Veterinary Medicine and Science

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Equine

Innate Immunity Toll-Like Triad TLR6-1-10 and Its Diversity in Distinct Horse Breeds

Karla Stejskalova[1] | Leona Vychodilova[1] | Eva Janova[1,2] | Jan Oppelt[2] | Petr Horin[1,2]

[1]Department of Animal Genetics, Faculty of Veterinary Medicine, University of Veterinary Sciences Brno, Brno, Czech Republic | [2]CEITEC VETUNI, RG Animal Immunogenomics, University of Veterinary Sciences Brno, Brno, Czech Republic

Correspondence: Petr Horin (horin@ics.muni.cz)

Received: 29 July 2024 | **Revised:** 22 November 2024 | **Accepted:** 17 January 2025

Funding: The project was funded by internal grant of VETUNI no. 2021ITA12.

Keywords: equine | haplotype | innate immunity | toll-like receptor

ABSTRACT

Toll-like receptors (TLRs) play important roles in innate immunity and developmental processes. Due to their nature as molecular pattern recognition receptors, their genetic diversity may reflect the effects of various pathogen pressures. Here, the extent of variability in the *TLR1-6-10* gene cluster in three geographically and historically distinct breeds of horses was analysed. A genetically diverse group of representatives of 14 other horse breeds provided additional information on the variability of this gene cluster in the domestic horse. Altogether, 25 SNPs were identified in the *TLR6-1-10* gene cluster across the 4 equine breed groups studied, of which 7 were synonymous and 18 non-synonymous. Twenty-eight inferred SNPs and 22 in silico translated amino acid haplotypes were identified. A predominant major haplotype present in all breed groups along with several group-specific haplotypes were identified. Strong linkage disequilibrium was detected for several SNPs, as well as effects of pervasive, site-specific selection. The existence of a major haplotype suggests it may confer a selective advantage across breeds. Less frequent breed-specific haplotypes may represent variability required or beneficial for responses to local pathogen pressures. Purifying site-specific selection was detected in the TIR domain and its vicinity in TLR6, whereas AA sites under diversifying selection were located in LRR domains and/or their surroundings in TLR1. Population structure models based on immune-related *TLR6-1-10* markers did not distinguish between breed groups, whereas in models based on neutral microsatellite markers, breed groups clustered separately. This supports the assumption that the diversity of the *TLR6-1-10* cluster is of adaptive value. The *TLR6-1-10* alleles and haplotypes identified represent potential candidate markers for disease association studies.

1 | Introduction

Toll-like receptors (TLRs) are evolutionarily old innate immunity molecules responsible for the recognition of various conserved molecular patterns associated with pathogens or released by damaged cells. Typically, N-terminal leucin-rich repeat domains (LRRs) recognize the ligand, whereas a conserved toll/interleukin-1 receptor (TIR) domain initiates downstream signalling (Behzadi et al. 2021). TLR signalling pathways are able to elicit the production of inflammatory cytokines and Type-I interferon (Kawai and Akira 2011), as well as adaptive immunity responses (Kumar 2022). TLRs are localized on the plasma or endosomal membranes of various immune, epithelial and endothelial cells (Vijay 2018).

From insects to mammals, TLRs play important roles in innate immunity and developmental processes (Imler and Zheng 2004; Kawai and Akira 2011; Lindsay and Wasserman 2014).

Ancestral TLR prototypes were identified in early Cnidarians and Molluscs (Brennan and Gilmore 2018). To date, 28 TLRs have been identified in vertebrates (of which TLR1-13 are found in mammals). These are divided into eight gene families (Liu et al. 2020). Conservation of synteny has been observed for *TLR* gene families (Roach et al. 2005). The *TLR1*, *6* and *10* gene cluster evolved through duplications and gene conversion (Kruithof et al. 2007); its genomic organization is similar across vertebrate species. In cattle, the *TLR1*, *TLR6* and *TLR10* genes are known to occupy an approximately 50 kb region of bovine chromosome 6, following the same gene order as in humans on HSA4, which is *5′-TLR6-1-10-3′* (Opsal et al. 2006).

