Selection Balancing at Innate Immune Genes: Adaptive Polymorphism Maintenance in Toll-Like Receptors

Piotr Minias [1,*] and Michal Vinkler [2,*]

[1]Department of Biodiversity Studies and Bioeducation, Faculty of Biology and Environmental Protection, University of Łódź, Banacha 1/3, 90-237 Łódź, Poland

[2]Faculty of Science, Department of Zoology, Charles University, Viničná 7, 128 43 Prague, Czech Republic

***Corresponding authors:** E-mails: pminias@op.pl; michal.vinkler@natur.cuni.cz.

Associate editor: Xuming Zhou

Abstract

Balancing selection is a classic mechanism for maintaining variability in immune genes involved in host–pathogen interactions. However, it remains unclear how widespread the mechanism is across immune genes other than the major histocompatibility complex (*MHC*). Although occasional reports suggest that balancing selection (heterozygote advantage, negative frequency-dependent selection, and fluctuating selection) may act on other immune genes, the current understanding of the phenomenon in non-MHC immune genes is far from solid. In this review, we focus on Toll-like receptors (*TLRs*), innate immune genes directly involved in pathogen recognition and immune response activation, as there is a growing body of research testing the assumptions of balancing selection in these genes. After reviewing infection- and fitness-based evidence, along with evidence based on population allelic frequencies and heterozygosity levels, we conclude that balancing selection maintains variation in *TLRs*, though it tends to occur under specific conditions in certain evolutionary lineages rather than being universal and ubiquitous. Our review also identifies key gaps in current knowledge and proposes promising areas for future research. Improving our understanding of host–pathogen interactions and balancing selection in innate immune genes are increasingly important, particularly regarding threats from emerging zoonotic diseases.

Key words: **balancing selection, host–pathogen interactions, innate immune genes, polymorphism, Toll-like receptors, TLR.**

Introduction

Pathogens are strong agents of selective pressure (Schmid-Hempel 2011) and genes coding for molecules that interact closely with pathogens are often key targets of natural selection (Nielsen et al. 2005; Fumagalli and Sironi 2014; Shultz and Sackton 2019). During infection, pathogens interact directly with host molecular components, primarily those aimed at antigen recognition and infection clearance. As such, immune defense is dependent on a myriad of molecular bonds between host and pathogen structures (Kaspers et al. 2022), resulting in either pathogen tolerance or triggering an immune response leading to resistance (Råberg et al. 2007). Although relatively subtle on the scale of an entire host phenotype, molecular variability could have crucial effects on host resistance to pathogens. However, evolutionary mechanisms maintaining variability in immune genes, here defined as any gene in an organism's genome essentially related to immune defense (i.e., being a part of the essential immunome sensu Ortutay et al. 2007), remain unclear. Though many immune genes show high levels of potentially functional variation (Minias et al. 2018; Velová et al. 2018), and a general theoretical framework to explain such variation has been proposed (Woolhouse et al. 2002), we still have little evidence to support current hypotheses explaining the evolutionary history of particular immune genes (Vinkler et al. 2022). Antigen-presenting genes of adaptive immunity, that is, the major histocompatibility complex (*MHC*), together with genes encoding proteins responsible for trimming and loading peptides for MHC presentation (e.g., *ERAP* genes) and receptors that bind loaded *MHC* molecules (e.g., *KIR* genes), represent a notable exception, being well-established targets of balancing selection (Hughes and Yeager 1998; Key et al. 2014; Radwan et al. 2020). In this review, we present current evidence on the evolutionary mechanisms (i.e., balancing selection) maintaining molecular variation in Toll-like receptors (*TLRs*), a group of innate immune genes crucial for triggering inflammation.

Toll-Like Receptors—Function, Selection, and Diversity

In contrast to the *MHC*, innate immune genes were, until recently, viewed as invariant and evolutionarily conserved, primarily as sequence variation in most innate immune genes is predominantly limited by negative (purifying) selection, being driven by functional constraints that select

Review

for general structural conservation. Nevertheless, recent insights into the immunogenetics of both domestic/wild animals and humans have revealed unexpected intraspecific and interspecific variability in several innate immune gene families, including *TLRs* (Lazarus et al. 2002; Novák 2014; Webb et al. 2015; Vinkler et al. 2022). *TLRs* are classed as pattern recognition receptors (PRRs) that form a direct molecular interface between the host and pathogen-derived structures (Vinkler and Albrecht 2009), allowing the host to detect infection and trigger an immune response against the pathogen. Animal taxa have differing sets of *TLRs*, ranging from around 30 in amphioxi (Ji et al. 2018) down to 10–13 genes in birds and mammals (Roach et al. 2005). Individual *TLR* molecules have evolved to sense a diverse range of danger signals through recognition of structurally distinct ligands, some of which are few and invariant and others numerous and highly variable. As an example, *TLR3* only binds to structurally invariant viral RNA, whereas *TLR4* detects host self (e.g., fibrinogen, heat-shock proteins, or endogenous oxidized phospholipids) or nonself (e.g., lipopolysaccharide, mannan, glycoinositol-phospholipids, pneumolysin, or viral envelope and fusion proteins) ligands (Piccinini and Midwood 2010; Kumar et al. 2011).

Timely and specific pathogen recognition, essential for efficient pathogen clearance, represents a strong selective force in host–pathogen interactions. As such, *TLR* genetic variability is expected to be driven by positive selection, promoting fixation and maintenance of diverse nonsynonymous substitutions and allelic variants across different taxa and evolutionary lineages. In general, ineffective pathogen recognition in hosts increases the costs of immune defense, both in terms of resource allocation and tissue damage (Ashley et al. 2012). Since structural variation in *TLR* ligands may evolve to allow a pathogen to escape the host's immune defense, possibly even leading to host mortality (Zdorovenko et al. 2007; Wang et al. 2015), reciprocal host–pathogen evolution is predicted as an outcome. This is consistent with the coevolutionary arms race dynamics expected under the "Red Queen model" (Woolhouse et al. 2002). *TLRs* represent a valuable model for studying such evolutionary mechanisms as (1) they show strong positive selection signatures at the interspecific level (Velová et al. 2018) and (2) unlike *MHC*, they are single-copy genes, allowing more effective tracking of adaptive evolution. Furthermore, their general structural conservation allows positive selection to be linked with protein structure molecular adaptations (Těšický et al. 2020), allowing functional identification of the phenotypic effects of molecular variation (Walsh et al. 2008; Levy et al. 2020; Fiddaman et al. 2022).

