
Publishers page: http://dx.doi.org/10.1073/pnas.1708597115
<http://dx.doi.org/10.1073/pnas.1708597115>

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.
Immunogenetic novelty confers a selective advantage in host-pathogen coevolution

Karl Phillips1, Joanne Cable2, Ryan Mohammed3, Magdalena Herdegen-Radwan1, Jaroslaw Raubic1, Karolina Przesmycka1, Cock van Oosterhout1, Jacek Radwan1

1Adam Mickiewicz University in Poznań, 2Cardiff University, 3University of the West Indies, 4University of East Anglia

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The major histocompatibility complex (MHC) is crucial to the adaptive immune response of vertebrates and is among the most polymorphic gene families known. Its high diversity is usually attributed to selection imposed by fast-evolving pathogens. Pathogens are thought to evolve to escape recognition by common immune alleles, and, hence, novel MHC alleles, introduced through mutation, recombination or gene flow, are predicted to give hosts superior resistance. Although this theoretical prediction underpins host-pathogen Red Queen coevolution, it has not been demonstrated in the context of natural MHC diversity. Here, we experimentally tested whether novel MHC variants (both alleles and functional genotypes) that are in resonance of guppy Poecilia reticulata to a common ectoparasite (Gyrodactylus turnbulli). We used exposure-controlled infection trials with wild-sourced parasites, and Gyrodactylus-naive host fish that were F2 descendants of crossed wild populations. Hosts carrying MHC variants (alleles or supertypes) that were new to a given parasite population experienced a 35-37% reduction in infection intensity, but the number of MHC variants carried by an individual, analogous to heterozygosity in single-locus systems, was not a significant predictor. Our results provide direct evidence of novel MHC advantage, confirming a fundamental mechanism underpinning the exceptional polymorphism of this gene family, and highlighting the role of immunogenetic novelty in host-pathogen coevolution.

Introduction

Host-pathogen coevolution is thought to drive the maintenance of genetic variation in immune genes, with consequences for important evolutionary processes including the evolution of virulence, the maintenance of sex, and sexual selection (1-4). One of the most striking examples of genetic polymorphism thought to be maintained by such processes is the vertebrate major histocompatibility complex (MHC), where dozens to hundreds of alleles may segregate in natural populations (5-7). The high polymorphism of this important immune gene family, which codes for proteins that present pathogen-derived antigens to T-cell receptors, has been a subject of research for decades (8), and understanding the processes that maintain this diversity has implications for areas outside evolutionary biology, from human health (9) to conservation biology (10, 11).

Despite almost fifty years of investigation, the processes driving evolution at the MHC are not fully understood (12). At the molecular level, the exceptionally high ratio of non-synonymous to synonymous nucleotide substitutions in MHC genes suggests that selection is not only maintaining polymorphism (‘balancing selection’), but is also actively promoting new polymorphism [positive selection’ (13, 14)]. Several mutually non-exclusive mechanisms may contribute to these selective pressures: heterozygote advantage (recognizing a wider spectrum of antigens); frequency-dependent selection from fast-evolving pathogens that favors rare or novel MHC variants; and variable selection in space and time (12). Recent theoretical work has suggested that frequency-dependent selection resulting from Red Queen dynamics may be the more important process, with the advantage conferred by novel alleles being particularly important in generating the patterns of allelic diversity observed at the MHC (15, 16). Novel allele advantage is an old hypothesis in MHC research, dating to the earliest days of observing the MHC’s extreme polymorphism (17). The mechanistic potential for novel MHC variants to confer adaptive advantage against pathogens has been demonstrated experimentally only relatively recently, using congenic mice and artificially selected virus lineages (18, 19). However, the number of MHC alleles segregating in wild populations can be upwards of two orders of magnitudes higher than that of the mouse virus system. It may be more difficult for pathogens to adapt to specific local variants, and novel variants will be competing in a much larger pool of alleles with potentially a wide range of antigen-binding properties. Experimentally testing novel variant advantage in more natural, ecological contexts is much harder, as the potential selective pressures acting on the MHC are notoriously hard to disentangle (12).