The role and importance of TLRs in immunity are well documented. TLR1 and TLR6, in the form of dimers with TLR2, are potent sensors of bacterial infection. TLR1 ligands are bacterial lipopeptides and TLR6 ligands are bacterial lipoproteins, lipoteichoic acid and zymosan (Behzadi et al. 2021). The function of TLR10 is less clear, as it can bind lipoproteins and viral glycoproteins as well as dsRNA, but it is also involved in the regulation of immune responses (Behzadi et al. 2021; Rodrigues et al. 2024).

Multiple single-nucleotide polymorphisms (SNPs) in TLR genes have been associated with resistance or susceptibility to infectious or immunity-related diseases in human and animal populations (Mukherjee et al. 2019). Pulmonary tuberculosis is associated with the polymorphic site H305L in TLR1 in humans (Meyer et al. 2016), whereas the I602S amino acid (AA) change in that receptor is associated with leprosy (Johnson et al. 2007). Variations in the TLR10 gene are known to be associated with the risk of asthma (Lazarus et al. 2004). Due to linkage disequilibrium (LD), not only solitary SNPs but also haplotypes across the whole *TLR6-1-10* gene cluster have been associated with diseases, for example, with tuberculosis in the African–American population (Ma et al. 2007) or prostate cancer risk (Stevens et al. 2008). Furthermore, there is evidence that several archaic haplotypes, originating from Denisovans and Neanderthals and contributing to increased microbial resistance, have been introduced by recurrent adaptive introgression to the modern human gene pool (Dannemann et al. 2016).

Equids possess 12 expressed TLR genes. The genomic organization and phylogenetic relationships of *TLR1-12* have been studied in the order Perissodactyla with a focus on the family Equidae (Stejskalova, Janova, Splichalova et al. 2023). However, little is known about the diversity of *TLR* genes within the *Equus caballus* species itself. In the current horse reference genome EquCab 3.0, all three *TLR* genes are localized on chromosome ECA3 in a region spanning 84 kbp and oriented in the same (forward) direction. *TLR6* is the first of the triad, followed by *TLR1* and finally the *TLR10* gene. The coding sequences (CDS) of *TLR6-1-10*, each comprising a single exon, are separated from each other by 33 and 18 kbp, respectively. The expression of *TLR6*, *1* and *10* has been confirmed and mRNA has been identified in equine white blood cells (Uddin et al. 2016). As TLR genes are immunity-related genes, their diversity and evolution reflect at least in part the history of host species–pathogen interactions (Bagheri and Zahmatkesh 2018).

The aim of this study was to evaluate the diversity and signatures of selection in the CDS of the *TLR6-1-10* gene cluster in distinct horse breeds adapted to different environments and to compare it with the diversity found in neutral microsatellite markers. The hypothesis was that neutral markers, as they are subjected only to low or no selective pressures, should characterize differences and/or similarities between breeds, whereas diversity in immunity-related TLR genes should reflect the adaptive value of these variants for the respective populations. The presence of breed-specific variants would suggest a history of interaction with environment-specific pathogens; the existence of shared variants and haplotypes would, on the contrary, suggest the presence of an important shared selective pressure and a general advantage provided by a particular haplotype.

The historic and centuries-old autochthonous Czech breed of Old Kladruber carriage horses, the ancient Akhal-Teke breed originating on the Turkmenistan steppes and representatives of the current European population of Arabian horses were analysed for this purpose. A mixed group of representatives of 14 other horse breeds was used as a source of additional diversity of the equine *TLR6*, *1* and *10* genes.

2 | Methods and Materials

2.1 | Blood Samples and DNA Extraction

Archived blood samples from 131 domestic horses, originally collected for the purposes of previous projects (Stejskalova, Cvanova et al. 2019; Vychodilova et al. 2013), were analysed. Arabian horses ($n = 31$), Old Kladruber grey horses ($n = 32$) and Akhal-Teke horses ($n = 30$) were studied as representatives of three different horse breeds. In addition, a panel of 2–3 samples from each of 14 other horse breeds was composed as follows ($n = 41$): Appaloosa, Camargue, Czech Warmblood, Espagnol, Icelandic horse, Lusitano, Murgese, Old Kladruber black, Quarter Horse, Romanian, Selle Francaise, Thoroughbred and Welsh pony. All blood samples were stored at −20°C. DNA extraction from 200 µL of whole blood was performed using a NucleoSpin Blood kit (Macherey-Nagel, Düren, Germany).