Although only around 5% of *TLR* codons experience pathogen-mediated positive selection (Grueber et al. 2014; Velová et al. 2018), this polymorphism may play a paramount role in pathogen recognition as it is mostly located in the ligand-binding regions of *TLR* exodomains, that is, at molecular surfaces crucial for pathogen recognition (Downing et al. 2010; Vinkler et al. 2014; Velová et al. 2018). *TLR* adaptations at these sites may affect the physicochemical properties of the receptor surface, including its electrostatic potential (Králová et al. 2018; Těšický et al. 2020). In pathogens, and particularly those with simple genomes such as bacteria or (especially) viruses, convergent evolution is common, indicating functional limits in their structural evolution (Wang et al. 2015; Gutierrez et al. 2019). Given these limits to the number of structural variants of pathogen-derived ligands, diversifying selection can also be predicted to select for a finite (but possibly multiple) number of *TLR* variants. At the interspecific level, constraints in host and pathogen structure variation may lead to evolutionary convergence (Walsh et al. 2008; Świderská et al. 2018; Těšický et al. 2020), whereas *TLR* polymorphism at the intraspecific level may be maintained within populations via balancing selection. Though nonnegligible levels of functional, and potentially functional, variability at *TLRs* has recently been revealed across and within phylogenetically diverse vertebrate lineages, including humans (Ferwerda et al. 2007), domestic animals (Leveque et al. 2003; Świderská et al. 2018), and wild species (Alcaide and Edwards 2011; Quéméré et al. 2015; Vinkler et al. 2015; Minias et al. 2021), the mechanisms of balancing selection actively maintaining this variation within populations are still debated.

Mechanisms of Balancing Selection

Although novel allelic variants are generated by de novo mutations, negative selection can subsequently remove alleles that become nonadvantageous, whereas positive selection can increase the frequency of advantageous alleles, possibly leading to fixation (according to the selective sweep mechanism; de Groot et al. 2002). Alternatively, before fixation is reached, balancing selection may intervene to maintain allelic variation at frequencies greater than expected under neutral evolution.

Balancing selection acts through several nonexclusive mechanisms. First, heterozygote advantage (overdominance) assumes that heterozygous genotypes have greater fitness than homozygous genotypes, which may select for maintenance of multiple alleles within populations (fig. 1). As in the case of the *MHC*, two different alleles at the same *TLR* locus should allow hosts to sense a broader spectrum of ligands, thereby triggering an immune response against a greater range of foreign pathogens (Hedrick 2012). Thus, according to theoretical predictions, higher heterozygosity across multiple *TLR* loci (or within a single *TLR* locus) should be associated with higher fitness resulting from more efficient protection against multiple pathogens.

Second, negative frequency-dependent selection assumes that low-frequency allelic variants confer greater fitness than more common alleles (rare allele advantage hypothesis; fig. 1). This is based on the premise of dynamic host–pathogen coevolution, where pathogens rapidly evolve to overcome (avoid) the most common immune defenses of their hosts (Takahata and Nei 1990). Thus,

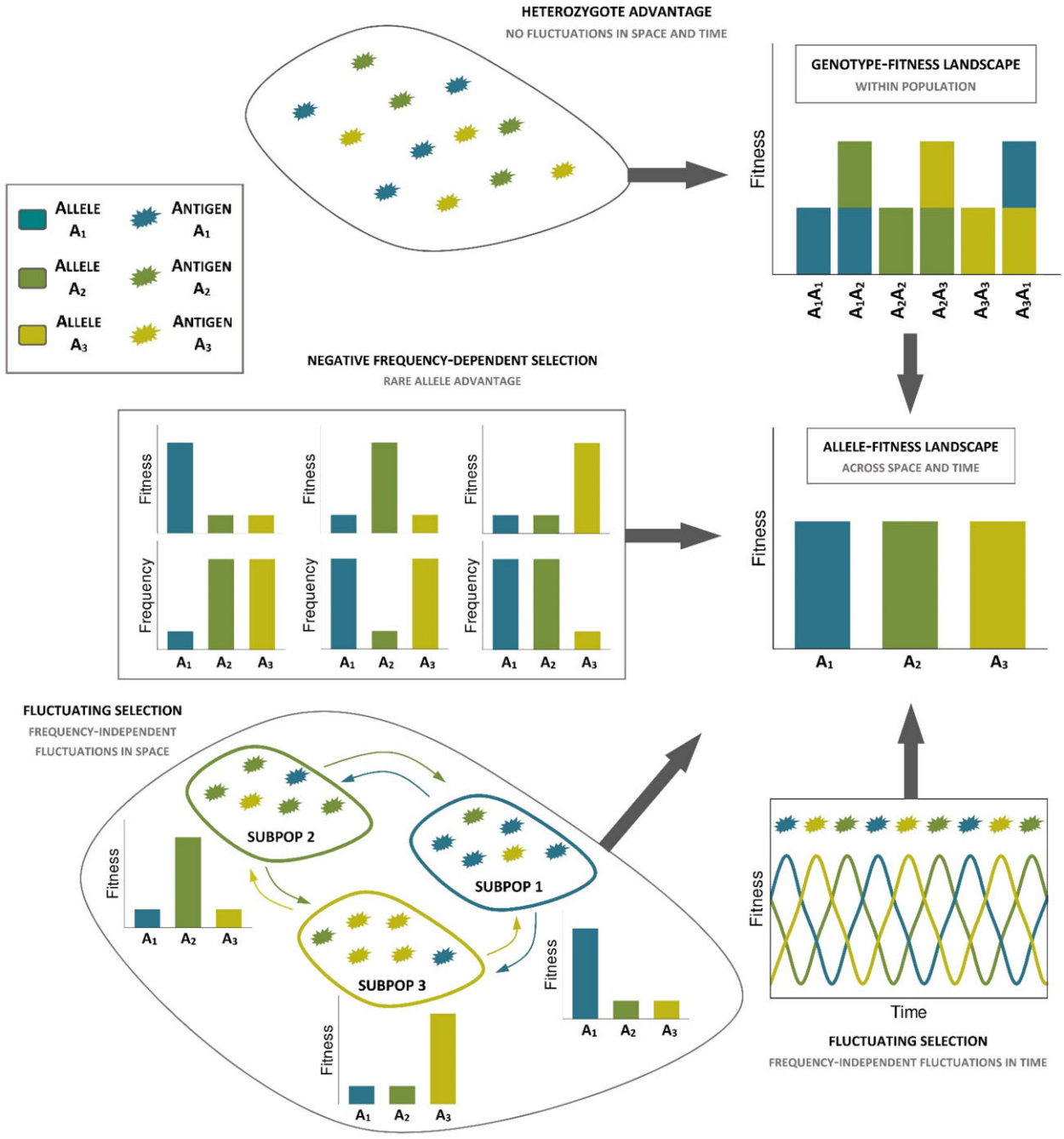

FIG. 1. Mechanisms of balancing selection that can maintain Toll-like receptor diversity within populations. Heterozygote advantage assumes that heterozygous genotypes have higher fitness than homozygous genotypes, as they recognize a broader spectrum of antigens (pathogens). Negative frequency-dependent selection assumes that low-frequency genotypes (alleles) have higher fitness, as they are not avoided by high frequency pathogens (rare allele advantage). Fluctuating selection assumes frequency-independent parasite-driven fluctuations in genotype or allele frequencies between subpopulations or years. Under all these mechanisms, diverse alleles may have similar fitness effects across space and time.

