > 0.05$).

Analyses of nucleotide variation at the DQ α locus revealed evidence of selection on this gene. Specifically, likelihood ratio tests of models M1a versus M2a and M7 versus M8 suggested that diversifying selection has strongly influenced patterns of molecular diversity at the DQ α locus. Regarding the comparison between M1a and M2a, the null model of neutral evolution was rejected in 870 of the 1000 tests with $p < 0.05$ (Fig. 2b). For the comparison M7 vs. M8, the null model was rejected in 899 of 1000 tests with $p < 0.05$ (Fig. 2c). The distribution of omega calculated using the 1000 replicate datasets was plotted (Fig. 2a). The mean omega calculated using M2 was 4.92, while the mean using M8 was 4.49. Using omegaMap, the posterior probability of positive selection in PEMA was .87 ($\omega = 1.29$), while the in PECA it was .86 ($\omega = 1.40$). The proportion of sites under significant diversifying selection ($\omega > 1$, Bayesian posterior probability $> .95$) was 0.207 in *P. maniculatus* and 0.146 in *P. californicus* (Table 1). In contrast, when the consensus sequence created for each species was used, no test in either software package demonstrated significant evidence of selection. We believe that this result is likely related to the analytical process through which the consensus is created (collapsing variation at individual nucleotides). Although for the majority of sites there is little effect, this process may blunt the signal at very variable, positively selected sites.

Branch site tests of lineage-specific selection strongly supported the hypothesis of enhanced selection in promiscuous species. When assigning the promiscuous *P. maniculatus* to the foreground lineage, 745 of the 1000 datasets run showed strong evidence of lineage-specific selection (Fig 2d), with selection enhanced in the promiscuous species at $p < 0.05$. In contrast, when setting the monogamous *P. californicus* as the foreground lineage, less than 5% of all tests were significant (Fig. 2f). The distribution of omega in foreground lineages differed markedly between analyses, with a large proportion of the tests resulting in $\omega > 1$ with

Table 1. Estimates of diversifying selection on exon 2 of the MHC DQ α locus.

	Gene	n	Nc	dN/dS	Selection parameters	Pos. Seln omegaMap
PEMA	DQ α	18	82	1.29 (.82–1.99)	$p_s = .207$ $w_s = 14.6$	HPD = .87
PECA	DQ α	14	82	1.40 (.76–2.53)	$p_s = .146$ $w_s = 29.1$	HPD = .86

PEMA = *P. maniculatus*, PECA = *P. californicus*. n = number of alleles recovered, N_c = number of codons, selection parameters. d_N/d_S averaged across all sites. w_s = omega at the proportion of sites under diversifying selection, p_s = the proportion of sites under selection. Pos. Seln. Indicates the Bayesian posterior probability that positive selection is occurring.

doi:10.1371/journal.pone.0037562.t001

promiscuity in the foreground (Fig. 2e) versus a very small proportion of tests with this value when monogamy in the foreground (Fig. 2g).

Codon level evidence of selection—Analyses of codon level variation revealed 19 codons (n = 17 for *P. maniculatus*, n = 12 for *P. californicus*) that were identified by omegaMap as being subject to significant diversifying selection (Bayesian posterior probability >.95). Eight of these codons were identified as subject to significant selection in both study species (Table 2), all of which are known peptide binding sites in other mammalian species [59].

Table 2. Codons in exon 2 of the DQ α locus inferred to be under significant positive selection in each study species.

Codon Site	Species		Function
	PEMA	PECA	
8		10.1	
9	16.0	24.7	PBR
10	6.1		
22	9.3	35.4	PBR
29	14.0	16.3	PBR
40	5.3		
50	45.3	49.9	PBR
56	15.9	66.8	
59	8.7		
60	5.2	11.9	PBR
63	10.2		PBR
64	30.0	47.7	PBR
66	12.9		PBR
67	10.9		PBR
69		5.1	
70	6.0		PBR
71	36.5	44.2	PBR
72	9.4	13.1	
77	6.8	24.3	PBR

Numbers represent values for d_N/d_S inferred by omegaMap. Codons experiencing significant positive selection in both species are highlighted by red numbers. PBR indicates sites that are thought to directly interact with pathogens.