2.2 | Sequencing and Sequence Analyses

Primers flanking exons bearing coding regions of *TLR1*, *6* and *10* were designed and cycling conditions set as described previously (Stejskalova, Janova, Splichalova et al. 2023). PCR reactions were performed in volume of 12.5 µL using EliZyme HS Robust MIX Red (Elisabeth Pharmacon, Brno, Czech Republic). NGS resequencing on a MiSeqTM System (Illumina, San Diego, California, USA) was performed as described previously (Stejskalova, Janova, Splichalova et al. 2023). Consensus sequences were generated from the final short-read alignments for each TLR and each sample and trimmed to contain only CDS. Variable positions were verified by SAMtools and GATK software. To solve ambiguities in heterozygotes, haplotypes were inferred using the PHASE algorithm in DnaSP (v6.12.03) and assigned to samples. Genotypes in eight neutral microsatellites used in the standard equine parentage test (HMS7, HTG4, VLH20, ASB23, ASB17, HMS2, HTG10 and AHT4) were assessed using standard multiplex PCR (Lee and Cho 2006).

2.3 | Population Structure Analyses

Allelic and haplotype frequencies were determined by direct counting using MS Excel. Observed and expected heterozygosities, the genetic differentiation between breeds (using the FST) and analysis of molecular variance (AMOVA) were assessed using GenAIEx v6.51 (Peakall and Smouse 2012). Mean observed and expected heterozygosities were compared by the Mann–Whitney test. To infer the population structure, principal coordinate analysis (PCoA-using GenAlEx) and the model-based Bayesian method in STRUCTURE (v2.3.4) were used (Pritchard et al. 2000). An admixture model with correlated allele frequencies was adopted for this purpose. Each parameter set was analysed with five replicates for $K = 1$ to $K = 6$, and all runs were performed with 10,000 burn-in periods and 50,000 MCMC repeats after burn-in. Clustering Markov Packager Across K (CLUMPAK) was used for the summation and graphical illustration of the results obtained by STRUCTURE (Kopelman et al. 2015).

2.4 | Selection Analyses

Effects of overall selection acting upon whole genes were assessed using the codon-based *Z*-test of selection in MEGA (v11.0.13). Hypotheses of strict neutrality, diversifying and purifying selection were tested (Kumar method, 1000 boot-strap replicates). Effects of pervasive site-specific selection were evaluated by FEL, FUBAR and SLAC methods; episodic site-specific diversifying selection was tested by MEME (all performed by Datamonkey.org; Weaver et al. 2018). LD was tested in DnaSP; significant pairwise comparisons were evaluated by Fisher's exact test with the Bonferroni correction.

3 | Results

3.1 | Sequence Polymorphisms

Full-length CDS of *TLR6*, *TLR1* and *TLR10* were obtained, and CDS, protein lengths and variable sites were identified for the four breed groups studied. Altogether, 25 SNPs were identified in the *TLR6-1-10* CDS. A brief list of all SNPs detected is presented in Table 1. A full list including the allelic frequencies as well as SNP counts in breed groups is provided in Tables S1 and S2. Eighteen SNPs were non-synonymous, five each were located in *TLR6* and *TLR10* and eight in *TLR1*. Three SNPs were specific to the Arabian breed, one to the Akhal-Teke breed and three SNPs were detected in the 14 breed panel only. All SNPs except one have already been described in the European Variation Archive (EVA). TLR1 showed the highest AA variation (1.02%) compared to TLR6 and TLR1 (0.63% and 0.62%, respectively).

All 25 SNPs were assembled into 28 inferred nucleotide haplotypes spanning the coding regions of the entire *TLR6-1-10* cluster. In silico translation into combined TLR6-1-10 AA sequence resulted in 22 AA haplotypes (Figure 1). In Arabian horses, seven of seven (nucleotide/AA) haplotypes were identified, out of which one of one was breed specific. In the Old Kladruber breed, there were eight of eight haplotypes, one of one of which was breed specific. In the Akhal-Teke breed group, 14/12 haplotypes were found, out of which 4/3 were breed specific. As expected, the highest variation was documented in the 14 breed panel with 21/17 haplotypes, 8/5 of which were panel specific. The most common haplotype was shared across all breed groups; it also showed the highest frequency within each breed (Figure 1). A full list of haplotypes including their frequencies in the breed groups is provided in Tables S3 and S4.