Here, we used direct experimentation to investigate how novel MHC class II alleles (which recognize extracellular pathogens) in tropical freshwater guppies (Poecilia reticulata) affected the infection trajectory of their monogenean parasite Gyrodactylus turnbulli. These ectoparasites are widespread across guppy populations, and exert significant selective pressure (20, 21). Heavy infections can kill hosts (22, 23), and some MHC class II genotypes have been linked to gyrodactyliid infection in the wild (21; see also 24). This host-parasite system is highly tractable: host exposure is easy to control, and infections can

Significance

The major histocompatibility complex (MHC) is one of the most polymorphic gene families in the vertebrate genome, with natural selection actively promoting and maintaining polymorphism. The exact mechanism/mechanisms responsible for these characteristics remain unclear, but identifying them is fundamental to our understanding of host-pathogen dynamics. Using targeted crosses of the model Trinidadian guppy, a tractable parasite, and exposure-controlled infection trials, we show that novel MHC variants are associated with less severe infections. Uniquely, our experimental design separates novel variant advantage from other modes of selection and confounding variables, such as individual MHC variability and genomic background. We thus demonstrate a fundamental process driving evolution of the vertebrate immune system, which helps explain the unique features of MHC genes.
be monitored through time without killing host or parasite (22, 23). Hosts in our experiments were F2 descendants of crosses between guppy populations that shared no MHC alleles (see methods), and gyrodactylid worms for each replicate were wild-caught and came from one of the populations used to found the respective cross. In addition to MHC novelty defined by amino acid sequences, we also considered novelty based on MHC supertypes, where MHC alleles are grouped into clusters with similar physicochemical properties (25-27). We predicted that hosts carrying MHC variants that were novel with respect to parasite origin would perform better in controlled gyrodactylid infection trials than hosts carrying ‘local’ MHC variants.

Results

Amongst guppies that survived to the end of the experiment (n = 209), fish carrying only novel alleles or supertypes (designated as N/N genotypes and N/N supergenotypes, respectively; see ‘Materials and methods’) experienced G. turnbulli infections that were significantly less severe than fish carrying only ‘local’ alleles or supertypes (L/L). Infection severity was measured in ‘worm days’ (the area under a graph of number of worms against time), and analyses used AICC-based multi-model inference (see methods). The N/N genotypes (n = 44) and N/N supergenotypes (n = 14) respectively experienced 35% and 37% fewer ‘worm days’ compared to the L/L genotype and supergenotype fish (n = 59 and 88, respectively; P = 0.003 and 0.012; Tables S2.1a,b, S2.2a,b). L/N genotypes (n = 106; i.e. fish carrying both novel and local alleles) experienced infections of comparable intensity to L/L
genotypes \((P = 0.65; \text{Table S2.1b})\), whereas L/N supergenotypes \((n = 107)\) experienced intermediate infection intensities that were only marginally non-significant relative to L/L \((P = 0.055; \text{Table S2.2b})\). Direct comparison of the best allele-based and supertype-based models of worm days, using a constant set of covariates, indicated that allele-based groupings produced the better fit \((\Delta \text{AIC}_C = 4.22; \text{Tables S2.1a, S2.2a})\); however, fish carrying at least one novel supertype \((n = 121)\) experienced 27.1% fewer worm days than fish with no novel supertypes but at least one novel allele \((n = 29)\); top-ranked model: \(P = 0.02; \text{Tables S2.3a,b}; \text{Fig. S2.1}\). We did not detect a significant interaction between replicate population and MHC genotype/supergenotype class \((\Delta \text{AIC}_C = +7.37/+5.71; P \geq 0.13/0.24)\). Neither the number of alleles nor the number of supertypes carried by a host – measures used as analogues of heterozygosity – were significantly associated with the number of worm days experienced by hosts \((\Delta \text{AIC}_C = +1.82/+1.28, P = 0.52/0.32; \text{Tables S2.1a,b, S2.2a,b})\).

No genetic variables were significant predictors of host mortality (Tables S2.5a-c, S2.6a-c), despite worm load itself being a significant predictor of mortality from infection day 3 onwards (Table S2.4).

### Discussion

Our results support the novel MHC variant hypothesis: N/N hosts experienced parasite infections that were significantly less severe than those of L/L hosts. This was the case whether novelty was defined by amino acid sequences (alleles) or by physicochemical functional groups (supertypes). Differences in parasite burden between the genotype classes did not translate into a detectable effect on host survival. However, this may reflect the relatively benign and stable conditions of the experiment: in the wild, fish weakened by infection may be more susceptible to predation (28) and secondary infections (29), and to environmental stressors such as river spates (30) – even one additional worm can reduce a wild guppy’s survival probability (30). Furthermore, besides reducing survival, parasites may reduce host fitness by affecting reproductive potential, as has previously been demonstrated in guppies (31).