doi:10.1371/journal.pone.0037562.t002

Discussion

A growing body of literature suggests that the evolution of MHC genes is influenced by a complex suite of selective [60–62] and neutral processes [63–65]. This complexity makes it challenging to infer the underlying reasons for patterns of MHC variation, particularly in natural populations. A priori, we had predicted (based upon interspecific differences in mating system and diversity of vaginal bacteria) that diversifying selection should be stronger in the promiscuous *P. maniculatus*. Indeed, our analyses strongly support this hypothesis. For the branch site tests, both the number of significant tests and the distribution of omega in the foreground lineage were much larger when the promiscuous *P. maniculatus* was identified as the foreground species, suggested enhanced selection on DQ α in this species compared to the monogamous *P. californicus*. This result is also supported by codon-level analyses, in which more codons were subject to diversifying selection in *P. maniculatus* than in *P. californicus*. Taken together, these outcomes provide convincing evidence that diversifying selection on the DQ α locus is enhanced in the promiscuous *P. maniculatus*.

Although several methods have been proposed for determining whether recombination has altered the signal of natural selection in DNA sequence data [46,55,66], analyses of population-level data that may be subject to recombination remain challenging. In our dataset, only weak evidence for recombination was detected. Whether this represents a lack of statistical power or the absence of biologically relevant recombination is unclear. Despite this outcome, we believe that our analyses are robust to the effects of recombination. The outcomes of all tests for selection, including omegaMap (which explicitly considers recombination) and the LRT (M7 vs. M8) conducted within PAML (reported to be more robust to recombination than other tests in PAML [55]) were concordant in identifying enhanced selection in the promiscuous study species. Given that recombination, which has been purported to be generator of allelic diversity [67,68], was only weakly evident in the background lineage for our branch site tests, we believe that the results of our analyses were not confounded by recombination and are instead indicative of the relative intensity of selection on the two study species.

Selection on specific codons—Codon-level analyses revealed that 30% more codons were under significant selection in *P. maniculatus* than in *P. californicus*, suggesting that selection on the DQ α locus is enhanced in the promiscuous species relative to its monogamous congener. Fifteen codons in exon 2 of the DQ α locus are thought to interact directly with antigens [59]; our analyses indicated that 12 of these codons were subject to diversifying selection in at least one of our study species. This concordance between sites that were subject to selection and the presumed functional role of these codons suggests that selection on immune response has contrib-

uted to variation at the DQ α locus in our study animals. Interestingly, the specific codons identified as being subject to selection differed between the study species. Possible explanations for this variation include interspecific differences in codon function [69] and response to a species-specific pathogen community [70]. Use of emerging genomic tools such as site-directed mutagenesis should help to elucidate the reasons for these differences in the specific DQ α codons subject to diversifying selection in *P. maniculatus* and *P. californicus*.

The role of demography—Natural selection occurs against the backdrop of demography and thus demographic parameters may play a critical role in shaping selection on MHC genes. Both historical (e.g., population expansion) and contemporary (e.g., dispersal) demographic processes can alter the nature of selection [71,72]. Additionally, differences in effective population size (N_e) can influence the strength of natural selection [63–65]. As a result, demography must be considered when comparing patterns or estimates of selection for different populations. In our study, effective population size was similar for the two study species. Nevertheless, other demographic factors may skew the signal of natural selection in these animals, including differences in mating system [73,74]. In particular, restricting reproduction to pair bonded individuals - a characteristic of monogamous species like *P. californicus* - is expected to lower effective population sizes relative to non-monogamous species [75–77], which may result in a ‘blunting’ of the effects of selection, or even a scenario in which selection operates within the context of relaxed constraint. Interestingly, we did not detect significant departures from neutrality using standard population genetic tests (e.g. Tajima’s D and Fu’s Fs); evidence of selection was detected only when using divergence-based analyses (omegaMap and PAML). Although the reason for this discrepancy in outcomes is unknown, we suspect that this disparity is related to the way in which each test incorporates signals of the historical demography of a population.

