HAL
open science

Avian blood parasite richness decreases with major histocompatibility complex class I loci number

Orsolya Vincze, Claire Loiseau, Mathieu Giraudeau

To cite this version:

Orsolya Vincze, Claire Loiseau, Mathieu Giraudeau. Avian blood parasite richness decreases with major histocompatibility complex class I loci number. Biology Letters, 2021, 17 (8), 10.1098/rsbl.2021.0253 . hal-03418441

HAL Id: hal-03418441

https://hal.science/hal-03418441v1

Submitted on 7 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BIOLOGY LETTERS

Avian blood parasite richness decreases with MHC class I loci number

Journal:	*Biology Letters*
Manuscript ID	RSBL-2021-0253.R2
Article Type:	Research
Date Submitted by the Author:	12-Jul-2021
Complete List of Authors:	Orsolya, Vincze; University Montpellier 1, ; Babeş-Bolyai University, Evolutionary Ecology Group, Hungarian Department of Biology and Ecology Loiseau, Claire; CIBIO-InBIO - Research Center in Biodiversity and Genetic Resources, InBIO Associate Laboratory; CEFE Giraudeau, Mathieu; University of Montpellier
Subject:	Evolution < BIOLOGY, Molecular Biology < BIOLOGY
Categories:	Evolutionary Biology
Keywords:	Haemosporidians, copy number variation, MHC diversity, parasite richness, pathogen-mediated selection, comparative analyses

SCHOLARONE™
Manuscripts

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
All data used in the analysis is available in Table S1 and S2.

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
O.V. & C.L. wrote the manuscript. O.V. performed the comparative analysis. C.L. and M.G. designed the analysis. All authors critically revised the manuscript, gave final approval for its publication and agree to be held accountable for the work performed therein.

For Review Only

Avian blood parasite richness decreases with MHC class I loci number

Orsolya Vincze[1,2,3,4*], Claire Loiseau[5,6*], Mathieu Giraudeau[1,2]

[1] CREEC, MIVEGEC, UMR IRD 224-CNRS 5290-Université de Montpellier, Montpellier, France and CREES Centre for Research on the Ecology and Evolution of Disease, Montpellier, France

[2] Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS-La Rochelle Université, 2 Rue Olympe de Gouges, FR-17000 La Rochelle, France

[3] Institute of Aquatic Ecology, Centre for Ecological Research, 4026 Debrecen, Hungary

[4] Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeş-Bolyai University, RO–400006 Cluj-Napoca, Clinicilor Street 5–7, Romania

[5]CIBIO-InBIO - Research Center in Biodiversity and Genetic Resources, InBIO Associate Laboratory, Campus de Vairão, 7 Rua Padre Armando Quintas, 4485-661 Vairão, Portugal.

[6] CEFE, Université de Montpellier, CNRS, Montpellier, France

* First Co-Authors

Correspondence: vincze.orsolya@ecolres.hu [OV]

Abstract

Major Histocompatibility Complex (MHC) genes are among the most polymorphic in the vertebrate genome. The high allele diversity is believed to be maintained primarily by sexual and pathogen-mediated balancing selection. The number of MHC loci also varies greatly across vertebrates, most notably across birds. MHC proteins play key roles in presenting antigens on the cell surface for recognition by T cells, with class I proteins

specifically targeting intracellular pathogens. Here we explore the hypothesis that MHC class I diversity (measured as loci number) co-evolves with haemosporidian parasite burden of the host. Using data on 54 bird species we demonstrate that high MHC class I diversity is associated with significantly lower richness of *Plasmodium*, *Haemoproteus* as well as overall haemosporidian parasite lineages, the former thus indicating a more efficient protection against intracellular pathogens. Nonetheless, the latter associations were only detected when MHC diversity was assessed using cloning and not 454 pyrosequencing-based studies, nor across all genotyping methods combined. Our results indicate that high MHC class I diversity might play a key role in providing qualitative resistance against diverse haemosporidian parasites in birds, but further clarification is needed for the origin of contrasting results when using different genotyping methods for MHC loci quantification.

Keywords

Haemosporidians, copy number variation, parasite richness, pathogen-mediated selection, comparative analyses, MHC diversity

1. Introduction

The Major Histocompatibility Complex (MHC) is known to be the most variable gene group in vertebrates, both in terms of allelic diversity and gene number [1]. MHC genes code for cell surface glycoproteins that bind to antigenic peptides and present them to cells of the immune system, initiating adaptive immune responses. Proteins encoded by the MHC genes ensure the immune presentation of a broad spectrum of peptides, both derived from the intracellular (class I proteins) or extracellular (class II proteins) space [2]. High MHC diversity, reflected by high allele diversity or loci number, is often related to the

ability to present a wider repertoire of antigens to the immune system, thus representing a genetic marker of disease resistance [3] but see [4]. Several, mutually non-exclusive hypotheses have been put forward to explain the maintenance of high MHC diversity, such as the heterozygote advantage (i.e., MHC diversity conferring disease resistance), negative frequency-dependent selection (benefit of rare MHC alleles in pathogen detection) or fluctuating selection (spatial or temporal change in the selective advantage of MHC alleles) [5,6].

The number of MHC loci varies greatly across vertebrates, being the result of a complex evolutionary history, involving gene duplications and translocations [7,8].For example, species from the Galliformes order possess only two MHC class I genes [9,10], which is considerably lower than in other bird lineages [11,12], such as passerines. Several studies reported that a low number of MHC loci could represent the ancestral state in birds, and that the highest diversification in copy numbers occurred relatively late in the avian radiation, reaching its highest diversity in passerines [12]. While, parasite-mediated selection has been considered as the most prominent component driving MHC diversification [13], it is now widely accepted that in addition to the pathogen burden, sexual selection and risk of developing autoimmune diseases or immunopathologies may together shape the evolution of an optimal rather than maximal MHC diversity [3,4,13–15].

Malaria and closely related haemosporidian parasites are intensively investigated vector-borne parasites in birds, partly because of the high diversity of hosts they are able to infect, and their nearly worldwide distribution [16]. A series of experimental studies [17–21], have demonstrated the substantial costs associated with haemosporidian infection, in terms of

reproductive success, health condition and survival. Hence, immunological adaptations targeting and eliminating such infections can provide considerable fitness benefits. Key components of these defenses are the MHC proteins. Nonetheless, knowledge on the co-evolution of MHC diversity and avian haemosporidian parasite burden across species remains limited [22–26]. However, other host-parasite systems found strong evidences that specific MHC alleles confer resistance to pathogens [27,28].

With the increasing use of molecular tools, a large number of bird species have been recently screened for haemosporidian parasites. The MalAvi database represents a unified database compiling these data, covering genetic and geographical data of avian blood parasites (*Plasmodium, Haemoproteus, Leucocytozoon* genera). It currently encompasses more than 14,400 records with 4,520 lineages detected in more than 2,100 host species [29]. Taking advantage of this unique database, here we explore whether MHC diversity co-evolves with haemosporidian lineages richness. Given that haemosporidian parasites are obligate intracellular parasites, here we focus on the MHC class I genes, exploring relationships between gene copy numbers (hereafter MHC-I diversity) and blood parasite lineages richness. Based on earlier species-specific studies on the link between MHC-I allelic diversity and blood parasites [23,26] and according to the heterozygote advantage hypothesis, we predict a negative association between MHC-I diversity and the diversity of intracellular parasites across host species. Nonetheless, an opposite association can also be predicted since studies focusing on extracellular pathogens, and thus MHC class II genes, detected positive associations between parasite burden and MHC diversity [10,30–32].

2. Material and Methods

(a) Database

We collected estimates of MHC-I diversity (i.e. loci number) from published databases [3,12,33]. In total, we compiled 81 different estimates of MHC-I diversity with corresponding blood parasite information (see final database in Table S1). For each bird species with MHC-I diversity, we collected data on the number of detected blood parasite lineages (hereafter lineages richness) of *Plasmodium*, *Haemoproteus* and *Leucocytozoon* parasites using the MalAvi database [29], (accessed on 11th June 2021. Moreover, given that the number of parasite lineages detected in a host is strongly dependent on sampling effort, we also extracted the number of publications (i.e. total number of studies exploring the given parasite genus in each particular host, see Table S2) and the total number of individuals tested for the given haemosporidian genus in each bird species. When data about the number of individuals tested was missing in the MalAvi database, this information was collected directly from the referenced publications. Given that body size is correlated with both parasite diversity and MHC diversity [12], we also collected information on species-specific body masses from literature [12].

(b) Statistical analyses

In order to explore the effect of sampling effort, we tested the association between parasite lineages richness and indicators of sampling effort (i.e. number of individuals or number of publications) for each parasite genus separately and overall (i.e. combining all three genera of blood parasites). Linear regressions indicated that the number of publications was a considerably and consistently better predictor of parasite lineages richness than the number of individuals inspected in all three inspected parasite genera, as well as for overall haemosporidian lineages richness (appendix SI: figure S1). Therefore, we used the number of publications as covariate in all subsequent analysis to control for sampling effort.