increasing allelic frequency should be accompanied by decreasing fitness, which is expected to maintain different alleles at moderate frequencies within populations and prevent a selective sweep.

Finally, fluctuating selection assumes that the landscape of pathogen-mediated selection on hosts and their immune genes should vary in space and time (Charbonnel and Pemberton 2005). These spatio-temporal fluctuations may be independent of allele frequency and, in contrast to the negative frequency-dependent selection hypothesis, exclusively reflect variation in environmental factors affecting the composition of local pathogen communities (Spurgin and Richardson 2010). Since different allelic variants may confer varying fitness in different habitats or locations (subpopulations), and intensity of directional positive selection on these alleles can change over time, multiple alleles are expected to be maintained within a metapopulation (across subpopulations; fig. 1).

So far, the most convincing empirical evidence for balancing selection comes from research on antigen-presenting genes of adaptive immunity, that is, the *MHC* (reviewed in Radwan et al. 2020). However, the *MHC* is specific in its genomic structure (multi-locus organization with frequent pseudogenisation; Sepil et al. 2012) and immunological function (functional dependency on T cell receptors; Migalska et al. 2019). Thus, it remains unclear how widespread balancing selection is in immune genes. It has been estimated that over half of genetic variability for resistance to infection is attributable to non-*MHC* genes (Acevedo-Whitehouse and Cunningham 2006), suggesting that other well-supported examples of genes evolving under balancing selection could yet be identified. Although balancing selection is recognized as the main force shaping *TLR* gene evolution in humans (Ferrer-Admetlla et al. 2008), and *TLRs* in many animal species show surprisingly high allelic variation (Alcaide and Edwards 2011; Świderská et al. 2018), it remains to be established whether the processes of balancing selection act at these key innate immune genes in nonhuman vertebrates and how general they are across different evolutionary lineages. Here, we review current empirical evidence for mechanisms of balancing selection acting on *TLRs* in natural populations of nonhuman vertebrates. Although precise determination of the processes mediating balancing selection on immune genes is difficult as they are nonexclusive and can operate in parallel (Spurgin and Richardson 2010), we attempt to identify those mechanisms that are likely to be of primary importance for the maintenance of *TLR* diversity.

TLR Heterozygote Advantage

Of all the mechanisms associated with balancing selection on TLRs, heterozygote advantage has received the most extensive scientific attention, possibly due to the methodological feasibility of testing for this mechanism. Traditionally, heterozygote advantage has been tested by examining covariation between *TLR* heterozygosity status and infection rate, fitness components (reproduction and survival), or fitness-related traits (e.g., condition or ornament expression). In general, heterozygous individuals are expected to show lower overall rates of infection by multiple pathogens, and thus display better condition, better expression of condition-dependent traits (e.g., ornaments) and, eventually, higher fitness.

Infection Rates

To date, most correlative research on *TLR* heterozygote advantage has focused on infection rates; however, the mechanism has received scant and/or indirect support. For example, Gavan et al. (2015) showed that different *TLR4* alleles in the water vole *Arvicola amphibius* were associated with lower infection rates by different parasites (*Megabothris* fleas, *Ixodes* tick larvae, and *Gamasidae* mites). Though not tested directly, the authors interpreted this pattern as heterozygote advantage at *TLR4*, a hypothesis further supported by the excess of heterozygote individuals observed before a population bottleneck (Gavan et al. 2015). A similar mechanism has been invoked in the roe deer *Capreolus capreolus*, in which different *TLR2* alleles provided resistance against different pathogens and the allele resistant against *Toxoplasma* infection conferred susceptibility to *Chlamydia*, and vice versa (Quéméré et al. 2021). The same study reported that individuals with a medium-frequency (31%) *TLR5* allele in a heterozygous (but not homozygous) state were less likely to be seropositive for *Chlamydia* than individuals lacking this allele (Quéméré et al. 2021). However, as *TLR5* is not expected to be directly involved in *Chlamydia* detection (Verweij et al. 2016), the nature of this association remains unclear. Despite evidence suggesting the benefits of heterozygosity (in line with the expected molecular mechanism of heterozygote advantage), numerous studies have provided support for the effects of specific *TLR* alleles (i.e., lower infection rates in individuals with specific resistance alleles, independent of heterozygosity status), some even suggesting a heterozygote disadvantage effect, with lower infection rates in homozygous individuals (table 1). Further, most studies have failed to find any significant association between *TLR* heterozygosity and infection rate, at least in some of the loci tested (table 1). It should be acknowledged, however, that most studies examined infections by a single pathogen, or tested for separate effects of multiple pathogens. Since heterozygote advantage at *TLRs* (and other antigen-recognition immune genes) is thought to confer an overall benefit when an organism is exposed to a diverse range of pathogens, single-pathogen experimental framework may be insufficient to test for this mechanism effectively.

Fitness Components and Fitness-Related Traits

As with infection rate, the heterozygote advantage mechanism receives little support when examined in relation to fitness components and fitness-related traits. For example, testing for correlations between *TLR* heterozygosity and reproductive success in the dunnock *Prunella modularis* conditionally supported heterozygote advantage, but only at a single *TLR3* locus (Lara et al. 2020). It has been suggested that *TLR3* heterozygous individuals may cope better against RNA viruses, such as avian influenza; however, heterozygote advantage has only been detected in males and this could possibly be explained by intricate sexual conflict (Santos et al. 2015). Indications of *TLR4* heterozygote advantage have also been found in a heavily bottlenecked population of another passerine bird, the Stewart Island robin *Petroica australis rakiura*. Here, a single heterozygous *TLR4* genotype conferred a large survival advantage, reducing mortality risk before maturity by more than half (Grueber et al. 2013). Strong selection advantage for this genotype was further confirmed by a 2-fold frequency excess in the population compared with the Hardy–Weinberg expectation (Grueber et al. 2013). Interestingly, elevated levels of *TLR* heterozygosity resulting from reciprocal translocation of robins between inbred

Table 1. List of Studies Testing for Heterozygote Advantage at *TLRs* in Nonhuman Vertebrates.