Potential environmental effects—Environmental factors are thought to play a significant role in shaping MHC variation, including spatial variation in MHC polymorphism [78–81]. By focusing on sympatric populations of mice, our study sought to minimize the effects of environmental differences on pathogen exposure. The study species did, however, vary markedly with respect to mating system. The greater number of reproductive partners per individual in *P. maniculatus* was associated with a greater diversity of vaginal bacteria in females of this species [35]. Number of reproductive partners is not the only aspect of reproductive behavior that may influence pathogen exposure and thus selection on MHC loci. For example, although contacts with reproductive partners are expected to be more numerous in *P. maniculatus*, these contacts are believed to be brief, with no extended periods of close proximity or nest sharing by adults. In contrast, *P. californicus* form extended male-female pair bonds, with

members of a pair nesting together throughout the year [82,83]. Time spent with conspecifics influences an animal’s risk of infection by some pathogens [84,85] and thus even if exposure to sexually transmitted pathogens is reduced in the monogamous *P. californicus*, exposure to pathogens with other modes of transmission may be enhanced due to the prolonged contact between paired individuals. Although the analyses presented here do not address the relative importance of sexual versus social transmission, the finding that selection is enhanced in promiscuous species (e.g. [86]) may shed light on the relative importance of these distinct modes of pathogen transmission. Future work that quantifies socially transmitted pathogens in our study species will clarify the impacts of sexual versus social contact on pathogen exposure and selection on MHC loci.

To assess the generality of our findings, comparative tests of the relationship between promiscuity and selection on immunogenes in mammals should be performed. Opportunities for such comparative studies, however, are limited since the necessary contrast between promiscuity and monogamy in closely related species is rare, given that monogamy is thought to occur in less than 5% of mammalian species [87]. Further, a truly replicate study would control for environmental exposure (i.e. sympatry), a factor that we believe is important when working with natural populations. Although laboratory studies of the relationship between mating system and selection on immunogenes are possible, both important behavioral parameters (e.g. mating behavior) and pathogen exposure are likely to vary between wild and captive animals [88], thereby likely altering the relationship under analysis. Future work may begin with the systematic identification of monogamous species whose congeners are both closely related and promiscuous. Although the distribution of mammalian monogamy suggests that these comparisons are rare, researchers interested in determining the generality of our finding may find them important foci for the study of the relationship between sexual behavior and selection on immunogenes.

Acknowledgments

Computational resources were supported in part by the National Science Foundation through TeraGrid resources provided by The Texas Advanced Computing Center under grant number TG-IBN100014. The paper was greatly improved by suggestions from Sonal Singhal and several anonymous reviewers. Lastly, MDM would like to thank his wife Kate and his children, whose support has been unwavering.

Author Contributions

Conceived and designed the experiments: MDM. Performed the experiments: MDM. Analyzed the data: MDM. Contributed reagents/materials/analysis tools: MDM. Wrote the paper: MDM EAL.

References

1. Endler J (1986) Natural Selection in the Wild. Princeton University Press.
2. Darwin C (1859) On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life. 1st ed. London: Murray.
3. Reznick D, Butler M, Rodd F, Ross P (1996) Life-history evolution in guppies (*Poecilia reticulata*) 6. Differential mortality as a mechanism for natural selection. *Evolution* 50: 1651–1660.
4. Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *P R Soc B* 277: 979–988. doi:10.1098/rspb.2009.2084.
5. van Oosterhout C (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proc Biol Sci* 276: 657–665. doi:10.1098/rspb.2008.1299.
6. Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G (2006) Complex MHC-based mate choice in a wild passerine. *P R Soc B* 273: 1111–1116. doi:10.1098/rspb.2005.3325.
7. Consuegra S, Garcia de Leaniz C (2008) MHC-mediated mate choice increases parasite resistance in salmon. *Proc Biol Sci* 275: 1397–1403. doi:10.1098/rspb.2008.0066.
8. Schwensow N, Eberle M, Sommer S (2008) Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *P R Soc B* 275: 555–564. doi:10.1098/rspb.2007.1433.
9. Zelano B, Edwards SV (2002) An MHC component to kin recognition and mate choice in birds: Predictions, progress, and prospects. *Am Nat* 160: S225–S237.
10. Froeschke G, Sommer S (2005) MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the southern Kalahari. *Mol Biol Evol* 22: 1254–1259. doi:10.1093/molbev/msi112.
11. 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. *Mol Ecol* 14: 85–91. doi:10.1111/j.1365-294X.2004.02402.x.