3.2 | Selection Analyses

The overall codon-based *Z*-test of selection showed that TLR1 was under positive selection ($p = 0.04$) when all horses were evaluated as a single group, as well as for the Arabian horses ($p = 0.03$). No significant results were observed in the other three breed groups. No signs of overall selection acting upon TLR6 or TLR10 were detected.

Site-specific pervasive purifying selection was detected in TLR6 (Table 2); AA site 536, located in the LRR-C-terminal domain, and AA site 715 in the TIR domain confirmed AA sites under the effects of negative selection previously identified in the Equidae (Stejskalova, Janova, Splichalova et al. 2023).

Site-specific positive selection was detected in TLR1 only; in each breed group, there was at least one positively selected AA site. No positively selected AA sites were identified in TLR6 or TLR10.

3.3 | Population Structure Analyses

LD was detected across all three *TLRs*. Significant values of both abs(D′) and R2 in each SNP combination are shown in Figure 2. Consistently strong LD was detected for non-synonymous SNPs, affecting AA sites I452V in an LRR domain of TLR6, H290R next to LRR domains of TLR1 and sites T117A and V492I in LRR domains of TLR10, along with synonymous SNPs coding for AA site E536 located in a LRR-CT domain and site E715 in the TIR domain of TLR6.

No significant differences between observed (Ho) and expected (He) mean heterozygosity (Table 3) in *TLR6-1-10* SNPs were found in any of the Arabian, Akhal-Teke or Old Kladruber horses breeds. When compared between breeds, no significant differences in He and/or Ho were found. Ho, He and other characteristics for each SNP in each breed are in Tables S2 and S5.

For the neutral microsatellite markers analysed, there was a significant difference between the mean Ho and He in Arabian horses as Ho in this breed was 100%. The difference in Ho between Arabians and the other two breeds was significant ($p < 0.01$), but no significant difference between breeds in He was observed (Table 3; Table S6).

The population structure analyses based on the two sets of markers—*TLR6-1-10* SNPs in functional genes and neutral microsatellite genotypes—showed differences between them. PCoA and structure analysis of *TLR6-1-10* markers in the three breeds revealed no breed-specific clustering, which would allow

TABLE 1 | An overview of single-nucleotide polymorphisms identified in coding sequences of TLR6, TLR1 and TLR10 in all four horse breed groups tested. Genomic positions shown refer to the EquCab3.0 assembly. Where available, the SNP identifier for the European Variation Archive is provided. Corresponding amino acid sites and their positions in functional protein domains are shown.

Gene	Position in the gene	ECA3 chromosome position	SNP	EVA RefSNP	AA site	Domain
TLR6 GeneID:100064554 2391 bp 796 AA	g.25911	90509655	C>T	rs1139032502	S374	
	g.26143	90509887	A>G	rs1140698326	I452V	LRR
	g.26246	90509990	C>T	rs395887193	A486V	LRR
	g.26370	90510114	G>C	rs396363017	G527	
	g.26397	90510141	G>A	rs1145170582	E536	LRR-CT
	g.26741	90510485	G>A	rs1149930392	R651H	TIR
	g.26934	90510678	A>G	rs1135913429	E715	TIR
	g.27104	90510848	T>A		L772H	TIR
	g.27164	90510908	A>G	rs1148651563	N792S	
TLR1 GeneID:100049577 2361 bp 786 AA	g.2674	90544177	C>A	rs1150405995	H115N	LRR
	g.2934	90544437	G>A	rs3432737291	P201	
	g.3013	90544516	G>T	rs1142778808	V228L	
	g.3038	90544541	A>G	rs1146945368	Y236C	
	g.3200	90544703	G>A	rs68512466	R290H	
	g.3682	90545185	A>G	rs1140498191	I451V	LRR
	g.3816	90545319	C>G	rs1143352175	S495R	LRR
	g.4048	90545551	G>A	rs1138594416	D573N	LRR-CT
	g.4543	90546046	A>G	rs1141161915	I738V	TIR
TLR10 GeneID:100064614 2436 bp 811 AA	g.23505	90565017	A>G	rs394415355	T117A	LRR
	g.24167	90565679	A>T	rs1145280186	Q337H	
	g.24630	90566142	G>A	rs1148006722	V492I	LRR
	g.24846	90566358	T>C	rs1141128290	L564	LRR-CT
	g.24963	90566475	C>T	rs1146892101	L603	
	g.25065	90566577	T>C	rs1147800666	F637L	TIR
	g.25267	90566779	A>G	rs1143116745	E704G	TIR