Trait Category	Lineage	Species	Pathogen/Trait	Heterozygote Advantage	Heterozygote Disadvantage	Single-Allele Effect	No Effect	Reference
Infection rate	Birds	*Coereba flaveola*	*Haemoproteus*		2B		1A, 7	1
	Mammals	*Apodemus sylvaticus*	*Toxoplasma gondii*				11, 12	2
		Arvicola amphibius	*Megabothris* flea			4		3
		Arvicola amphibius	*Ixodes* larvae, Gamasidae mites		4	4		3
		Arvicola amphibius	*Bartonella*, *Ixodes* nymphs				4	3
		Arvicola amphibius	Coinfection	4[a]				3
		Capra ibex	*Brucella melitensis*		1	1	2, 4	4
		Capreolus capreolus	*Toxoplasma*		2	2	4, 5	5
		Capreolus capreolus	*Chlamydia*	5		2	4	5
		Capreolus capreolus	Coinfection	2[a]			4,5	5
		Myodes glareolus	*Borrelia afzelii*		2	2		6,7
Reproduction	Birds	*Acrocephalus sechellensis*	Lifetime reproductive success		3	3		8
		Prunella modularis	Reproductive success	3			1, 2, 4, 5, 15	9
Survival	Birds	*Acrocephalus sechellensis*	Lifetime survival			3		8
		Atlapetes pallidiceps	Annual survival		ML			10
		Melospiza melodia	Annual survival				ML	11
		Petroica australis	Annual survival	4			ML	12
		Tympanuchus cupido attwateri	Annual survival			1B	4, 5, 15	13
Other	Birds	*Chroicocephalus ridibundus*	Physiological condition	1B, 3, ML			4, 5	14
		Sterna hirundo	Colony size choice				1, 3, 4	15
		Tympanuchus cupido attwateri	Immune function			1B	4, 5, 15	13
	Mammals	*Capreolus capreolus*	Natal dispersal	3, 4			2, 5	16

Note.—Numbers indicate identity of *TLR* loci associated with different effects (heterozygote advantage, heterozygote disadvantage, single-allele effect). ML, multi-locus associations.

[a]Associations inferred, but not explicitly tested for.

1, Antonides et al. (2019); 2, Morger et al. (2014); 3, Gavan et al. (2015); 4, Quéméré et al. (2020); 5, Quéméré et al. (2021); 6, Tschirren et al. (2013); 7, Cornetti et al. (2018); 8, Davies et al. (2021); 9, Lara et al. (2020); 10, Hartmann et al. (2014); 11, Nelson-Flower et al. (2018); 12, Grueber et al. (2013); 13, Bateson et al. (2016); 14, Podlaszczuk et al. (2021); 15, Drzewińska-Chańko et al. (2021); 16, Vanpé et al. (2016).

populations may have contributed to observed effects of increased individual survival and recruitment following active conservation measures (Grueber et al. 2017). In the black-headed gull *Chroicocephalus ridibundus*, multi-locus *TLR* heterozygosity (and within-individual sequence diversity) correlated with different physiological condition indices; however, all these associations prevailed at the nucleotide rather than amino acid level, suggesting molecular mechanisms indirectly linked to antigen recognition (e.g., modifications of translation level) (Podlaszczuk et al. 2021).

To date, heterozygosity-fitness correlations at *TLRs* have almost exclusively been tested in birds. In mammals, we are aware of just one study that tested for correlations with natal dispersal, a trait that may have important fitness consequences (Nevoux et al. 2013). This study found that roe deer dispersal propensity and distance increased with heterozygosity at *TLR3* (all individuals) and *TLR4* (heavy individuals only), which could be associated with fitness advantage (Vanpé et al. 2016). Aside from this, information on correlations between *TLR* heterozygosity and fitness components (reproduction and survival) are lacking for mammals.

Other Mechanisms of Balancing Selection at TLRs

Owing to fundamental methodological difficulties, including the need for a long-term spatio-temporal sampling design, any direct empirical evidence for mechanisms of balancing selection acting on *TLRs* other than heterozygote advantage (i.e., negative frequency-dependent and fluctuating selection) is highly limited. Below, we provide a summary of the available evidence.

Infection Rates

Association of *TLR* variation with infection rate may not necessarily provide support for the heterozygote advantage hypothesis; instead, it may reflect other mechanisms of balancing selection, such as spatial or temporal variation in allelic frequency. A striking example has been provided

through a genome-wide association study of Berthelot's pipit *Anthus berthelotii*, which revealed relationships between two single nucleotide polymorphisms (SNPs) in *TLR4* with malaria *Plasmodium* infection (Armstrong et al. 2019). Surprisingly, one of the SNPs was negatively associated with malaria infection risk in the Canary Islands, but positively associated in the Madeira archipelago. No evidence for heterozygote advantage was found in either population, clearly indicating a local adaptation to fluctuating pathogen-driven selection in the pipits, maintaining *TLR4* variation across populations (Armstrong et al. 2019). Similarly, a negative correlation was found between *Amblyomma* tick prevalence and *TLR5* nucleotide and amino acid diversity across fragmented populations of a tropical bird, the wedge-billed woodcreeper *Glyphorynchus spirurus* (Perrin et al. 2021). Since ticks are key vectors of animal infectious diseases (Jongejan and Uilenberg 2004), it has been suggested that *TLR* variation may have been maintained by balancing selection (possibly through different mechanisms). On the other hand, the *TLR5* variation paralleled neutral genetic variation in woodcreepers, suggesting a dominant role of drift (Perrin et al. 2021). This is similar to the distribution of *TLR5* polymorphism in humans and other primates, where, despite selection for variation at the interspecific level, nonfunctional pseudogene alleles may be maintained in a polymorphic state by drift when negative selection is weak (Wlasiuk et al. 2009).