125 To test for associations between MHC-I diversity and blood parasite diversity, we
126 constructed phylogenetic generalized linear mixed model, using a Bayesian approximation
127 in the R package '*MCMCglmm*' [34]. We used parasite lineages richness as dependent
128 variables (separately for each genus, and overall haemosporidian lineages richness) in
129 models with poisson error distribution [34]. Models included MHC-I diversity, species
130 body mass and the number of publication of the respective group of blood parasites as
131 covariates. MHC-I diversity data was methodologically heterogeneous, due to a large
132 variety of genotyping methods used to quantify it (e.g., cloning, 454 pyrosequencing, long-
133 reads, see table S1). Given that these methods frequently provide divergent results [33,35],
134 we included genotyping method as a random factor in the models. Moreover, MHC-I
135 diversity was repeatedly assessed in some species using the same genotyping method,
136 providing slightly conflicting estimates. Given that the accuracy of these estimates is
137 difficult to determine, we included all estimates in the models, but using species as a
138 random factor to control for pseudoreplication. Phylogeny was included in all models to
139 control for the shared evolutionary history of the host species. We used weakly informative
140 priors in all models [34]. Model convergence and absence of auto-correlation was assessed
141 by visual examination of the posterior parameter distributions. Since MHC diversity might
142 not be linearly related to parasite burden, we also tested quadratic effects in these models,
143 by incorporating MHC class I diversity as second-degree orthogonal polynomials. In
144 addition, because it is often ill-advised to combine data from various genotyping methods
145 [11,12,36], we also performed models separately for the two most frequently used
146 genotyping methods (cloning: n = 22 records, 21 species; 454 pyrosequencing: n = 35
147 records, 22 species). The limited number of species assessed by the other genotyping
148 methods did not allow separate analysis (n species ≤ 9 for all other methods). For models,
149 only using cloning or pyrosequencing-based studies, we used the same MCMCglmm model

150 structure as described above but without the genotyping method random factor. Genotyping
151 effort can influence estimates of loci diversity, therefore to confirm the lack of such bias in
152 our results, we performed sensitivity analysis using the number of genotyped individuals as
153 model weights. All results were highly consistent irrespective of control for genotyping
154 effort (Table S6).
155
156 To account for phylogenetic non-independence we used trees from http://birdtree.org
157 [37]. We downloaded 1000 random trees using the Hackett backbone tree [38], and
158 calculated a rooted, ultrametric consensus tree using the SumTrees software [39]. All
159 statistical analyses were conducted using R version 3.6.3 [40].
160
161 **3. Results**
162 We compiled complete data on parasite lineages richness and MHC-I diversity for 54 bird
163 species, from 29 families and 7 orders, with *Passeriformes* being the most represented (*n* =
164 39 species; appendix SI: table S1). MHC class I loci numbers in our database varied widely,
165 from one to 33 loci. MHC-I diversity was highly divergent in species repeatedly genotyped
166 using different methods (e.g. *Acrocephalus arundinaceus*, cloning: 4 loci, 454
167 pyrosequencing: 9-10 loci). Cloning generally detected less MHC class I loci than 454
168 pyrosequencing (1-6, and 4-23 respectively). Observed parasite lineages richness also
169 varied greatly across host species (*Plasmodium*: 0-38; *Haemoproteus*: 0-42; *Leucocytozoon*
170 0-45; All blood parasites genera: 1-78).
171
172 The overall analysis, with genotyping method included as random factor (appendix SI: table
173 S3a) revealed no association between MHC-I diversity and any of the three haemosporidian
174 genera (*Plasmodium*: p = 0.578, *Haemoproteus*: p = 0.467, *Leucocytozoon*: p = 0.356), or in

175 overall blood parasite lineages richness (p=0.533). None of the quadratic effects of MHC
176 class I diversity were significant (Table S3b).
177
178 Models based on cloning assessment of MHC-I diversity (appendix SI: table S4) indicated
179 that a higher MHC-I diversity was associated with a significantly lower lineages richness of
180 *Plasmodium* (posterior mean = - 0.47, p = 0.0161; figure 1a), *Haemoproteus* (posterior
181 mean = -0.51, p = 0.0148; figure 1b), and overall haemosporidian lineages richness
182 (posterior mean = -0.39, p = 0.0120, figure 1d). No association was found between MHC-I
183 diversity and *Leucocytozoon* lineages richness (posterior mean = -0.05, p = 0.9047; figure
184 1c). In contrast, MHC diversity quantified using 454 pyrosequencing, revealed no
185 correlation between MHC-I diversity and *Plasmodium* (posterior mean = -0.01, p =
186 0.5843), *Haemoproteus* (posterior mean = -0.01, p = 0.5897), *Leucocytozoon* (posterior
187 mean = -0.05, p = 0.1446), or overall haemosporidian lineages richness (posterior mean = -
188 0.01, p = 0.4264, appendix SI: figure S2a-d, table S5).
189
190 **4. Discussion**
191 Here we performed a cross-species comparative analysis on MHC-I loci diversity in birds
192 and explored its association with the diversity of intracellular blood parasites infecting
193 these hosts. We did not detect an overall association between MHC-I diversity and blood
194 parasite diversity across species, but methodological heterogeneity in MHC genotyping
195 limits the reliability of these combined analyses [41]. On the contrary, analysis restricted to
196 cloning assessment of MHC-I diversity (but not 454 pyrosequencing) indicated that high
197 MHC-I diversity is associated with lower blood parasite diversity, regarding overall
198 haemosporidian, *Plasmodium* and *Haemoproteus*, but not *Leucocytozoon* lineages richness.
199

200 Comparative studies exploring the link between MHC class II and extracellular parasites
201 (helminths or ectoparasites) indicated that high parasite diversity was positively associated
202 with MHC nucleotide [30], or allelic diversity [31]. Bolnik et al. [42] however found a
203 negative correlation between MHC class II diversity and overall microbial diversity. One
204 recent study focusing for the first time on MHC loci number found a positive association
205 between helminth diversity and MHC class II [32]. Positive associations are commonly
206 explained by the benefit of maintaining high MHC diversity under strong parasite pressure.
207 Contrary to these results, our study points out a different dynamic linking intracellular
208 blood parasite burden and MHC-I diversity. We show that MHC-I diversity and blood
209 parasite richness is not positively, but rather negatively correlated, indicating that high
210 MHC-I diversity may provide qualitative resistance against diverse avian haemosporidian
211 parasites. Experimental studies support the possibility of full resistance against strains of
212 blood parasites [17]. In fact, after inoculation with a generalist *Plasmodium* lineage (SGS1),
213 bird species showed strikingly different parasite burdens, varying from full resistance or to
214 high susceptibility [17]. The role of MHC protein contributing to this resistance is
215 supported by the fact that individual flycatchers with high number of functional MHC
216 alleles were shown to have lower probability of malarial infection [43]. Other studies
217 demonstrate that the presence of specific MHC alleles could significantly decrease
218 probabilities of being infected by certain blood parasite strains [23,24]. Contrasting results
219 between parasite burden and MHC class I or class II diversity might be explained by
220 various non-exclusive factors: (i) the distinct evolutionary history of these genes and their
221 capacity to affect resistance and/or tolerance [12,41,44], (ii) the highly divergent species
222 composition of the mentioned studies (e.g. patterns detected across broad taxonomic scales
223 might not be present with narrow taxonomic coverage and vice versa), and (iii) the different