Abbreviations: LRR, leucine-rich repeat domain; LRR-CT, C-terminal leucine-rich repeat domain; TIR, toll/interleukin-1 receptor domain.

one to distinguish between breeds (Figure 3a). Most of the molecular variance was observed within breed groups, while the differences between breeds were small (9%). FST values showed significant, low to intermediate (0.068–0.172) genetic distance between breeds with the exception of FST values for TLR10 between Akhal-Teke and Arabian horses, which were non-significant (Table 4). The most distant from each other were Arabians and Old Kladruber horses.

When neutral microsatellites were analysed, both PCoA and structure analyses showed that the three breeds were separated (Figure 3b). Arabians were clearly distinct, whereas for Akhal-Teke and Old Kladruber horses, some clusters indicated mixed contributions from inferred ancestral populations. Most of the molecular variance was observed within breeds (88%). For inter-breed distances, the highest FST value (0.178) calculated was between Arabian and Akhal-Teke horses (Table 4).

4 | Discussion

The origin, occurrence, course and spread of infectious diseases may be understood as a manifestation of evolutionarily fixed properties of both pathogens and their hosts, resulting in complex host–pathogen interactions. Innate immunity is the phylogenetically older component of anti-infective defence (Bryant and Monie 2012). The mechanisms of innate immunity focus on the detection of evolutionarily conserved determinants common to larger groups of pathogens through pattern recognition receptors, such as the TLRs (Gasteiger et al. 2017). Although mammalian species differ in the number of *TLR* genes present in their genome and/or in the number of TLRs expressed, the *TLR* gene families show considerable conservation synteny (Roach et al. 2005), for example, *TLR7* and *TLR8*, located on the X chromosome, and the TLR6-1-10 gene cluster, whose genomic organization follows the same pattern in humans, mice, cattle and horses (Opsal et al. 2006; Rodrigues et al. 2024). In an evolutionary