Population Genetics and Population Demographic History

Although multiple studies have identified putative functional polymorphism in *TLRs* in wild animal populations (Alcaide and Edwards 2011; Tschirren et al. 2011; Vinkler et al. 2015) and domestic species (Darfour-Oduro et al. 2016; Świderská et al. 2018), there is virtually no information available on temporal variation in *TLR* allelic frequencies, which would clearly support the negative frequency-dependent selection mechanism. Despite this, several studies have documented notable variation in *TLR* allelic composition across subpopulations, consistent with fluctuating selection. For example, patterns of *TLR2* polymorphism in the bank vole *Myodes glareolus* suggest spatial variation in selective pressures (Tschirren et al. 2012), probably due to *TLR2* allelic variation being linked to resistance to *Borrelia afzelii*, a tick-borne pathogen widespread in wild rodents (Tschirren et al. 2013). Similarly, a large-scale study on the gentoo penguin *Pygoscelis papua* detected population-specific adaptations at *TLR5* to divergent polar environments and local pathogen assemblages (Levy et al. 2020), also invoking selection that fluctuates in space. This population diversification pattern, with gene flow contributing to polymorphism maintenance, is also known for several human *TLRs*, particularly *TLR4* and the *TLR10-1-6* cluster (Ferwerda et al. 2007; Barreiro et al. 2009).

Balancing selection acting within populations could lead to an overrepresentation of several common alleles compared with neutral equilibrium, a phenomenon that can be tested for with a suite of related statistics, such as Tajima's D or Fu and Li's D (Fijarczyk and Babik 2015). Significantly positive values of these statistics have been reported for several vertebrate taxa at different *TLR* loci, suggesting the action of balancing selection (Kloch et al. 2018; Ham-Dueñas et al. 2020; Xu et al. 2020). On the other hand, many other *TLR* studies have reported no deviation from neutrality for Tajimas' D (e.g., Hartmann et al. 2014; Levy et al. 2020; Podlaszczuk et al. 2021). It should be acknowledged, however, that these neutrality tests are extremely sensitive to nonequilibrium demography and population structure (Fijarczyk and Babik 2015). This background noise can be filtered out, to certain extent, by comparing allele frequencies in genes predicted to evolve under balancing selection with neutral markers. Application of such an approach in a bottlenecked population of the Seychelles warbler *Acrocephalus sechellensis* showed a heterozygote excess in *TLR15*, which retained more variation than would be predicted based on observations of neutral markers, supporting the role of balancing selection (Gilroy, van Oosterhout, et al. 2017). Further, demographic simulations in this species revealed that variation observed at three *TLR* genes (*TLR1LB*, *TLR3*, and *TLR15*) could not be explained by neutral evolution exclusively, and was more likely generated through past balancing selection (Gilroy, Phillips, et al. 2017). Unfortunately, this evidence for balancing selection cannot effectively distinguish between rare allele advantage and heterozygote advantage mechanisms since heterozygote excess (when compared with neutral Hardy–Weinberg expectations) can be driven by either or both processes (Spurgin and Richardson 2010). Finally, there is limited evidence for balancing selection in *TLRs* originating from genomic scans of nonsynonymous polymorphisms within populations. In brief, balancing selection is expected to maintain excess diversity not only at the targets of selection but also at neighboring neutral loci, detectable through comparisons with a genome-wide distribution of nonsynonymous diversity (Fijarczyk and Babik 2015). This pattern was revealed at three *TLR* genes (*TLR1*, *TLR2*, and *TLR6*) in the bank vole, though interpretation of the results was complicated by paralogous gene conversion at two of the loci (Lundberg et al. 2020). Interestingly, of the 135 PRR signaling pathway genes in bank voles, the *TLR* family (together with C-type lectin receptors) appeared to be a key target of balancing selection (Lundberg et al. 2020).

Conclusions

Consistent with the "One Health" concept, an understanding of the mechanisms responsible for polymorphism maintenance in innate immune genes of wild animals is essential for predicting risks associated with animal to human disease transmission and preventing future disease outbreaks (Lebov et al. 2017). Though still fragmentary, multiple lines of evidence from a wide range of species suggest that balancing selection is likely to be a common

mechanism for polymorphism maintenance in *TLRs*. At the same time, contrasting evidence for different taxa and *TLR* genes (either supporting balancing selection or not) indicates that the phenomenon is far from being universal and omnipresent, but occurs under specific conditions in certain evolutionary lineages. Although distinguishing between different types of balancing selection in wild species is difficult, the best empirical evidence relates to fluctuating selection maintaining *TLR* variation across populations. In contrast, evidence for heterozygote advantage is still relatively soft, and there is no evidence for negative frequency-dependent selection. Here, we identify several limitations currently restricting our understanding of balancing selection acting in vertebrate *TLRs*:

- Lack of evidence for functional variation in *TLR* alleles—though selection operates through phenotypic variation, any links between natural *TLR* genetic variation, differences in molecular function, and variance in immune responsiveness are missing for wild species. Notably, these associations may be predicted as positive (strengthening pathogen recognition) or negative (potentially leading to immunological tolerance).
- Shortage of association studies between *TLR* variation and infection rates—as immunological variation should manifest as altered disease resistance in the natural environment, there is a need for data identifying links between *TLR* variation and diverse pathogens (coinfections). It should be stressed, however, that *TLR* variation may also be selected for by interactions with nonpathogenic symbionts.
- Insufficient association studies between *TLR* variation and fitness components—whereas linking polymorphism maintained in *TLRs* with fitness varying across different contexts (via quantification of allele- and genotype-fitness landscapes) can serve as the most straightforward evidence for balancing selection, relatively limited effort has been made to test for this association across species.
- Low phylogenetic coverage—until now, research on intraspecific variation in *TLRs* has focused primarily on birds and mammals, whereas other vertebrate lineages with more diversified *TLRs* (e.g., fish) are underrepresented. Improved knowledge of pathogen community diversity across evolutionarily divergent host species could guide the search for promising models on balancing selection in *TLRs*.
- Lack of wider spatial datasets and time series—broader information on the spatio-temporal dynamics of *TLR* allele frequency changes potentially linked with changes in pathogen composition would provide a basis for robust inferences on the roles of negative frequency-dependent selection and fluctuating selection in shaping *TLR* diversity.