224 association of MHC diversity to different parasite groups (e.g. MHC class II diversity
225 relates negatively to bacterial but positively to helminth diversity). Additional studies are
226 needed to elucidate the mechanisms behind these divergent associations between MHC and
227 pathogen diversity.
228
229 Our results indicating fewer intracellular pathogens in host species with more MHC-I loci
230 support the heterozygote advantage hypothesis, which relies on the advantage provided by
231 heterozygosity of the MHC alleles, through the recognition of a wider range of antigens
232 compared to homozygotes. While it was suggested that pathogen richness alone seems
233 insufficient to explain the maintenance of extremely high MHC diversity (especially in
234 some passerines) [25], our results (restricted to cloning assessment of MHC) show a strong
235 correlation between MHC and parasite diversity. Importantly however, our results indicate
236 that the association between MHC-I diversity and parasite lineages richness is not
237 consistent when different genotyping techniques are used. The reason for this discrepancy
238 might be manifold. First, cloning quantification of MHC-I diversity in our database covered
239 a relatively wide taxonomic range (*Galliformes* to *Passeriformes*), while 454
240 pyrosequencing was only applied in passerines. It is thus possible that pathogen-mediated
241 evolution contributed to major taxonomic differences in MHC loci numbers, but different
242 selection processes are at play at the narrower taxonomic scales of e.g. passerines. The
243 latter is also supported by the greatly accelerated MHC duplication rate in songbirds, which
244 is unlikely to be the result of parasite-mediated selection alone [12]. Second, it is now
245 known that significantly fewer MHC class I alleles are expressed than present in the
246 genome [3], which is particularly concerning in species with highly duplicated MHC genes
247 such as many songbirds [3,11,12]. Non-expressed MHC alleles (pseudogenes) are
248 particularly problematic while using genotyping methods relying on genomic DNA

assembled from short segments [3], such as pyrosequencing, which potentially inflates estimated of copy numbers. Consistently, pyrosequencing detected higher MHC loci than cloning in all species genotyped by both techniques. Third, the association between MHC-I diversity and parasite burden is not always linear [26], sometimes indicating a negative association with moderate increase in MHC diversity, but not across high MHC diversity ranges. Consequently, an initial decline in parasite burden with moderate increase in MHC diversity (e.g. cloning analyses: loci number varies from 1 to 6) does not exclude the possibility of a different association at high loci numbers (e.g., analyses restricted to pyrosequencing: loci number varies from 4 to 23). Indeed, high MHC diversity can lead to a significant reduction of the T cell receptor repertoire, having detrimental effects on the efficiency of the immune response [45].

Overall, these results indicate a potentially important role of intracellular parasites in shaping cross-species variance in MHC-I diversity. Nonetheless, evidence is mixed when using different genotyping techniques, highlighting the need for better quality data on MHC diversity to better understand immune gene evolution and selection shaping this. We urge future research to quantify diversity in expressed MHC loci and to benefit from advanced genotyping technologies, such as long-read sequencing.

Data accessibility.

All data are available in Table S1. Data and all code used to generate the results will be made available in the Dryad Digital Repository upon acceptance.

Authors' contributions.

O.V. & C.L. wrote the manuscript. O.V. performed the comparative analysis. C.L. and M.G. designed the analysis. All authors contributed to revisions and approved the final version.

Funding.

This work is funded by National Funds through FCT - Foundation for Science and Technology under the IF/00744/2014/CP1256/CT0001 Exploratory Research Project (C.L) and the PTDC/BIA-EVL/29390/2017 DEEP Research Project (C.L.). OV was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the New National Excellence Programme of the Hungarian Ministry of Innovation and Technology.

Acknowledgments

We are grateful to Eric Pante for consultations on genotyping techniques and to the two anonymous reviewers for their helpful comments..

References

1. Edwards S V., Hedrick PW. 1998 Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends Ecol. Evol.* **13**, 305–311. (doi:10.1016/S0169-5347(98)01416-5)
2. Klein J. 1986 *Natural history of the major histocompatibility complex*. New York, NY: Wiley.
3. O'Connor EA, Westerdahl H. 2021 Trade-offs in expressed major histocompatibility complex diversity seen on a macroevolutionary scale among songbirds. *Evolution* **75**, 1061-1069. (doi:10.1111/evo.14207)
4. Wegner KM. 2003 Parasite Selection for Immunogenetic Optimality. *Science (80-.).* **301**, 1343–1343. (doi:10.1126/science.1088293)

5. Bernatchez L, Landry C. 2003 MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* **16**, 363–377. (doi:10.1046/j.1420-9101.2003.00531.x)

6. Spurgin LG, Richardson DS. 2010 How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. B Biol. Sci.* **277**, 979–988. (doi:10.1098/rspb.2009.2084)

7. Kulski JK, Shiina T, Anzai T, Kohara S, Inoko H. 2002 Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* **190**, 95–122. (doi:10.1034/j.1600-065X.2002.19008.x)

8. Balakrishnan CN *et al.* 2010 Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biol.* **8**, 29. (doi:10.1186/1741-7007-8-29)

9. Kaufman J, Milne S, Göbel TWF, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S. 1999 The chicken B locus is a minimal essential major histocompatibility complex. *Nature* **401**, 923–925. (doi:10.1038/44856)

10. Westerdahl H. 2007 Passerine MHC: genetic variation and disease resistance in the wild. *J. Ornithol.* **148**, 469–477. (doi:10.1007/s10336-007-0230-5)

11. O'Connor EA, Strandh M, Hasselquist D, Nilsson J-Å, Westerdahl H. 2016 The evolution of highly variable immunity genes across a passerine bird radiation. *Mol. Ecol.* **25**, 977–989. (doi:10.1111/mec.13530)

12. Minias P, Pikus E, Whittingham LA, Dunn PO. 2019 Evolution of Copy Number at the MHC Varies across the Avian Tree of Life. *Genome Biol. Evol.* **11**, 17–28. (doi:10.1093/gbe/evy253)

13. Bentkowski P, Radwan J. 2019 Evolution of major histocompatibility complex gene copy number. *PLOS Comput. Biol.* **15**, e1007015. (doi:10.1371/journal.pcbi.1007015)

14. Nowak MA, Tarczy-Hornoch K, Austyn JM. 1992 The optimal number of major histocompatibility complex molecules in an individual. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10896–10899. (doi:10.1073/pnas.89.22.10896)

15. Minias P, Whittingham LA, Dunn PO. 2017 Coloniality and migration are related to selection on MHC genes in birds. *Evolution (N. Y).* **71**, 432–441. (doi:10.1111/evo.13142)

16. Valkiunas G. 2004 *Avian Malaria Parasites and other Haemosporidia*. CRC Press.

17. Palinauskas V, Valkiūnas G, Bolshakov C V., Bensch S. 2008 Plasmodium relictum (lineage P-SGS1): Effects on experimentally infected passerine birds. *Exp. Parasitol.* **120**, 372–380. (doi:10.1016/j.exppara.2008.09.001)

336 18. Palinauskas V, Valkiūnas G, Bolshakov C V., Bensch S. 2011 Plasmodium relictum
337 (lineage SGS1) and Plasmodium ashfordi (lineage GRW2): The effects of the co-
338 infection on experimentally infected passerine birds. *Exp. Parasitol.* **127**, 527–533.
339 (doi:10.1016/j.exppara.2010.10.007)

340 19. Marzal A, Lope F de, Navarro C, Möller AP. 2005 Malarial parasites decrease
341 reproductive success: an experimental study in a passerine bird. *Oecologia* **142**, 541–
342 545. (doi:10.1007/s00442-004-1757-2)

343 20. KNOWLES SCL, PALINAUSKAS V, SHELDON BC. 2010 Chronic malaria
344 infections increase family inequalities and reduce parental fitness: experimental
345 evidence from a wild bird population. *J. Evol. Biol.* **23**, 557–569.
346 (doi:10.1111/j.1420-9101.2009.01920.x)

347 21. Puente JM la, Merino S, Tomás G, Moreno J, Morales J, Lobato E, García-Fraile S,
348 Belda EJ. 2010 The blood parasite Haemoproteus reduces survival in a wild bird: a
349 medication experiment. *Biol. Lett.* **6**, 663–665. (doi:10.1098/rsbl.2010.0046)

350 22. Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G. 2006 MAJOR
351 HISTOCOMPATIBILITY ALLELES ASSOCIATED WITH LOCAL
352 RESISTANCE TO MALARIA IN A PASSERINE. *Evolution (N. Y).* **60**, 383–389.
353 (doi:10.1111/j.0014-3820.2006.tb01114.x)

354 23. Loiseau C, Zoorob R, Garnier S, Birard J, Federici P, Julliard R, Sorci G. 2008
355 Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecol.*
356 *Lett.* **11**, 258–265. (doi:10.1111/j.1461-0248.2007.01141.x)

357 24. Westerdahl H, Asghar M, Hasselquist D, Bensch S. 2012 Quantitative disease
358 resistance: to better understand parasite-mediated selection on major
359 histocompatibility complex. *Proc. R. Soc. B Biol. Sci.* **279**, 577–584.
360 (doi:10.1098/rspb.2011.0917)

361 25. Biedrzycka A, Bielański W, Ćmiel A, Solarz W, Zając T, Migalska M, Sebastian A,
362 Westerdahl H, Radwan J. 2018 Blood parasites shape extreme major
363 histocompatibility complex diversity in a migratory passerine. *Mol. Ecol.* **27**, 2594–
364 2603. (doi:10.1111/mec.14592)

365 26. MADSEN T, UJVARI B. 2006 MHC class I variation associates with parasite
366 resistance and longevity in tropical pythons. *J. Evol. Biol.* **19**, 1973–1978.
367 (doi:10.1111/j.1420-9101.2006.01158.x)

368 27. Eizaguirre C, Lenz TL, Kalbe M, Milinski M. 2012 Rapid and adaptive evolution of
369 MHC genes under parasite selection in experimental vertebrate populations. *Nat.*
370 *Commun.* **3**, 1–6. (doi:10.1038/ncomms1632)