The novel variant advantage that we observed could, in theory, result in either balanced polymorphism or fixation of the novel variant – i.e. it is consistent with both balancing and positive selection. When a novel variant is introduced into a natural population (by point mutation, microrecombination or introgression), both processes are co-occurring and indistinguishable, likely resulting in an increase of the novel allele’s frequency. We explored this potential using computer simulations parameterized from our current data on the effects of novel alleles/supertypes on gyrodactylid load, and from the effect of gyrodactylid load on the survival of guppies in the field from a mark-recapture study (30).

These simulations show that upwards of 11% of novel variants should successfully establish in a population, and upwards of 54% if the variant is a novel supertype (compared to \(<0.1%\) in neutral simulations; see Appendix S10).

In the long term, the same process that leads to novel variant advantage – adaptation of parasites to local MHC genotypes – should diminish the advantage of the variant, leading to balanced polymorphism via negative frequency-dependent selection (12, 16, 17, 32). Such dynamics, whereby a novel (or rare) allele increases in frequency, loses its advantage, and decreases in frequency again have yet to be demonstrated. Alternatively, in the absence of balancing selection, novel variants could spread to fixation in a population, but such a scenario is inconsistent with the high MHC polymorphism observed in most study systems being coupled with strong signatures of positive selection.

Novel variant advantage may also explain the striking trans-species polymorphism observed at the MHC (33): if such polymorphisms are derived from hybridization as opposed to being ancestral (34), novel variant advantage may accelerate introgression and promote interspecies sharing of polymorphism. Furthermore, novel allele advantage might also affect the evolution of MHC-based mating preferences, an important factor in shaping MHC diversity (35-37). Our results suggest that preferences for partners with MHC alleles that are novel in a population, rather than those that just differ from self MHCs, should be strongly favored.

Negative frequency-dependent dynamics (favoring both novel and [sufficiently] rare alleles) and heterozygote advantage are two important types of balancing selection maintaining MHC diversity (12). Although frequency dependence and heterozygote advantage are not mutually exclusive, their relative influences are notoriously hard to test independently (12) and may be impossible to separate by observational studies alone (38). Other processes that can contribute to MHC diversity further complicate the separation (2, 12, 39). However, the crosses in our experiment produced genotypes/supertypologies that would not normally be found in natural populations at the time at which novelty enters/arthes. In particular, we generated hosts that were ‘homozygous’ with respect to MHC novelty (N/N) while controlling for genetic background. Hence, we were able to separate the effects of novelty from the effects of simply carrying more MHC variants, and this showed that the number of alleles or supertypes carried by an individual was not significantly associated with infection intensity (Tables S2.1a,b, S2.2a,b).

Furthermore, we did not detect an overdominance-type advantage for L/N fish (Tables S2.1b, S2.2b). The simplicity of the guppy-Gyrodactylus system may explain the lack of heterozygote advantage, which previous work has shown to be particularly important in multi-pathogen systems (40). Importantly, however, our results show that, against a natural host-pathogen genetic background, novel MHC variants can be selectively advantageous by virtue of properties arising from their novelty/extreme rarity (17, 32), rather than by simply being present in heterozygotes.

Direct comparison of the best allele-based and supertype-based models of infection intensity indicated that allele-based groupings produced the better fit (Tables S2.1a, S2.2a). On the other hand, fish with at least one novel supertype experienced infections that were significantly less severe than fish with no novel supertypes but at least one novel allele (i.e. a novel amino acid sequence variant within a shared supertype). This suggests that functional novelty may be more important than simple allelic novelty, whereby the unique binding properties of the novel supertypes are more likely to fill an immune response void (41, 42). The disparity between the statistical model and empirical observations on parasite loads of guppies with and without novel supertypes highlights that the relative fitness contributions of novel supertypes and novel alleles within supertypes remain to be determined. Despite this uncertainty, our experiment demonstrates a general advantage of novel MHC variants.

Whilst the breeding design of our study controls for population-level linkage between MHC class II genes and other genes that may affect immune responses (28 chromosome pairs (43), plus recombination when F1s reproduce), we cannot exclude possible effects from genes that may be in close physical linkage with the MHC without knockdown experiments or isogenic guppy lineages. Unlike for tetrapods, linkage with MHC class I can be ruled out for teleost fish (44) – a pertinent point because, although MHC class I usually targets intracellular pathogens, these genes have been co-opted into roles more typical of class II in some teleosts (45). Concerning the exact mechanism by which the MHC may influence responses to skin ectoparasites, one possibility is through antigen-presenting skin cells (dendritic cells) mediating production of pro-inflammatory cytokines. This has been demonstrated in vitro with zebrafish skin tissue (46), and implied by gene expression studies in salmonids infected with sea...
lice (47-49). That supertypes are associated with lower infection intensities also suggests a functional rather than linkage-based in-
fluence of MHC.