Although filling the gaps outlined above will be both technically and financially demanding, focusing on *TLRs* as a model genetic system for investigating balancing selection offers several advantages. First, being directly involved in interactions with structurally diversified pathogen structures, *TLRs* are among those proteins where polymorphism maintenance is predictable. Second, their general conservation enhances effective screening of within- and between-population polymorphism, allowing efficient development of molecular tools (primers, probes), robust protein modeling, and prediction of variation in protein molecular phenotypes. Third, their well-understood immunological function makes *TLRs* easy to investigate functionally, including the adoption of cutting edge experimental approaches such as allelic variant manipulation. Fourth, given specific interactions of individual *TLRs* with pathogen-derived structures, precise predictions can be made about their associations with infection agents. Taking into consideration the existing body of evidence for and against balancing selection in *TLRs*, new research should be targeted at gaining a better understanding of the ecological and evolutionary contexts that promote (or inhibit) its role in shaping innate immune gene variation. Further, we recommend that future studies should focus primarily on overcoming the limitations listed above to provide comprehensive evidence for the broad action of balancing selection in immune genes other than the *MHC*, which still remains a key target of evolutionary immunological research.

Acknowledgements

We are grateful to Kevin F. Roche for his language correction. M.V. was supported by the Czech Science Foundation (Grant No. P502/19-20152Y). We thank two anonymous reviewers for constructive comments on an earlier draft of the manuscript.

Data availability

No new data were generated or analyzed in support of this research.

References

Acevedo-Whitehouse K, Cunningham AA. 2006. Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol.* **21**: 433–438.

Alcaide M, Edwards SV. 2011. Molecular evolution of the toll-like receptor multigene family in birds. *Mol Biol Evol.* **28**:1703–1715.

Antonides J, Mathur S, Sundaram M, Ricklefs R, DeWoody JA. 2019. Immunogenetic response of the bananaquit in the face of malarial parasites. *BMC Evol Biol.* **19**:107.

Armstrong C, Davies RG, González-Quevedo C, Dunne M, Spurgin LG, Richardson DS. 2019. Adaptive landscape genetics and malaria across divergent island bird populations. *Ecol Evol.* **9**: 12482–12502.

Ashley NT, Weil ZM, Nelson RJ. 2012. Inflammation: mechanisms, costs, and natural variation. *Annu Rev Ecol Evol Syst.* **43**:385–406.

Barreiro LB, Ben-Ali M, Quach H, Laval G, Patin E, Pickrell JK, Bouchier C, Tichit M, Neyrolles O, Gicquel B, *et al.* 2009. Evolutionary dynamics of human toll-like receptors and their different contributions to host defense. *PLoS Genet.* **5**:e1000562.

Bateson ZW, Hammerly SC, Johnson JA, Morrow ME, Whittingham LA, Dunn PO. 2016. Specific alleles at immune genes, rather than genome-wide heterozygosity, are related to immunity and survival in the critically endangered Attwater's prairie-chicken. *Mol Ecol.* **25**:4730–4744.

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

Cornetti L, Hilfiker D, Lemoine M, Tschirren B. 2018. Small-scale spatial variation in infection risk shapes the evolution of a *Borrelia* resistance gene in wild rodents. *Mol Ecol.* **27**:3515–3524.

Darfour-Oduro KA, Megens HJ, Roca A, Groenen MA, Schook LB. 2016. Evidence for adaptation of porcine Toll-like receptors. *Immunogenetics* **68**:179–189.

Davies CS, Taylor MI, Hammers M, Burke T, Komdeur J, Dugdale HL, Richardson DS. 2021. Contemporary evolution of the innate immune receptor gene TLR3 in an isolated vertebrate population. *Mol Ecol.* **30**:2528–2542.

de Groot NG, Otting N, Doxiadis GG, Balla-Jhagjhoorsingh SS, Heeney JL, van Rood JJ, Gagneux P, Bontrop RE. 2002. Evidence for an ancient selective sweep in the MHC class I gene repertoire of chimpanzees. *Proc Natl Acad Sci U S A.* **99**:11748–11753.

Downing T, Lloyd AT, O'Farrelly C, Bradley DG. 2010. The differential evolutionary dynamics of avian cytokine and TLR gene classes. *J Immunol.* **184**:6993–7000.

Drzewińska-Chańko J, Włodarczyk R, Gajewski A, Rudnicka K, Dunn PO, Minias P. 2021. Immunocompetent birds choose larger breeding colonies. *J Anim Ecol.* **90**:2325–2335.

Ferrer-Admetlla A, Bosch E, Sikora M, Marquès-Bonet T, Ramírez-Soriano A, Muntasell A, Navarro A, Lazarus R, Calafell F, Bertranpetit J, Casals F. 2008. Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol.* **181**:1315–1322.

Ferwerda B, McCall MB, Alonso S, Giamarellos-Bourboulis EJ, Mouktaroudi M, Izagirre N, Syafruddin D, Kibiki G, Cristea T, Hijmans A, *et al.* 2007. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. *Proc Natl Acad Sci U S A.* **104**:16645–16650.

Fiddaman SR, Vinkler M, Spiro SG, Levy H, Emerling CA, Boyd AC, Dimopoulos EA, Vianna JA, Cole TL, Pan H, *et al.* 2022. Adaptation and cryptic pseudogenization in penguin toll-like receptors. *Mol Biol Evol.* **39**:msab354.

Fijarczyk A, Babik W. 2015. Detecting balancing selection in genomes: limits and prospects. *Mol Ecol.* **24**:3529–3545.

Fumagalli M, Sironi M. 2014. Human genome variability, natural selection and infectious diseases. *Curr Opin Immunol.* **30**:9–16.

Gavan MK, Oliver MK, Douglas A, Piertney SB. 2015. Gene dynamics of toll-like receptor 4 through a population bottleneck in an insular population of water voles (*Arvicola amphibius*). *Conserv Genet.* **16**:1181–1193.

Gilroy DL, Phillips KP, Richardson DS, van Oosterhout C. 2017. Toll-like receptor variation in the bottlenecked population of the Seychelles warbler: computer simulations see the 'ghost of selection past'and quantify the 'drift debt'. *J Evol Biol.* **30**:1276–1287.

Gilroy DL, van Oosterhout C, Komdeur J, Richardson DS. 2017. Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Anim Conserv.* **20**:235–250.

Grueber CE, Sutton JT, Heber S, Briskie JV, Jamieson IG, Robertson BC. 2017. Reciprocal translocation of small numbers of inbred individuals rescues immunogenetic diversity. *Mol Ecol.* **26**:2660–2673.

Grueber CE, Wallis GP, Jamieson IG. 2013. Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a reintroduced population of a threatened species. *Mol Ecol.* **22**:4470–4482.

Grueber CE, Wallis GP, Jamieson IG. 2014. Episodic positive selection in the evolution of avian Toll-like receptor innate immunity genes. *PLoS One.* **9**:e89632.

Gutierrez B, Escalera-Zamudio M, Pybus OG. 2019 Parallel molecular evolution and adaptation in viruses. *Curr Opin Virol.* **34**:90–96.