371 28. Hill AVS *et al.* 1991 Common West African HLA antigens are associated with
372 protection from severe malaria. *Nature* **352**, 595–600. (doi:10.1038/352595a0)

29. BENSCH S, HELLGREN O, PÉREZ-TRIS J. 2009 MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome b lineages. *Mol. Ecol. Resour.* **9**, 1353–1358. (doi:10.1111/j.1755-0998.2009.02692.x)

30. Šimková A, Ottová E, Morand S. 2006 MHC variability, life-traits and parasite diversity of European cyprinid fish. *Evol. Ecol.* **20**, 465–477. (doi:10.1007/s10682-006-0014-z)

31. GOÜY DE BELLOCQ J, CHARBONNEL N, MORAND S. 2008 Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *J. Evol. Biol.* **21**, 1144–1150. (doi:10.1111/j.1420-9101.2008.01538.x)

32. Minias P, Gutiérrez JS, Dunn PO. 2020 Avian major histocompatibility complex copy number variation is associated with helminth richness. *Biol. Lett.* **16**, 20200194. (doi:10.1098/rsbl.2020.0194)

33. He K, Minias P, Dunn PO. 2021 Long-Read Genome Assemblies Reveal Extraordinary Variation in the Number and Structure of MHC Loci in Birds. *Genome Biol. Evol.* **13**. (doi:10.1093/gbe/evaa270)

34. Hadfield J. 2010 MCMC methods for Multi–response Generalised Linear Mixed Models: The MCMCglmm R Package. *J. Stat. Softw.* **33**, 1–22. (doi:10.1002/ana.22635)

35. BABIK W. 2010 Methods for MHC genotyping in non-model vertebrates. *Mol. Ecol. Resour.* **10**, 237–251. (doi:10.1111/J.1755-0998.2009.02788.X)

36. Rekdal SL, Anmarkrud JA, Johnsen A, Lifjeld JT. 2018 Genotyping strategy matters when analyzing hypervariable major histocompatibility complex-Experience from a passerine bird. *Ecol. Evol.* **8**, 1680–1692. (doi:10.1002/ece3.3757)

37. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012 The global diversity of birds in space and time. *Nature* **491**, 444–448. (doi:10.1038/nature11631)

38. Hackett SJ *et al.* 2008 A Phylogenomic Study of Birds Reveals Their Evolutionary History. *Science (80-.).* **320**, 1763–8. (doi:10.1126/science.1157704)

39. Sukumaran J, Holder MT. 2010 DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**, 1569–1571. (doi:10.1093/bioinformatics/btq228)

40. R Core Team. 2020 R: a language and environment for statistical computing.

41. Minias P, Włodarczyk R, Remisiewicz M, Cobzaru I, Janiszewski T. 2021 Distinct evolutionary trajectories of MHC class I and class II genes in Old World finches and buntings. *Heredity* **126**, 974–990. (doi:10.1038/s41437-021-00427-8)

42. Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014 Major Histocompatibility Complex class IIb polymorphism influences gut

409 microbiota composition and diversity. *Mol. Ecol.* **23**, 4831–4845.
410 (doi:10.1111/MEC.12846)

411 43. RADWAN J, ZAGALSKA-NEUBAUER M, CICHOŃ M, SENDECKA J, KULMA
412 K, GUSTAFSSON L, BABIK W. 2012 MHC diversity, malaria and lifetime
413 reproductive success in collared flycatchers. *Mol. Ecol.* **21**, 2469–2479.
414 (doi:10.1111/j.1365-294X.2012.05547.x)

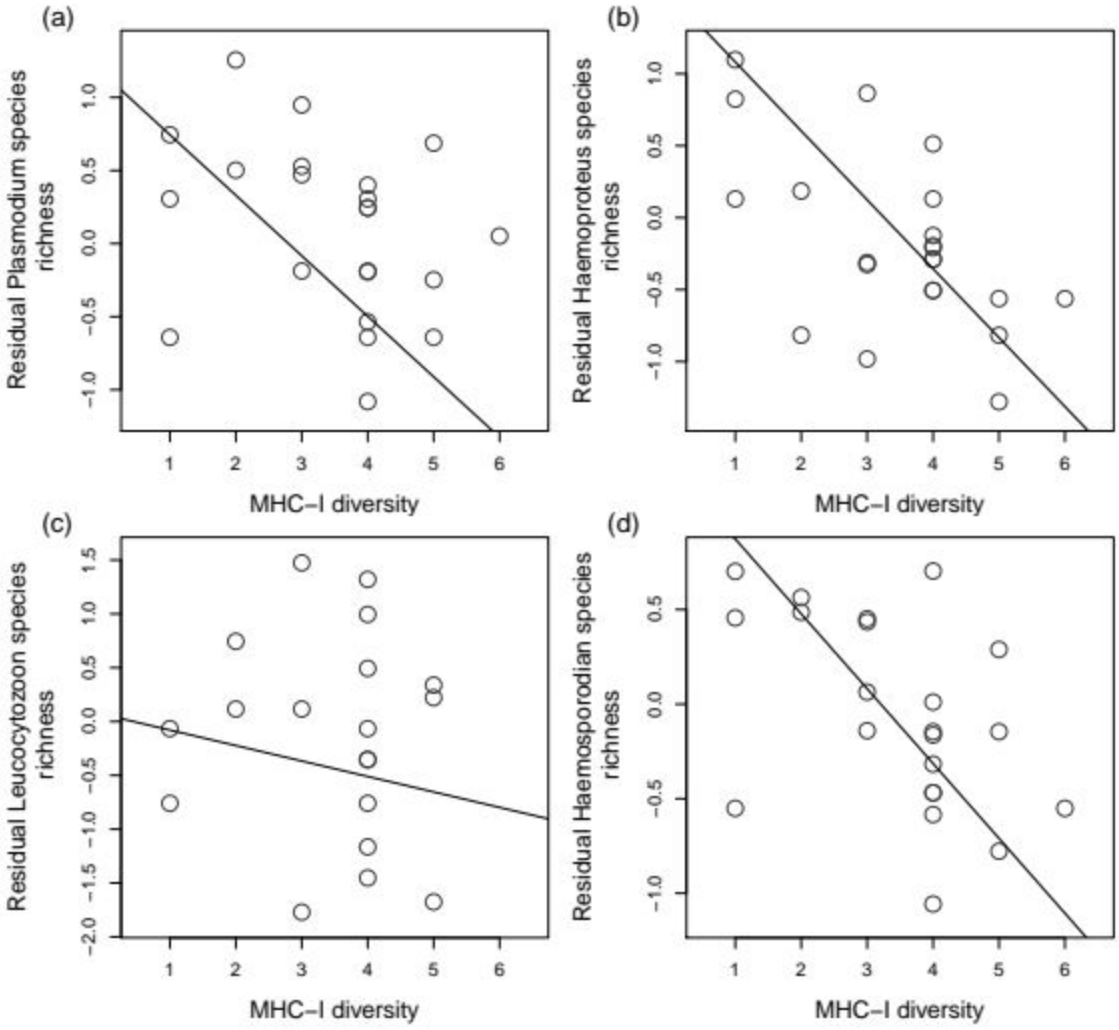
415 44. Råberg L. 2014 How to Live with the Enemy: Understanding Tolerance to Parasites.
416 *PLoS Biol.* **12**, e1001989. (doi:10.1371/journal.pbio.1001989)

417 45. Migalska M, Sebastian A, Radwan J. 2019 Major histocompatibility complex class I
418 diversity limits the repertoire of T cell receptors. *Proc. Natl. Acad. Sci.* **116**, 5021–
419 5026. (doi:10.1073/pnas.1807864116)

420

421

Figure 1. Association between residual blood parasite species richness and MHC-I diversity for (a) *Plasmodium*, (b) *Haemoproteus*, (c) *Leucocytozoon* and (d) overall haemosporidian lineages richness. Parasite diversity residuals were obtained from a log-log linear regressions between parasite lineages richness (dependent) and research effort (explanatory). Slopes were obtained from phylogenetic MCMCglmms with poisson error distribution between parasite lineages richness and research effort as well as MHC-I diversity.



Avian blood parasite richness decreases with MHC class I loci number

Orsolya Vincze[1,2,3,4*], Claire Loiseau[5,6*], Mathieu Giraudeau[1,2]

[1] CREEC, MIVEGEC, UMR IRD 224-CNRS 5290-Université de Montpellier, Montpellier, France and CREES Centre for Research on the Ecology and Evolution of Disease, Montpellier, France

[2] Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS-La Rochelle Université, 2 Rue Olympe de Gouges, FR-17000 La Rochelle, France

[3] Institute of Aquatic Ecology, Centre for Ecological Research, 4026 Debrecen, Hungary

[4] Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeş-Bolyai University, RO–400006 Cluj-Napoca, Clinicilor Street 5–7, Romania

[5]CIBIO-InBIO - Research Center in Biodiversity and Genetic Resources, InBIO Associate Laboratory, Campus de Vairão, 7 Rua Padre Armando Quintas, 4485-661 Vairão, Portugal.