Our results suggest that novel variants rather than locally adapted variants are associated with lower levels of parasite infection, and the lack of a significant interaction with replicate pop-
ulation suggests this finding may be generalisable. This contrasts with several studies on stickleback (Gasterosteus aculeatus) MHC-
parasite interactions. When comparing or crossing lake and river
stickleback populations, these studies have variously shown local adaptation (24), immigrant advantage (50), and a mixture of signals (51). This variation likely reflects the highly divergent
docologies and parasite faunas of lake and river sticklebacks, which may exert a complex suite of selection pressures (24, 50-52).

Our study builds on the important insights from these studies by over-
coming some of their limitations, such as the use of only a single
population pair (24, 50); no experimental control of population-
level linkage ([50]; although this study’s use of statistical control means it is uniquely able to demonstrate effects of such linkage); and the stock caveats of snapshot observational studies ([51]; e.g., uncertainty regarding where an allele may be in a frequency-
dependent dynamic, and the difficulty of separating effects of rare
alleles from heterozygosity (42)). Also, these previous studies did not test for effects of supertypes. The most pertinent stickleback
experiment to our result presents the simplest finding: MHC vari-
ants that confer resistance, irrespective of being rare or common, tend to increase in frequency (55). Many traits and circumstances
may make a variant resistant; our study shows that novelty is likely to be one of them.

Based on our results, we predict that novel variants, including those coming from immigrants, have a reasonable probability of establishing and spreading. However, inferring the consequences of novel variant advantage from MHC-based population genetic
structure is more challenging, especially in populations connected
by gene flow (42). Furthermore, genuine novelty of a variant can only be ascertained by exhaustively sampling a population’s MHC diversity over an extended period of time. Introgression of novel alleles may be easier to observe if it occurs among well-diversified populations coming into secondary contact, including between species (e.g. 56) or between allopatric populations (such as those we investigated). We suspect that the small number of alleles shared between Trinidad and Tobago (Appendix S5) might come from fish introduced to Tobago by humans, as most of these alleles were found in a population close to human settlements and communication hubs. If so, our simulations suggest that these alleles are likely to spread across Tobago in the nearby future, subject to migration rate to other Tobaganian populations.

Our data constitute empirical demonstration of a long-
positoned, fundamental model of the evolution of MHC variability: novel immune variants confer a selective advantage, consistent with Red Queen scenarios in which parasites adapt to local host immune genotypes (8, 15, 16). We show this effect using wild-sourced host and pathogen genetic variation, indicating that demonstrations of novel MHC advantage in congeneric laboratory systems (18, 19) may be applicable in the context of natural MHC diversity. Furthermore, because we show this with replicable parasite populations across controlling for population genetic
background, our finding should be pertinent regardless of the source of novelty (e.g. mutation, recombination, introgression) or the rarity with which novelty enters a system. Looking beyond the MHC, although Red Queen dynamics have been shown in several non-vertebrate host-pathogen systems (57-59), examples of experimentally tested molecular mechanisms are rare, even for simple systems such as bacteria-phaeage interactions (60). In con-
trast, we explicitly link phenotype (infection intensity) to genotype
and linkage (in novel MHC) with an a priori hypothesis, using a gene family with a well-characterized immunological function. Overall,
population, and allocated F2s to three genotype groups based on these designations. After rejection of two local haplotype motifs and the LN/NN (mixed/heterozygous). Although a nominally viable allele could be present in the local population at a frequency too low for our sample to detect, population genetic analyses on a larger dataset suggested that this is likely to be extremely rare, if not completely absent in our data. The three replicate populations are prezent on islands (Appendix 55). Because of the breeding design, all three genotype groups should have the same average genetic background with respect to population genetic variation and heterozygosity (23 chromosome pairs (43), plus recombination when F1s reproduce).