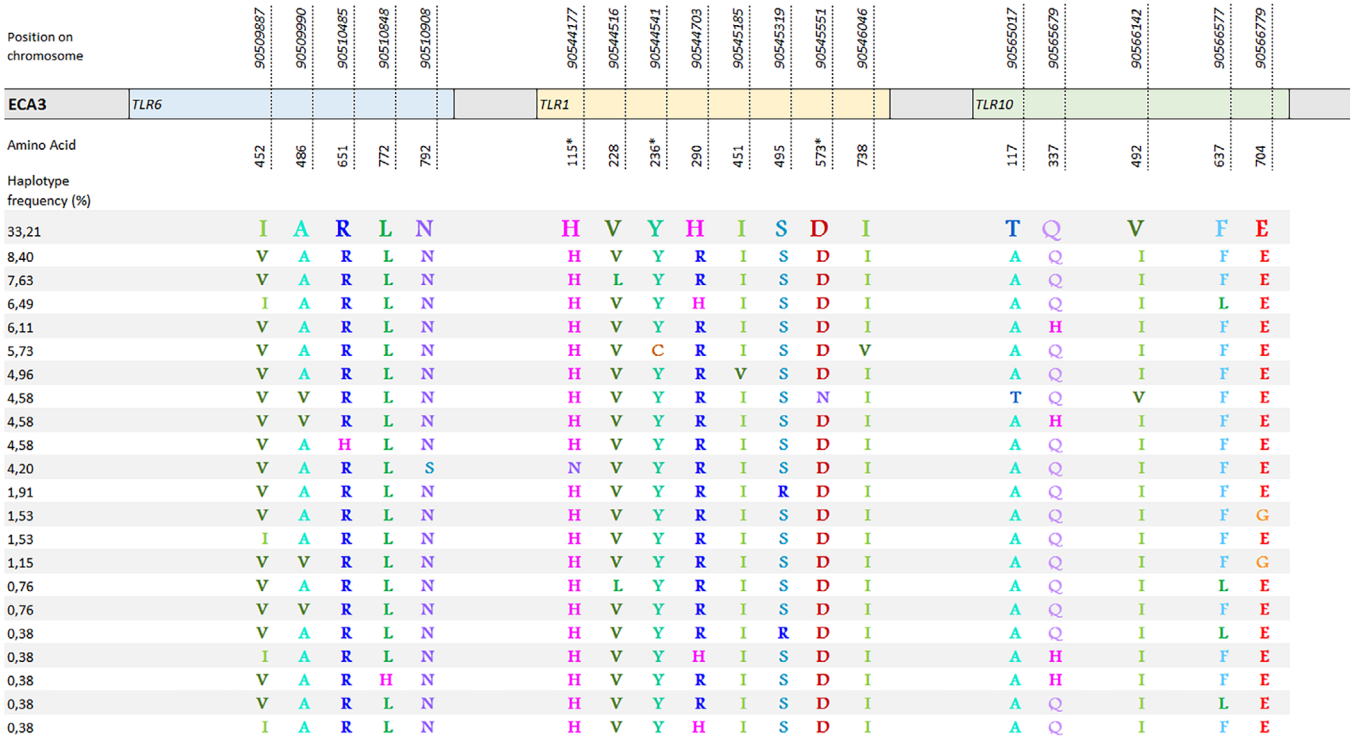

FIGURE 1 | Amino acid haplotypes translated from inferred SNP haplotypes identified in the TLR6-1-10 gene cluster in all four breed groups studied. Overall haplotype frequencies are shown; the major haplotype was also the most frequent one in each breed group when evaluated separately. Genomic positions shown refer to the assembly EquCab3.0. * denotes positively selected amino acid site.

TABLE 2 | An overview of amino acid (AA) sites under the effect of pervasive selection forces in the four horse breed groups studied. AA site localization in functionally important protein domains is shown. LRR (leucine-rich repeat domain) or LRR-CT (C-terminal leucine-rich repeat domain) in bold and TIR (toll/interleukin-1 receptor domain) in italics.

	Arabian	Akhal-Teke	Old Kladruber	14 Breeds panel	Across all breeds
AA sites under pervasive purifying selection in TLR6				527, 536	527, 536, *715*
AA sites under pervasive diversifying selection in TLR1	236, 290, 451, 573	115	290, 451, 573	115, 573	115, 236, 573

context, the *TLR6-1-10* gene cluster on human chromosome 4 carries archaic haplotypes of functional significance acquired by adaptive introgression from Neanderthals and Denisovans, associated with increased resistance to bacterial infections but also with increased susceptibility to allergic diseases (Dannemann et al. 2016). Other *TLR6-1-10* haplotypes as well as solitary SNPs have been associated with susceptibility to infections (Mukherjee et al. 2019) or sarcoidosis (Veltkamp et al. 2012). Based on the conservative nature of TLRs across mammals, we may assume that the TLR6-1-10 cluster is subject to the selective pressures of pathogens also in horses. However, there is little information on *TLR* gene variation in horses and its adaptive value. In a study of the kinetics of West Nile Virus (WNV)–induced transcripts in PBMCs and other tissues, multiple TLRs including TLR1 and TLR6 were found to be upregulated in infected horses (Bielefeldt-Ohmann et al. 2017; Uddin et al. 2016). Outside the TLR6-1-10 cluster, TLR4 in compound genotypes was associated with WNV outcome in horses as well (Stejskalova, Cuanova et al. 2019; Stejskalova, Janova, Horecky et al. 2019).