[6] CEFE, Université de Montpellier, CNRS, Montpellier, France

* First Co-Authors

Correspondence: vincze.orsolya@ecolres.hu [OV]

Abstract

Major Histocompatibility Complex (MHC) genes are among the most polymorphic in the vertebrate genome. The high allele diversity is believed to be maintained primarily by sexual and pathogen-mediated balancing selection. The number of MHC loci also varies greatly across vertebrates, most notably across birds. MHC proteins play key roles in presenting antigens on the cell surface for recognition by T cells, with class I proteins

specifically targeting intracellular pathogens. Here we explore the hypothesis that MHC class I diversity (measured as loci number) co-evolves with haemosporidian parasite burden of the host. Using data on 54 bird species we demonstrate that high MHC class I diversity is associated with significantly lower richness of *Plasmodium*, *Haemoproteus* as well as overall haemosporidian parasite lineages, the former thus indicating a more efficient protection against intracellular pathogens. Nonetheless, the latter associations were only detected when MHC diversity was assessed using cloning and not 454 pyrosequencing-based studies, nor across all genotyping methods combined. Our results indicate that high MHC class I diversity might play a key role in providing qualitative resistance against diverse haemosporidian parasites in birds, but further clarification is needed for the origin of contrasting results when using different genotyping methods for MHC loci quantification.

Keywords

Haemosporidians, copy number variation, parasite richness, pathogen-mediated selection, comparative analyses, MHC diversity

1. Introduction

The Major Histocompatibility Complex (MHC) is known to be the most variable gene group in vertebrates, both in terms of allelic diversity and gene number [1]. MHC genes code for cell surface glycoproteins that bind to antigenic peptides and present them to cells of the immune system, initiating adaptive immune responses. Proteins encoded by the MHC genes ensure the immune presentation of a broad spectrum of peptides, both derived from the intracellular (class I proteins) or extracellular (class II proteins) space [2]. High MHC diversity, reflected by high allele diversity or loci number, is often related to the

ability to present a wider repertoire of antigens to the immune system, thus representing a genetic marker of disease resistance [3] but see [4]. Several, mutually non-exclusive hypotheses have been put forward to explain the maintenance of high MHC diversity, such as the heterozygote advantage (i.e., MHC diversity conferring disease resistance), negative frequency-dependent selection (benefit of rare MHC alleles in pathogen detection) or fluctuating selection (spatial or temporal change in the selective advantage of MHC alleles) [5,6].

The number of MHC loci varies greatly across vertebrates, being the result of a complex evolutionary history, involving gene duplications and translocations [7,8].For example, species from the Galliformes order possess only two MHC class I genes [9,10], which is considerably lower than in other bird lineages [11,12], such as passerines. Several studies reported that a low number of MHC loci could represent the ancestral state in birds, and that the highest diversification in copy numbers occurred relatively late in the avian radiation, reaching its highest diversity in passerines [12]. While, parasite-mediated selection has been considered as the most prominent component driving MHC diversification [13], it is now widely accepted that in addition to the pathogen burden, sexual selection and risk of developing autoimmune diseases or immunopathologies may together shape the evolution of an optimal rather than maximal MHC diversity [3,4,13–15].

Malaria and closely related haemosporidian parasites are intensively investigated vector-borne parasites in birds, partly because of the high diversity of hosts they are able to infect, and their nearly worldwide distribution [16]. A series of experimental studies [17–21], have demonstrated the substantial costs associated with haemosporidian infection, in terms of

reproductive success, health condition and survival. Hence, immunological adaptations targeting and eliminating such infections can provide considerable fitness benefits. Key components of these defenses are the MHC proteins. Nonetheless, knowledge on the co-evolution of MHC diversity and avian haemosporidian parasite burden across species remains limited [22–26]. However, other host-parasite systems found strong evidences that specific MHC alleles confer resistance to pathogens [27,28].

With the increasing use of molecular tools, a large number of bird species have been recently screened for haemosporidian parasites. The MalAvi database represents a unified database compiling these data, covering genetic and geographical data of avian blood parasites (*Plasmodium, Haemoproteus, Leucocytozoon* genera). It currently encompasses more than 14,400 records with 4,520 lineages detected in more than 2,100 host species [29]. Taking advantage of this unique database, here we explore whether MHC diversity co-evolves with haemosporidian lineages richness. Given that haemosporidian parasites are obligate intracellular parasites, here we focus on the MHC class I genes, exploring relationships between gene copy numbers (hereafter MHC-I diversity) and blood parasite lineages richness. Based on earlier species-specific studies on the link between MHC-I allelic diversity and blood parasites [23,26] and according to the heterozygote advantage hypothesis, we predict a negative association between MHC-I diversity and the diversity of intracellular parasites across host species. Nonetheless, an opposite association can also be predicted since studies focusing on extracellular pathogens, and thus MHC class II genes, detected positive associations between parasite burden and MHC diversity [10,30–32].

2. Material and Methods

(a) Database

We collected estimates of MHC-I diversity (i.e. loci number) from published databases [3,12,33]. In total, we compiled 81 different estimates of MHC-I diversity with corresponding blood parasite information (see final database in Table S1). For each bird species with MHC-I diversity, we collected data on the number of detected blood parasite lineages (hereafter lineages richness) of *Plasmodium*, *Haemoproteus* and *Leucocytozoon* parasites using the MalAvi database [29], (accessed on 11th June 2021. Moreover, given that the number of parasite lineages detected in a host is strongly dependent on sampling effort, we also extracted the number of publications (i.e. total number of studies exploring the given parasite genus in each particular host, see Table S2) and the total number of individuals tested for the given haemosporidian genus in each bird species. When data about the number of individuals tested was missing in the MalAvi database, this information was collected directly from the referenced publications. Given that body size is correlated with both parasite diversity and MHC diversity [12], we also collected information on species-specific body masses from literature [12].

(b) Statistical analyses

In order to explore the effect of sampling effort, we tested the association between parasite lineages richness and indicators of sampling effort (i.e. number of individuals or number of publications) for each parasite genus separately and overall (i.e. combining all three genera of blood parasites). Linear regressions indicated that the number of publications was a considerably and consistently better predictor of parasite lineages richness than the number of individuals inspected in all three inspected parasite genera, as well as for overall haemosporidian lineages richness (appendix SI: figure S1). Therefore, we used the number of publications as covariate in all subsequent analysis to control for sampling effort.

To test for associations between MHC-I diversity and blood parasite diversity, we constructed phylogenetic generalized linear mixed model, using a Bayesian approximation in the R package '*MCMCglmm*' [34]. We used parasite lineages richness as dependent variables (separately for each genus, and overall haemosporidian lineages richness) in models with poisson error distribution [34]. Models included MHC-I diversity, species body mass and the number of publication of the respective group of blood parasites as covariates. MHC-I diversity data was methodologically heterogeneous, due to a large variety of genotyping methods used to quantify it (e.g., cloning, 454 pyrosequencing, long-reads, see table S1). Given that these methods frequently provide divergent results [33,35], we included genotyping method as a random factor in the models. Moreover, MHC-I diversity was repeatedly assessed in some species using the same genotyping method, providing slightly conflicting estimates. Given that the accuracy of these estimates is difficult to determine, we included all estimates in the models, but using species as a random factor to control for pseudoreplication. Phylogeny was included in all models to control for the shared evolutionary history of the host species. We used weakly informative priors in all models [34]. Model convergence and absence of auto-correlation was assessed by visual examination of the posterior parameter distributions. Since MHC diversity might not be linearly related to parasite burden, we also tested quadratic effects in these models, by incorporating MHC class I diversity as second-degree orthogonal polynomials. In addition, because it is often ill-advised to combine data from various genotyping methods [11,12,36], we also performed models separately for the two most frequently used genotyping methods (cloning: n = 22 records, 21 species; 454 pyrosequencing: n = 35 records, 22 species). The limited number of species assessed by the other genotyping methods did not allow separate analysis (n species ≤ 9 for all other methods). For models, only using cloning or pyrosequencing-based studies, we used the same MCMCglmm model