12. Oliver M, Piertney S (2010) Beyond splitting hares and rabbiting on about major histocompatibility complex complexity. *Mol Ecol* 19: 4099–4101. doi:10.1111/j.1365-294X.2010.04812.x.
13. Holmes K (1994) Human ecology and behavior and sexually transmitted bacterial infections. *P Natl Acad Sci Usa* 91: 2448–2455.
14. Padian N, Shiboski S, Glass S, Vittinghoff E (1997) Heterosexual transmission of human immunodeficiency virus (HIV) in Northern California: Results from a ten-year study. *Am J Epidemiol* 146: 350–357.
15. Doherty IA, Padian NS, Marlow C, Aral SO (2005) Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. *J Infect Dis* 191 Suppl 1: S42–S4. doi:10.1086/425277.
16. Potterat JJ, Zimmerman-Rogers H, Muth SQ, Rothenberg RB, Green DL, et al. (1999) Chlamydia transmission: Concurrency, reproduction number, and the epidemic trajectory. *Am J Epidemiol* 150: 1331–1339.
17. Morris M, Kretzschmar M (1997) Concurrent partnerships and the spread of HIV. *Aids* 11: 641–648.
18. Morris M, Kretzschmar M (1995) Concurrent partnerships and transmission dynamics in networks. *Social Networks* 17: 299–318.
19. Kraut-Becher JR, Aral SO (2003) Gap length: an important factor in sexually transmitted disease transmission. *Sex Transm Dis* 30: 221–225.
20. Eames KTD, Keeling MJ (2004) Monogamous networks and the spread of sexually transmitted diseases. *Mathematical Biosciences* 189: 115–130. doi:10.1016/j.mbs.2004.02.003.
21. Hurst GD, Sharpe RG, Broomfield AH, Walker LE, Majerus TM, et al. (1995) Sexually transmitted disease in a promiscuous insect, *Adalia bipunctata*. *Ecol Entomol* 20: 230–236.
22. Bester-Meredith JK, Marler C (2003) The association between male offspring aggression and paternal and maternal behavior of *Peromyscus* mice. *Ethology* 109: 797–808.
23. Bredy TW, Brown RE, Meaney MJ (2007) Effect of resource availability on biparental care, and offspring neural and behavioral development in the California mouse (*Peromyscus californicus*). *Eur J Neurosci* 25: 567–575. doi:10.1111/j.1460-9568.2006.05266.x.
24. Kalcounis-Rueppell MC, Metheny JD, Vonhof MJ (2006) Production of ultrasonic vocalizations by *Peromyscus* mice in the wild. *Front Zool* 3: 3. doi:10.1186/1742-9994-3-3.
25. Storz JF (1999) Genetic consequences of mammalian social structure. *J Mammal* 80: 553–569.
26. Ribble DO (1992) Lifetime reproductive success and its correlates in the monogamous rodent, *Peromyscus californicus*. *J Anim Ecol* 61: 457–468.
27. Turner LM, Hoekstra HE (2006) Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice (*Peromyscus*). *Mol Biol Evol* 23: 1656–1669. doi:10.1093/molbev/msl035.
28. Vandegrift KJ, Raffel TR, Hudson PJ (2008) Parasites prevent summer breeding in white-footed mice, *Peromyscus leucopus*. *Ecology* 89: 2251–2258.
29. Birdsall DA, Nash D (1973) Occurrence of successful multiple insemination of females in natural populations of deer mice (*Peromyscus maniculatus*). *Evolution* 27: 106–110.
30. Ribble DO, Millar J (1996) The mating system of northern populations of *Peromyscus maniculatus* as revealed by radiotelemetry and DNA fingerprinting. *Ecoscience* 3: 423–428.
31. Shurtliff QR, Pearse DE, Rogers DS (2005) Parentage analysis of the canyon mouse (*Peromyscus crinitus*): Evidence for multiple paternity. *J Mammal* 86: 531–540.
32. Foltz DW (1981) Genetic evidence for long-term monogamy in a small rodent, *Peromyscus polionotus*. *Am Nat* 117: 665–675.
33. Ribble DO (1991) The monogamous mating system of *Peromyscus californicus* as revealed by DNA fingerprinting. *Behav Ecol Sociobiol* 29: 161–166.
34. Gubernick DJ, Teferi T (2000) Adaptive significance of male parental care in a monogamous mammal. *P Roy Soc Lond B Bio* 267: 147–150.
35. MacManes MD (2011) Promiscuity in mice is associated with increased vaginal bacterial diversity. *Naturwissenschaften* 98: 951–960. http://dx.doi.org/10.1007/s00114-011-0848-2. PMID 21964973.
36. Sikes RS, Gannon WL, Animal Care and Use Committee of the American Society of Mammalogists (2011) Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *J Mammal* 92: 235–253. doi:10.1644/10-MAMM-F-355.1.
37. Miller S, Dykes D, Polesky H (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16: 1215–1215.
38. Chirhart SE, Honeycutt RL, Greenbaum IF (2000) Microsatellite markers for the deer mouse *Peromyscus maniculatus*. *Mol Ecol* 9: 1669–1671.
39. Mullen LM, Hirschmann RJ, Prince KL, Glenn TC, Dewey MJ, et al. (2006) Sixty polymorphic microsatellite markers for the oldfield mouse developed in *Peromyscus polionotus* and *Peromyscus maniculatus*. *Mol Ecol Notes* 6: 36–40. doi:10.1111/j.1471-8286.2005.01128.x.
40. Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8: 103–106. doi:10.1111/j.1471-8286.2007.01931.x.