Ham-Dueñas JG, Canales-del-Castillo R, Voelker G, Ruvalcaba-Ortega I, Aguirre-Calderón CE, González-Rojas JI. 2020. Adaptive genetic diversity and evidence of population genetic structure in the endangered Sierra Madre Sparrow (*Xenospiza baileyi*). *PLoS One.* **15**:e0232282.

Hartmann SA, Schaefer HM, Segelbacher G. 2014. Genetic depletion at adaptive but not neutral loci in an endangered bird species. *Mol Ecol.* **23**:5712–5725.

Hedrick PW. 2012. What is the evidence for heterozygote advantage selection? *Trends Ecol Evol.* **27**:698–704.

Hughes AL, Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet.* **32**:415–435.

Ji J, Ramos-Vicente D, Navas-Pérez E, Herrera-Úbeda C, Lizcano JM, Garcia-Fernàndez J, Escrivà H, Bayés À, Roher N. 2018. Characterization of the TLR family in *Branchiostoma lanceolatum* and discovery of a novel TLR22-like involved in dsRNA recognition in amphioxus. *Front Immunol.* **9**:2525.

Jongejan F, Uilenberg G. 2004. The global importance of ticks. *Parasitology* **129**:S3–S14.

Kaspers B, Schat KA, Göbel TW, Vervelde L. 2022. *Avian immunology*. 3rd ed. London: Academic Press.

Key FM, Teixeira JC, de Filippo C, Andrés AM. 2014. Advantageous diversity maintained by balancing selection in humans. *Curr Opin Genet Dev.* **29**:45–51.

Kloch A, Wenzel MA, Laetsch DR, Michalski O, Bajer A, Behnke JM, Welc-Falęciak R, Piertney SB. 2018. Signatures of balancing selection in toll-like receptor (TLRs) genes—novel insights from a free-living rodent. *Sci Rep.* **8**:8361.

Králová T, Albrecht T, Bryja J, Hořák D, Johnsen A, Lifjeld JT, Novotný M, Sedláček O, Velová H, Vinkler M. 2018. Signatures of diversifying selection and convergence acting on passerine Toll-like receptor 4 in an evolutionary context. *Mol Ecol.* **27**:2871–2883.

Kumar H, Kawai T, Akira S. 2011. Pathogen recognition by the innate immune system. *Int Rev Immunol.* **30**:16–34.

Lara CE, Grueber CE, Holtmann B, Santos ES, Johnson SL, Robertson BC, Castaño-Villa GJ, Lagisz M, Nakagawa S. 2020. Assessment of the dunnocks' introduction to New Zealand using innate immune-gene diversity. *Evol Ecol.* **34**:803–820.

Lazarus R, Vercelli D, Palmer LJ, Klimecki WJ, Silverman EK, Richter B, Riva A, Ramoni M, Martinez FD, Weiss ST, Kwiatkowski DJ. 2002. Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease. *Immunol Rev.* **190**:9–25.

Lebov J, Grieger K, Womack D, Zaccaro D, Whitehead N, Kowalcyk B, MacDonald PDM. 2017. A framework for one health research. *One Health.* **3**:44–50.

Leveque G, Forgetta V, Morroll S, Smith AL, Bumstead N, Barrow P, Loredo-Osti JC, Morgan K, Malo D. 2003. Allelic variation in TLR4 is linked to susceptibility to *Salmonella enterica* serovar Typhimurium infection in chickens. *Infect Immun.* **71**:1116–1124.

Levy H, Fiddaman SR, Vianna JA, Noll D, Clucas GV, Sidhu JK, Polito MJ, Bost CA, Phillips RA, Crofts S, *et al.* 2020. Evidence of pathogen-induced immunogenetic selection across the large geographic range of a wild seabird. *Mol Biol Evol.* **37**:1708–1726.

Lundberg M, Zhong X, Konrad A, Olsen RA, Råberg L. 2020. Balancing selection in pattern recognition receptor signalling pathways is associated with gene function and pleiotropy in a wild rodent. *Mol Ecol.* **29**:1990–2003.

Migalska M, Sebastian A, Radwan J. 2019. Major histocompatibility complex class I diversity limits the repertoire of T cell receptors. *Proc Natl Acad Sci U S A.* **116**:5021–5026.

Minias P, Drzewinska-Chanko J, Wlodarczyk R. 2021. Evolution of innate and adaptive immune genes in a non-model waterbird, the common tern. *Infect Genet Evol.* **95**:105069.

Minias P, Pikus E, Whittingham LA, Dunn PO. 2018. A global analysis of selection at the avian MHC. *Evolution.* **72**:1278–1293.

Morger J, Bajnok J, Boyce K, Craig PS., Rogan MT, Lun ZR, Hide G, Tschirren B. 2014. Naturally occurring Toll-like receptor 11 (TLR11) and Toll-like receptor 12 (TLR12) polymorphisms are not associated with *Toxoplasma gondii* infection in wild wood mice. *Infect Genet Evol.* **26**:180–184.

Nelson-Flower MJ, Germain RR, MacDougall-Shackleton EA, Taylor SS, Arcese P. 2018. Purifying selection in the toll-like receptors of song sparrows *Melospiza melodia*. *J Hered.* **109**: 501–509.

Nevoux M, Arlt D, Nicoll M, Jones C, Norris K. 2013. The short-and long-term fitness consequences of natal dispersal in a wild bird population. *Ecol Lett.* **16**:438–445.

Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, *et al.* 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3**:e170.

Novák K. 2014. Functional polymorphisms in toll-like receptor genes for innate immunity in farm animals. *Vet Immunol Immunopathol.* **157**:1–11.

Ortutay C, Siermala M, Vihinen M. 2007. Molecular characterization of the immune system: emergence of proteins, processes, and domains. *Immunogenetics* **59**:333–348.

Perrin A, Khimoun A, Faivre B, Ollivier A, de Pracontal N, Théron F, Loubon M, Leblond G, Duron O, Garnier S. 2021. Habitat fragmentation differentially shapes neutral and immune gene variation in a tropical bird species. *Heredity* **126**:148–162.

Piccinini AM, Midwood KS. 2010. DAMPening inflammation by modulating TLR signalling. *Mediat Inflamm.* **2010**:672395.

Podlaszczuk P, Indykiewicz P, Kamiński M, Minias P. 2021. Physiological condition reflects polymorphism at the toll-like receptors in a colonial waterbird. *Ornithology* **138**:ukab052.