150 structure as described above but without the genotyping method random factor. Genotyping
151 effort can influence estimates of loci diversity, therefore to confirm the lack of such bias in
152 our results, we performed sensitivity analysis using the number of genotyped individuals as
153 model weights. All results were highly consistent irrespective of control for genotyping
154 effort (Table S6).
155
156 To account for phylogenetic non-independence we used trees from http://birdtree.org
157 [37]. We downloaded 1000 random trees using the Hackett backbone tree [38], and
158 calculated a rooted, ultrametric consensus tree using the SumTrees software [39]. All
159 statistical analyses were conducted using R version 3.6.3 [40].
160
161 **3. Results**
162 We compiled complete data on parasite lineages richness and MHC-I diversity for 54 bird
163 species, from 29 families and 7 orders, with *Passeriformes* being the most represented (*n* =
164 39 species; appendix SI: table S1). MHC class I loci numbers in our database varied widely,
165 from one to 33 loci. MHC-I diversity was highly divergent in species repeatedly genotyped
166 using different methods (e.g. *Acrocephalus arundinaceus*, cloning: 4 loci, 454
167 pyrosequencing: 9-10 loci). Cloning generally detected less MHC class I loci than 454
168 pyrosequencing (1-6, and 4-23 respectively). Observed parasite lineages richness also
169 varied greatly across host species (*Plasmodium*: 0-38; *Haemoproteus*: 0-42; *Leucocytozoon*
170 0-45; All blood parasites genera: 1-78).
171
172 The overall analysis, with genotyping method included as random factor (appendix SI: table
173 S3a) revealed no association between MHC-I diversity and any of the three haemosporidian
174 genera (*Plasmodium*: p = 0.578, *Haemoproteus*: p = 0.467, *Leucocytozoon*: p = 0.356), or in

overall blood parasite lineages richness (p=0.533). None of the quadratic effects of MHC class I diversity were significant (Table S3b).

Models based on cloning assessment of MHC-I diversity (appendix SI: table S4) indicated that a higher MHC-I diversity was associated with a significantly lower lineages richness of *Plasmodium* (posterior mean = - 0.47, p = 0.0161; figure 1a), *Haemoproteus* (posterior mean = -0.51, p = 0.0148; figure 1b), and overall haemosporidian lineages richness (posterior mean = -0.39, p = 0.0120, figure 1d). No association was found between MHC-I diversity and *Leucocytozoon* lineages richness (posterior mean = -0.05, p = 0.9047; figure 1c). In contrast, MHC diversity quantified using 454 pyrosequencing, revealed no correlation between MHC-I diversity and *Plasmodium* (posterior mean = -0.01, p = 0.5843), *Haemoproteus* (posterior mean = -0.01, p = 0.5897), *Leucocytozoon* (posterior mean = -0.05, p = 0.1446), or overall haemosporidian lineages richness (posterior mean = -0.01, p = 0.4264, appendix SI: figure S2a-d, table S5).

4. Discussion

Here we performed a cross-species comparative analysis on MHC-I loci diversity in birds and explored its association with the diversity of intracellular blood parasites infecting these hosts. We did not detect an overall association between MHC-I diversity and blood parasite diversity across species, but methodological heterogeneity in MHC genotyping limits the reliability of these combined analyses [41]. On the contrary, analysis restricted to cloning assessment of MHC-I diversity (but not 454 pyrosequencing) indicated that high MHC-I diversity is associated with lower blood parasite diversity, regarding overall haemosporidian, *Plasmodium* and *Haemoproteus*, but not *Leucocytozoon* lineages richness.

Comparative studies exploring the link between MHC class II and extracellular parasites (helminths or ectoparasites) indicated that high parasite diversity was positively associated with MHC nucleotide [30], or allelic diversity [31]. Bolnik et al. [42] however found a negative correlation between MHC class II diversity and overall microbial diversity. One recent study ~~also~~focusing for the first time on MHC loci number found a positive association between helminth diversity ~~and MHC class II loci number~~ and MHC class II [32]. ~~The~~Positive~~se~~ associations are commonly explained by the benefit of maintaining high MHC diversity under strong parasite pressure. Contrary to these results, our study points out a different dynamic linking intracellular blood parasite burden and MHC-I diversity. We show that MHC-I diversity and blood parasite richness is not positively, but rather negatively correlated, indicating that high MHC-I diversity may provide qualitative resistance against diverse avian haemosporidian parasites. Experimental studies support the possibility of full resistance against strains of blood parasites [17]. In fact, after inoculation with a generalist *Plasmodium* lineage (SGS1), bird species showed strikingly different parasite burdens, varying from full resistance or to high susceptibility [17]. The role of MHC protein contributing to this resistance is supported by the fact that individual flycatchers with high number of functional MHC alleles were shown to have lower probability of malarial infection [43~~2~~]. Other studies demonstrate that the presence of specific MHC alleles could significantly decrease probabilities of being infected by certain blood parasite strains [23,24]. C~~Finally, c~~ontrasting results between parasite burden and MHC class I or class II diversity might be explained by various non-exclusive factors: (i) the distinct evolutionary history of these genes and their capacity to affect resistance and/or tolerance [12,41,44], (ii) the highly divergent species composition of the mentioned studies (e.g. patterns detected across broad taxonomic scales might not be present with narrow

taxonomic coverage and vice versa), and (iii) the different association of MHC diversity to different parasite groups (e.g. MHC class II diversity relates negatively to bacterial but positively to helminth diversity). Additional studies are needed to elucidate the mechanisms behind these divergent associations between MHC and pathogen diversity. ~~It appears that MHC class I copy numbers were primarily shaped by accelerated evolution ('burst') and stabilizing selection, while the evolution of MHC class IIB loci were likely governed by fluctuating selection and drift [12].~~

Our results indicating fewer intracellular pathogens in host species with more MHC-I loci support the heterozygote advantage hypothesis, which relies on the advantage provided by heterozygosity of the MHC alleles, through the recognition of a wider range of antigens compared to homozygotes. While it was suggested that pathogen richness alone seems insufficient to explain the maintenance of extremely high MHC diversity (especially in some passerines) [25], our results (restricted to cloning assessment of MHC) show a strong correlation between MHC and parasite diversity. Importantly however, our results indicate that the association between MHC-I diversity and parasite lineages richness is not consistent when different genotyping techniques are used. The reason for this discrepancy might be manifold. First, cloning quantification of MHC-I diversity in our database covered a relatively wide taxonomic range (*Galliformes* to *Passeriformes*), while 454 pyrosequencing was only applied in passerines. It is thus possible that pathogen-mediated evolution contributed to major taxonomic differences in MHC loci numbers, but different selection processes are at play at the narrower taxonomic scales of e.g. passerines. The latter is also supported by the greatly accelerated MHC duplication rate in songbirds, which is unlikely to be the result of parasite-mediated selection alone [12]. Second, it is now known that significantly fewer MHC class I alleles are expressed than present in the

genome [3], which is particularly concerning in species with highly duplicated MHC genes such as many songbirds [3,11,12]. Non-expressed MHC alleles (pseudogenes) are particularly problematic while using genotyping methods relying on genomic DNA assembled from short segments [3], such as pyrosequencing, which potentially inflates estimated of copy numbers. Consistently, pyrosequencing detected higher MHC loci than cloning in all species genotyped by both techniques. Third, the association between MHC-I diversity and parasite burden is not always linear [26], sometimes indicating a negative association with moderate increase in MHC diversity, but not across high MHC diversity ranges. Consequently, an initial decline in parasite burden with moderate increase in MHC diversity (e.g. cloning analyses: loci number varies from 1 to 6) does not exclude the possibility of a different association at high loci numbers (e.g., analyses restricted to pyrosequencing: loci number varies from 4 to 23). Indeed, high MHC diversity can lead to a significant reduction of the T cell receptor repertoire, having detrimental effects on the efficiency of the immune response [45~~3~~].

Overall, these results indicate a potentially important role of intracellular parasites in shaping cross-species variance in MHC-I diversity. Nonetheless, evidence is mixed when using different genotyping techniques, highlighting the need for better quality data on MHC diversity to better understand immune gene evolution and selection shaping this. We urge future research to quantify diversity in expressed MHC loci and to benefit from advanced genotyping technologies, such as long-read sequencing.

Data accessibility.

All data are available in Table S1. Data and all code used to generate the results will be made available in the Dryad Digital Repository upon acceptance.

Authors' contributions.

O.V. & C.L. wrote the manuscript. O.V. performed the comparative analysis. C.L. and M.G. designed the analysis. All authors contributed to revisions and approved the final version.

Funding.

This work is funded by National Funds through FCT - Foundation for Science and Technology under the IF/00744/2014/CP1256/CT0001 Exploratory Research Project (C.L) and the PTDC/BIA-EVL/29390/2017 DEEP Research Project (C.L.). OV was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the New National Excellence Programme of the Hungarian Ministry of Innovation and Technology.

Acknowledgments

We are grateful to Eric Pante for consultations on genotyping techniques and to the two anonymous reviewers for their helpful comments..