41. Raymond M, Rousset F (1995) GENEPOP (version 1.2): population-genetics software for exact tests and ecumenicism. *J Hered* 86: 248–249.
42. Excoffier L, Lischer HEL (2010) Arlequin suite ver3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567. doi:10.1111/j.1755-0998.2010.02847.x.
43. Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Mol Ecol Resour* 8: 299–301. doi:10.1111/j.1471-8286.2007.01997.x.
44. Pfau RS, Van Den Bussche RA, McBee K, Lochmiller RL (1999) Allelic diversity at the MHC-DQA locus in cotton rats (*Sigmodon hispidus*) and a comparison of DQA sequences within the family Muridae (Mammalia : Rodentia). *Immunogenetics* 49: 886–893.
45. Drummond A, Ashton B, Buxton S, Cheung M, Cooper A, et al. (2010) Geneious. Available: http://www.geneious.com. Accessed: 15 February 2011.
46. Wilson DJ, McVean G (2006) Estimating diversifying selection and functional constraint in the presence of recombination. *Genetics* 172: 1411–1425. doi:10.1534/genetics.105.044917.
47. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351: 652–654. doi:10.1038/351652a0.
48. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452. doi:10.1093/bioinformatics/btp187.
49. Fu Y, Li W (1993) Statistical Tests of Neutrality of Mutations. *Genetics* 133: 693–709.
50. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
51. Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24: 1586–1591. doi:10.1093/molbev/msm088.
52. Kosakovsky Pond SL, Frost SD, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21: 676–679. doi:10.1093/bioinformatics/bti079.
53. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* 32: 1792–1797. doi:10.1093/nar/gkh340.
54. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574. doi:10.1093/bioinformatics/btg180.
55. Anisimova M, Nielsen R, Yang Z (2003) Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 164: 1229–1236.
56. Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22: 2472–2479. doi:10.1093/molbev/msi237.
57. Yang Z, Reis dos M (2011) Statistical Properties of the Branch-Site Test of Positive Selection. *Mol Biol Evol* 28: 1217–1228. doi:10.1093/molbev/msq303.
58. Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. *Theoret Appl Genetics* 38: 226–231. doi:10.1007/BF01245622.
59. Bondinas GP, Moustakas AK, Papadopoulos GK (2007) The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. *Immunogenetics* 59: 539–553. doi:10.1007/s00251-007-0224-8.
60. Oppelt C, Starkloff A, Rausch P, Holst von D, Rödel HG (2010) Major histocompatibility complex variation and age-specific endoparasite load in subadult European rabbits. *Mol Ecol* 19: 4155–4167. doi:10.1111/j.1365-294X.2010.04766.x.
61. Schad J, Dechmann DKN, Voigt CC, Sommer S (2011) MHC class II DRB diversity, selection pattern and population structure in a neotropical bat species, *Noctilio albiventris*. *Heredity* 107: 115–126. doi:10.1038/hdy.2010.173.
62. Fraser BA, Weadick CJ, Janowitz I, Rodd H, Hughes KA (2011) Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. *BMC Genomics* 12: 202. doi:10.1186/1471-2164-12-202.
63. Hansen MM, Skaala O, Jensen LF, Bekkevold D, Mensberg K-LD (2007) Gene flow, effective population size and selection at major histocompatibility complex genes: brown trout in the Hardanger Fjord, Norway. *Mol Ecol* 16: 1413–1425. doi:10.1111/j.1365-294X.2007.03255.x.
64. Hansen MM, Ruzzante DE, Nielsen EE, Bekkevold D, Mensberg K-LD (2002) Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. *Mol Ecol* 11: 2523–2535.
65. Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 10: 195–205. doi:10.1038/nrg2526.
66. Scheffler K, Martin DP, Seoighe C (2006) Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 22: 2493–2499. doi:10.1093/bioinformatics/btl427.
67. Schaschl H, Suchentrunk F, Hammer S, Goodman SJ (2005) Recombination and the origin of sequence diversity in the DRB MHC class II locus in chamois (*Rupicapra* spp.). *Immunogenetics* 57: 108–115. doi:10.1007/s00251-005-0784-4.
68. Schaschl H, Wandeler P, Suchentrunk F, Obexer-Ruff G, Goodman SJ (2006) Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. *Heredity* 97: 427–437. doi:10.1038/sj.hdy.6800892.
69. Lavner Y, Kotlar D (2005) Codon bias as a factor in regulating expression via translation rate in the human genome. *Gene* 345: 127–138. doi:10.1016/j.gene.2004.11.035.
70. Kvennefors ECE, Sampayo E, Ridgway T, Barnes AC, Hoegh-Guldberg O (2010) Bacterial communities of two ubiquitous great barrier reef corals reveals both site- and species-specificity of common bacterial associates. *PLoS ONE* 5: e10401. doi:10.1371/journal.pone.0010401.
71. Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39: 197–218. doi:10.1146/annurev.genet.39.073003.112420.