Quéméré E, Galan M, Cosson J-F, Klein F, Aulagnier S, Gilot-Fromont E, Merlet J, Bonhomme M, Hewison AJM, Charbonnel N. 2015. Immunogenetic heterogeneity in a widespread ungulate: the European roe deer (*Capreolus capreolus*). *Mol Ecol.* **24**: 3873–3887.

Quéméré E, Hessenauer P, Galan M, Fernandez M, Merlet J, Chaval Y, Morellet N, Verheyden H, Gilot-Fromont E, Charbonnel N. 2021. Pathogen-mediated selection favours the maintenance of innate immunity gene polymorphism in a widespread wild ungulate. *J Evol Biol.* **34**:1156–1166.

Quéméré E, Rossi S, Petit E, Marchand P, Merlet J, Game Y, Galan M, Gilot-Fromont E. 2020. Genetic epidemiology of the Alpine ibex reservoir of persistent and virulent brucellosis outbreak. *Sci Rep.* **10**:4400.

Råberg L, Sim D, Read AF. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science* **318**:812–814.

Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J. 2020. Advances in the evolutionary understanding of MHC polymorphism. *Trends Genet.* **36**:298–311.

Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A. 2005. The evolution of vertebrate toll-like receptors. *Proc Natl Acad Sci U S A.* **102**:9577–9582.

Santos ES, Santos LLS, Lagisz M, Nakagawa S, Sheldon B. 2015. Conflict and cooperation over sex: the consequences of social and genetic polyandry for reproductive success in dunnocks. *J Anim Ecol.* **84**:1509–1519.

Schmid-Hempel P. 2011. *Evolutionary parasitology: the integrated study of infections, immunology, ecology, and genetics*. New York: Oxford University Press.

Sepil I, Moghadam HK, Huchard E, Sheldon BC. 2012. Characterization and 454 pyrosequencing of major histocompatibility complex class I genes in the great tit reveal complexity in a passerine system. *BMC Evol Biol.* **12**:68.

Shultz AJ, Sackton TB. 2019. Immune genes are hotspots of shared positive selection across birds and mammals. *Elife* **8**:e41815.

Spurgin LG, Richardson DS. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc R Soc B.* **277**:979–988.

Świderská Z, Šmídová A, Buchtová L, Bryjová A, Fabiánová A, Munclinger P, Vinkler M. 2018. Avian Toll-like receptor allelic diversity far exceeds human polymorphism: an insight from domestic chicken breeds. *Sci Rep.* **8**:17878.

Takahata N, Nei M. 1990. Allelic genealogy under over-dominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**:967–978.

Těšický M, Velová H, Novotný M, Kreisinger J, Beneš V, Vinkler M. 2020. Positive selection and convergent evolution shape molecular phenotypic traits of innate immunity receptors in tits (Paridae). *Mol Ecol.* **29**:3056–3070.

Tschirren B, Andersson M, Scherman K, Westerdahl H, Mittl PR, Råberg L. 2013. Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population. *Proc R Soc B.* **280**:20130364.

Tschirren B, Andersson M, Scherman K, Westerdahl H, Råberg L. 2012. Contrasting patterns of diversity and population differentiation at the innate immunity gene toll-like receptor 2 (TLR2) in two sympatric rodent species. *Evolution* **66**:720–731.

Tschirren B, Råberg L, Westerdahl H. 2011. Signatures of selection acting on the innate immunity gene Toll-like receptor 2 (TLR2) during the evolutionary history of rodents. *J Evol Biol.* **24**:1232–1240.

Vanpé C, Debeffe L, Galan M, Hewison AM, Gaillard JM, Gilot-Fromont E, Morellet N, Verheyden H, Cosson JF, Cargnelutti B, *et al.* 2016. Immune gene variability influences roe deer natal dispersal. *Oikos* **125**:1790–1801.

Velová H, Gutowska-Ding MW, Burt DW, Vinkler M. 2018. Toll-like receptor evolution in birds: gene duplication, pseudogenization, and diversifying selection. *Mol Biol Evol.* **35**:2170–2184.

Verweij SP, Karimi O, Pleijster J, Lyons JM, de Vries HJ, Land JA, Morré SA, Ouburg S. 2016. TLR2. TLR 4, and TLR9 genotypes and haplotypes in the susceptibility to and clinical course of Chlamydia trachomatis infections in Dutch women. *FEMS Pathog Dis.* **74**:ftv107.

Vinkler M, Adelman JS, Ardia DR. 2022. Evolutionary and ecological immunology. In: Kaspers B, Schat KA, Göbel TW, Vervelde L (eds) *Avian immunology*. 3rd ed. Academic Press, London, pp 519–557.

Vinkler M, Albrecht T. 2009. The question waiting to be asked: Innate immunity receptors in the perspective of zoological research. *Folia Zool.* **58**:15–28.

Vinkler M, Bainová H, Bryja J. 2014. Protein evolution of toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genet Sel Evol.* **46**:72.

Vinkler M, Bainová H, Bryjová A, Tomášek O, Albrecht T, Bryja J. 2015. Characterisation of Toll-like receptors 4, 5, and 7 and their genetic variation in the grey partridge. *Genetica* **143**:101–112.

Walsh C, Gangloff M, Monie T, Smyth T, Wei B, McKinley TJ, Maskell D, Gay N, Bryant C. 2008. Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J Immunol.* **181**:1245–1254.

Wang X, Quinn PJ, Yan A. 2015. Kdo 2-lipid A: structural diversity and impact on immunopharmacology. *Biol Rev.* **90**:408–427.

Webb AE, Gerek ZN, Morgan CC, Walsh TA, Loscher CE, Edwards SV, O'Connell MJ. 2015. Adaptive evolution as a predictor of species-specific innate immune response. *Mol Biol Evol.* **32**:1717–1729.

Wlasiuk G, Khan S, Switzer WM, Nachman MW. 2009. A history of recurrent positive selection at the toll-like receptor 5 in primates. *Mol Biol Evol.* **26**:937–949.

Woolhouse MEJ, Webster JP, Domingo E, Charlesworth B, Levin BR. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet.* **32**:569–577.

Xu W, Zhou X, Fang W, Chen X. 2020. Genetic diversity of toll-like receptor genes in the vulnerable Chinese egret (*Egretta eulophotes*). *PLoS One.* **15**:e0233714.

Zdorovenko GM, Zdorovenko EL, Varbanets LD. 2007. Composition, structure, and biological properties of lipopolysaccharides from different strains of *Pseudomonas syringae pv. atrofaciens*. *Microbiology* **76**:683–697.