References

1. Edwards S V., Hedrick PW. 1998 Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends Ecol. Evol.* **13**, 305–311. (doi:10.1016/S0169-5347(98)01416-5)

2. Klein J. 1986 *Natural history of the major histocompatibility complex*. New York, NY: Wiley.-

3. O'Connor EA, Westerdahl H. 2021 Trade-offs in expressed major histocompatibility complex diversity seen on a macroevolutionary scale among songbirds. *Evolution* **75**, 1061-1069 ~~*(N. Y).* , evo.14207~~. (doi:10.1111/evo.14207)

4. Wegner KM. 2003 Parasite Selection for Immunogenetic Optimality. *Science (80-.).* **301**, 1343–1343. (doi:10.1126/science.1088293)

5. Bernatchez L, Landry C. 2003 MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* **16**, 363–377. (doi:10.1046/j.1420-9101.2003.00531.x)

6. Spurgin LG, Richardson DS. 2010 How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. B Biol. Sci.* **277**, 979–988. (doi:10.1098/rspb.2009.2084)

7. Kulski JK, Shiina T, Anzai T, Kohara S, Inoko H. 2002 Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* **190**, 95–122. (doi:10.1034/j.1600-065X.2002.19008.x)

8. Balakrishnan CN *et al.* 2010 Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biol.* **8**, 29. (doi:10.1186/1741-7007-8-29)

9. Kaufman J, Milne S, Göbel TWF, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S. 1999 The chicken B locus is a minimal essential major histocompatibility complex. *Nature* **401**, 923–925. (doi:10.1038/44856)

10. Westerdahl H. 2007 Passerine MHC: genetic variation and disease resistance in the wild. *J. Ornithol.* **148**, 469–477. (doi:10.1007/s10336-007-0230-5)

11. O'Connor EA, Strandh M, Hasselquist D, Nilsson J-Å, Westerdahl H. 2016 The evolution of highly variable immunity genes across a passerine bird radiation. *Mol. Ecol.* **25**, 977–989. (doi:10.1111/mec.13530)

12. Minias P, Pikus E, Whittingham LA, Dunn PO. 2019 Evolution of Copy Number at the MHC Varies across the Avian Tree of Life. *Genome Biol. Evol.* **11**, 17–28. (doi:10.1093/gbe/evy253)

13. Bentkowski P, Radwan J. 2019 Evolution of major histocompatibility complex gene copy number. *PLOS Comput. Biol.* **15**, e1007015. (doi:10.1371/journal.pcbi.1007015)

14. Nowak MA, Tarczy-Hornoch K, Austyn JM. 1992 The optimal number of major histocompatibility complex molecules in an individual. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10896–10899. (doi:10.1073/pnas.89.22.10896)

15. Minias P, Whittingham LA, Dunn PO. 2017 Coloniality and migration are related to selection on MHC genes in birds. *Evolution (N. Y).* **71**, 432–441. (doi:10.1111/evo.13142)

16. Valkiunas G. 2004 *Avian Malaria Parasites and other Haemosporidia*. CRC Press.

17. Palinauskas V, Valkiūnas G, Bolshakov C V., Bensch S. 2008 Plasmodium relictum (lineage P-SGS1): Effects on experimentally infected passerine birds. *Exp. Parasitol.* **120**, 372–380. (doi:10.1016/j.exppara.2008.09.001)

18. Palinauskas V, Valkiūnas G, Bolshakov C V., Bensch S. 2011 Plasmodium relictum (lineage SGS1) and Plasmodium ashfordi (lineage GRW2): The effects of the co-infection on experimentally infected passerine birds. *Exp. Parasitol.* **127**, 527–533. (doi:10.1016/j.exppara.2010.10.007)

19. Marzal A, Lope F de, Navarro C, Mö-ller AP. 2005 Malarial parasites decrease reproductive success: an experimental study in a passerine bird. *Oecologia* **142**, 541–545. (doi:10.1007/s00442-004-1757-2)

20. KNOWLES SCL, PALINAUSKAS V, SHELDON BC. 2010 Chronic malaria infections increase family inequalities and reduce parental fitness: experimental evidence from a wild bird population. *J. Evol. Biol.* **23**, 557–569. (doi:10.1111/j.1420-9101.2009.01920.x)

21. Puente JM la, Merino S, Tomás G, Moreno J, Morales J, Lobato E, García-Fraile S, Belda EJ. 2010 The blood parasite Haemoproteus reduces survival in a wild bird: a medication experiment. *Biol. Lett.* **6**, 663–665. (doi:10.1098/rsbl.2010.0046)

22. Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G. 2006 MAJOR HISTOCOMPATIBILITY ALLELES ASSOCIATED WITH LOCAL RESISTANCE TO MALARIA IN A PASSERINE. *Evolution (N. Y).* **60**, 383–389. (doi:10.1111/j.0014-3820.2006.tb01114.x)

23. Loiseau C, Zoorob R, Garnier S, Birard J, Federici P, Julliard R, Sorci G. 2008 Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecol. Lett.* **11**, 258–265. (doi:10.1111/j.1461-0248.2007.01141.x)

24. Westerdahl H, Asghar M, Hasselquist D, Bensch S. 2012 Quantitative disease resistance: to better understand parasite-mediated selection on major histocompatibility complex. *Proc. R. Soc. B Biol. Sci.* **279**, 577–584. (doi:10.1098/rspb.2011.0917)

25. Biedrzycka A, Bielański W, Ćmiel A, Solarz W, Zając T, Migalska M, Sebastian A, Westerdahl H, Radwan J. 2018 Blood parasites shape extreme major histocompatibility complex diversity in a migratory passerine. *Mol. Ecol.* **27**, 2594–2603. (doi:10.1111/mec.14592)

26. MADSEN T, UJVARI B. 2006 MHC class I variation associates with parasite resistance and longevity in tropical pythons. *J. Evol. Biol.* **19**, 1973–1978. (doi:10.1111/j.1420-9101.2006.01158.x)

27. Eizaguirre C, Lenz TL, Kalbe M, Milinski M. 2012 Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat. Commun.* **3**, 1–6. (doi:10.1038/ncomms1632)

28. Hill AVS *et al.* 1991 Common West African HLA antigens are associated with protection from severe malaria. *Nature* **352**, 595–600. (doi:10.1038/352595a0)

29. BENSCH S, HELLGREN O, PÉREZ-TRIS J. 2009 MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome b lineages. *Mol. Ecol. Resour.* **9**, 1353–1358. (doi:10.1111/j.1755-0998.2009.02692.x)

30. Šimková A, Ottová E, Morand S. 2006 MHC variability, life-traits and parasite diversity of European cyprinid fish. *Evol. Ecol.* **20**, 465–477. (doi:10.1007/s10682-006-0014-z)

31. GOÜY DE BELLOCQ J, CHARBONNEL N, MORAND S. 2008 Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *J. Evol. Biol.* **21**, 1144–1150. (doi:10.1111/j.1420-9101.2008.01538.x)

32. Minias P, Gutiérrez JS, Dunn PO. 2020 Avian major histocompatibility complex copy number variation is associated with helminth richness. *Biol. Lett.* **16**, 20200194. (doi:10.1098/rsbl.2020.0194)

33. He K, Minias P, Dunn PO. 2021 Long-Read Genome Assemblies Reveal Extraordinary Variation in the Number and Structure of MHC Loci in Birds. *Genome Biol. Evol.* **13**. (doi:10.1093/gbe/evaa270)

34. Hadfield J. 2010 MCMC methods for Multi–response Generalised Linear Mixed Models: The MCMCglmm R Package. *J. Stat. Softw.* **33**, 1–22. (doi:10.1002/ana.22635)

35. BABIK W. 2010 Methods for MHC genotyping in non-model vertebrates. *Mol. Ecol. Resour.* **10**, 237–251. (doi:10.1111/J.1755-0998.2009.02788.X)

36. Rekdal SL, Anmarkrud JA, Johnsen A, Lifjeld JT. 2018 Genotyping strategy matters when analyzing hypervariable major histocompatibility complex-Experience from a passerine bird. *Ecol. Evol.* **8**, 1680–1692. (doi:10.1002/ece3.3757)

37. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012 The global diversity of birds in space and time. *Nature* **491**, 444–448. (doi:10.1038/nature11631)

38. Hackett SJ *et al.* 2008 A Phylogenomic Study of Birds Reveals Their Evolutionary History. *Science (80-.).* **320**, 1763–8. (doi:10.1126/science.1157704)

39. Sukumaran J, Holder MT. 2010 DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**, 1569–1571. (doi:10.1093/bioinformatics/btq228)

40. R Core Team. 2020 R: a language and environment for statistical computing.

41. Minias P, Włodarczyk R, Remisiewicz M, Cobzaru I, Janiszewski T. 2021 Distinct evolutionary trajectories of MHC class I and class II genes in Old World finches and buntings. *Heredity* ~~*(Edinb)*~~**126**, 974–990~~., 1–17~~. (doi:10.1038/s41437-021-00427-8)