72. Otto SP (2000) Detecting the form of selection from DNA sequence data. *Trends Genet* 16: 526–529.
73. Broquet T, Jaquière J, Perrin N (2009) Opportunity for sexual selection and effective population size in the lek-breeding European treefrog (*Hyla arborea*). *Evolution* 63: 674–683. doi:10.1111/j.1558-5646.2008.00586.x.
74. Stiver JR, Apa AD, Remington TE, Gibson RM (2008) Polygyny and female breeding failure reduce effective population size in the lekking Gunnison sage-grouse. *Biol Conserv* 141: 472–481. doi:10.1016/j.biocon.2007.10.018.
75. Bellinger MR, Johnson JA, Toepfer J, Dunn P (2003) Loss of genetic variation in Greater Prairie Chickens following a population bottleneck in Wisconsin, USA. *Conserv Biol* 17: 717–724.
76. Johnson JA, Bellinger MR, Toepfer JE, Dunn P (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. *Mol Ecol* 13: 2617–2630. doi:10.1111/j.1365-294X.2004.02264.x.
77. Bouzat JL, Johnson K (2004) Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chickens. *Mol Ecol* 13: 499–505. doi:10.1046/j.1365-294X.2003.02068.x.
78. Alcaide M, Lemus JA, Blanco G, Tella JL, Serrano D, et al. (2010) MHC diversity and differential exposure to pathogens in kestrels (Aves: Falconidae). *Mol Ecol* 19: 691–705. doi:10.1111/j.1365-294X.2009.04507.x.
79. Wegner K, Reusch T, Kalbe M (2003) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J Evol Biol* 16: 224–232.
80. Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, et al. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Mol Ecol* 16: 1439–1451. doi:10.1111/j.1365-294X.2007.03281.x.
81. Landry C, Bernatchez L (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Mol Ecol* 10: 2525–2539.
82. Ribble DO (1992) Dispersal in a monogamous rodent, *Peromyscus californicus*. *Ecology* 73: 859–866.
83. Gubernick DJ, Nordby JC (1993) Mechanisms of sexual fidelity in the monogamous California mouse, *Peromyscus californicus*. *Behav Ecol Sociobiol* 32: 211–219.
84. Altizer S, Nunn CL, Thrall PH, Gittleman JL, Antonovics J, et al. (2003) Social organization and parasite risk in mammals: Integrating theory and empirical studies. *Annu Rev Ecol Evol S* 34: 517–547. doi:10.1146/annurev.ecolsys.34.030102.151725.
85. Kerness Whiteman N, Parker PG (2004) Effects of host sociality on ectoparasite population biology. *J Parasitol* 90: 939–947.
86. Wlasiuk G, Nachman MW (2010) Promiscuity and the rate of molecular evolution at primate immunity genes. *Evolution* 64: 2204–2220. doi:10.1111/j.1558-5646.2010.00989.x.
87. Kleiman DG (1977) Monogamy in mammals. *Q Rev Biol* 52: 39–69.
88. Calisi RM, Bentley GE (2009) Lab and field experiments: Are they the same animal? *Horm Behav* 56: 1–10. doi:10.1016/j.yhbeh.2009.02.010.