Amino acid (AA) substitutions vary in their functional consequences for an MHC molecule’s antigen-binding profile, such that alleles with different AA sequences may be functionally similar. These MHC ‘superotypes’ are predicted to bind to similar antigenic ‘supermotifs’, and may better characterize the breadth of host defense than alleles (25-27). Using 15 guppy MHC codons previously identified as being under positive selection (69), five physicochemical descriptors of each AA (70), and discriminant analysis of principal components (71,72), we reduced the list of allele sequences to 14 supertype clusters (full description in Appendix S1.3). Supertype designations were used to assign fish to LN, NN, and supertypes analogous to allele-based groupings, except that supertypes shared between pairs of crossed populations (all pairs shared at least one supertype; Appendix 55) were treated as ‘local’. ‘Supergenotypes’ thus do not describe the haplotype makeup of individual fish, as they may include alleles shared across supertypes. Only nine individuals, over three replicate populations, had a number of supertypes lower than their number of alleles (i.e. they carried 2+ alleles of the same supertype). Of these, only one changed novel/local categorization between allelic and supertype-based analyses (from LN to L/L). Novel and local alleles did not differ systematically in the supertype categorization between allelic and supertype-based analyses (from L/N to N/N). We included interactions of the highest ranked model to include MHC group and the experimental population. We also included interactions elsewhere in the analysis.

We used mitochondrial barcoding to identify 2-4 gyno-replicates to species level (full details in Appendix S1.3). All sequences showed their strongest matches (98-100% identity) against published G. turnbulli (GT) sequences, except for those representing 35 fish from the AVPPS population, all infected on the same day, which matched G. baltaruss (Gb; 98-100%). These fish also showed markedly different pathology, consistent with Gb (Appendix S1.2). We also found no patho logical differences between Gb and Gt and the absence of indicators of Gb elsewhere in the experiment, we excluded these fish from our analyses.

The analyses described in the next two paragraphs were performed separately for alleles and supertypes. We first tested if MHC group (UL, LN, NN) predicted whether or not an individual fish survived the experiment. We used corrected Akaike information criterion (AICc) ranking (73) of logistic regressions to explore all combinations of the following main effects: MHC group; number of MHC variants (alleles+supertypes, continuous, 1-4); standard length (continuous, z-transformed); age (one variable, ‘male’ if in possession of a fully-shaped gonopodium, i.e. ‘hook’ and ‘hood’ visible; ‘female’ if length > 13.0 mm and no gonopodium evident; ‘juvenile’ for all others); temperature (mean daily maximum over monitored period, z-transformed); and experimental population. We also included interactions between experimental population and both MHC group and number of MHC variants (alleles+supertypes, continuous, 1-4); and interactions of the highest ranked model to include MHC group and the experimental population. For clarity of presentation in the main text, we only report the ΔAICc of the highest ranked model to include MHC group and the P-values (two-tailed) for certain contrasts, but in Appendix S2 we give a fully nuanced account of the analysis.

We used a similar process to test whether MHC group predicted infection burden among fish that survived the experiment. We used ‘worm days’ as the response variable, calculated as the total area under a fish’s 17- day infection trajectory line. Worm days are tractable to analyze (no zero-inflation or random effects) and provide an ecologically relevant summary measure of infection severity. We compared infected versus uninfected worms. We then examined models comprising the top two units of ranked models. We found that the highest ranked model to include MHC group and the experimental population. We also included interactions elsewhere in the analysis.

We thank staff of the Environmental Research Institute Charlotteville (ERIC), Tobago, for support during the field station, A Pilastro for training in guppy artificial insemination, members of the Cardiff University parasite research group who helped with parasite work, J Lighten for sharing an as-then unpublished manuscript, and M Migalska for Fig. 1. We thank M Konczal, M Migalska, DS Richardson, J Kaufmann, LG Spurgin, W Babik, CA Morrison, J Lighten, W Smallbone, AR Ellison, M Hammers, MG Gage, and three anonymous reviewers for comments on earlier manuscript versions. This project was funded by a Polish National Science Centre Harmony grant to JRDwn, CG and JC (grant number UMO-2013/10/M/NZ8/00253). JRDwn, CG and JC conceived the study with input from JRDwn and RSM on experimental design. RSM and KPP collected and reared the fish; MHR and JRDwn performed artificial inseminations; KPP performed infection trials, with assistance from JC, KJP and JRDwn; and JK performed the statistical analyses. MHR and KJP; KPP analyzed the data and performed computer simulations; JRDwn and KPP wrote the manuscript, with input from all co-authors. The authors declare no competing interests. Data accessibility All data and scripts will be made available in accordance with the publisher’s requirements.