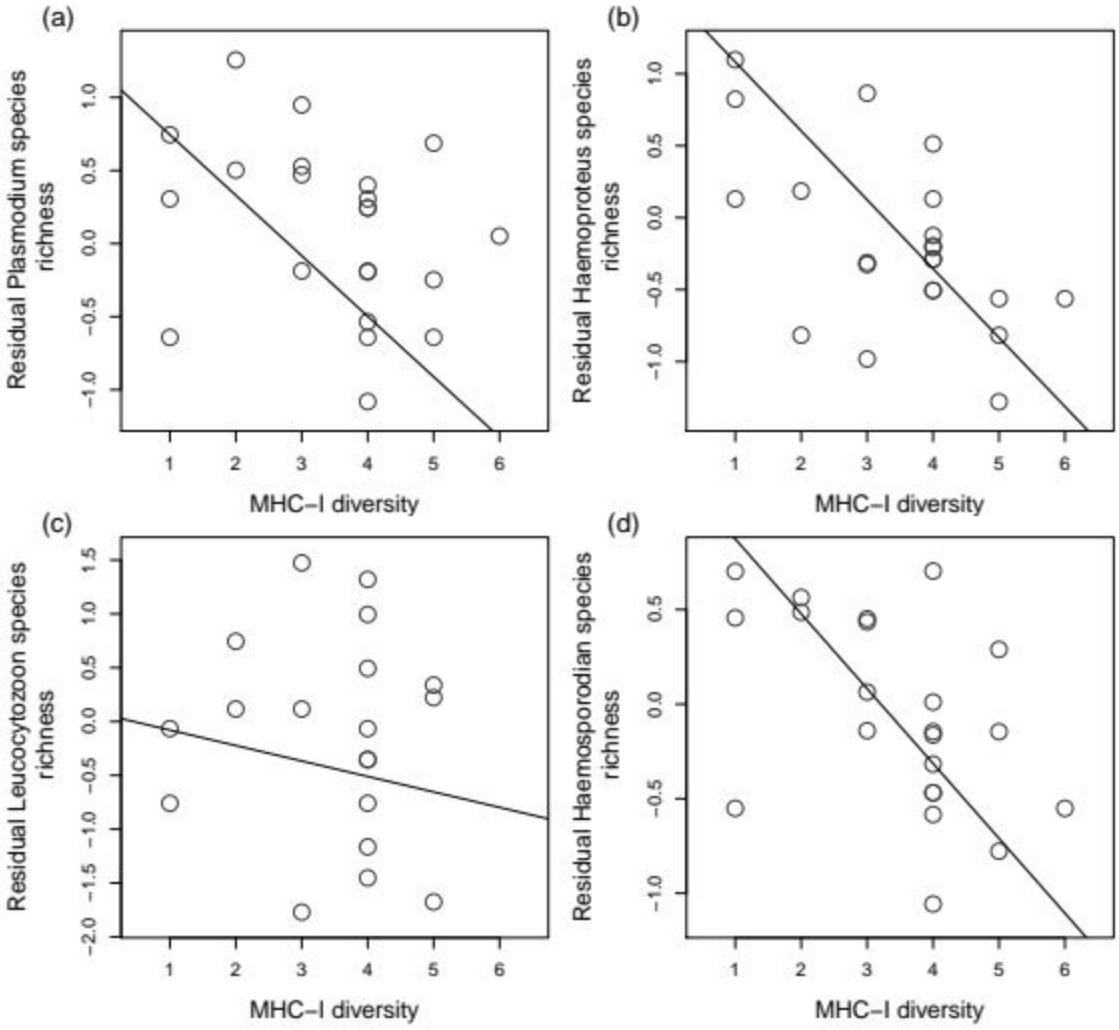
42. Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014 Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Mol. Ecol.* **23**, 4831–4845. (doi:10.1111/MEC.12846)

43~~2~~. RADWAN J, ZAGALSKA-NEUBAUER M, CICHOŃ M, SENDECKA J, KULMA K, GUSTAFSSON L, BABIK W. 2012 MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Mol. Ecol.* **21**, 2469–2479. (doi:10.1111/j.1365-294X.2012.05547.x)

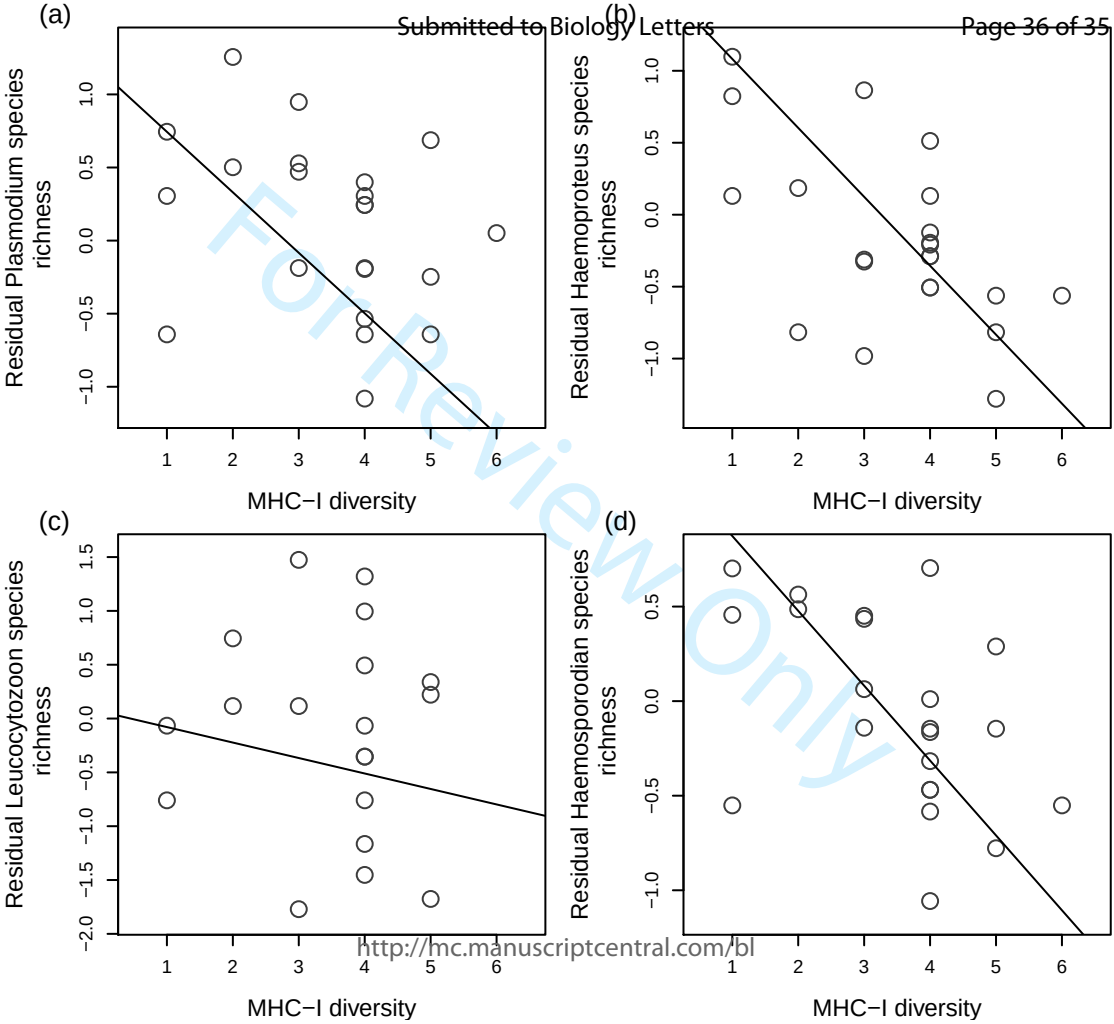
44. Råberg L. 2014 How to Live with the Enemy: Understanding Tolerance to Parasites. *PLoS Biol.* **12**, e1001989. (doi:10.1371/journal.pbio.1001989)

45~~3~~. Migalska M, Sebastian A, Radwan J. 2019 Major histocompatibility complex class I diversity limits the repertoire of T cell receptors. *Proc. Natl. Acad. Sci.* **116**, 5021–5026. (doi:10.1073/pnas.1807864116)

425 **Figure 1.** Association between residual blood parasite species richness and MHC-I
426 diversity for (a) *Plasmodium*, (b) *Haemoproteus*, (c) *Leucocytozoon* and (d) overall
427 haemosporidian lineages richness. Parasite diversity residuals were obtained from a log-log
428 linear regressions between parasite lineages richness (dependent) and research effort
429 (explanatory). Slopes were obtained from phylogenetic MCMCglmms with poisson error
430 distribution between parasite lineages richness and research effort as well as MHC-I
431 diversity.



433



For Review Only