In this study, we assessed the extent of the variability in the *TLR1-6-10* gene cluster and compared it to the variability in a set of neutral microsatellites in three geographically and culturally distinct breeds of horses. The Akhal-Teke is an ancient primitive breed native to arid areas of Central Asia (Cieslak et al. 2010), and its main breeding centres have remained until now in Turkmenistan. The Arabian horse was domesticated in the deserts of the Arabian Peninsula, but over the centuries many lineages have arisen in Egypt and Europe, and mtDNA sequence analysis shows that modern Arabian horses may have heterogeneous origins (Głażewska 2010). The Kladruber horse was developed as a heavy carriage horse 400 years ago in mild weather of Central Europe. Since its establishment, strict selective breeding has been applied. Despite intensive inbreeding, several

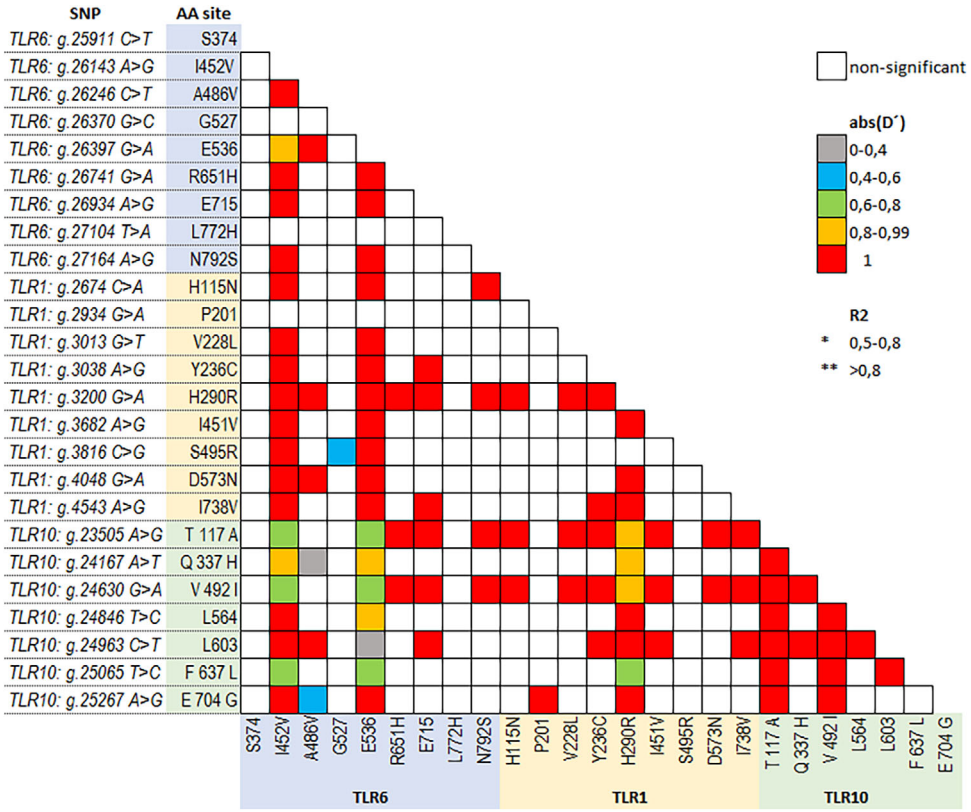

FIGURE 2 | A graphic overview of linkage disequilibrium detected between 25 SNPs in all four horse breed groups. Only significant values are shown ($p < 0.01$; number of significant pairwise comparisons by Fisher's exact test: 102, number of significant comparisons using the Bonferroni procedure: 54).

TABLE 3 | Mean heterozygosity observed (Ho) and expected (He) in two sets of markers—*TLR6-1-10* SNPs and neutral microsatellites—in three solid-breed groups of horses.

	Arabian	Akhal-Teke	Old Kladruber
Mean Ho (SE) in TLR6-1-10 SNPs	0.261 (0.059)	0.192 (0.040)	0.189 (0.046)
Mean He (SE) in TLR6-1-10 SNPs	0.209 (0.045)	0.182 (0.039)	0.179 (0.042)
Mean Ho (SE) in eight neutral microsatellites	1	0.697 (0.081)	0.787 (0.025)
Mean He (SE) in eight neutral microsatellites	0.774 (0.014)	0.688 (0.074)	0.731 (0.016)

Abbreviation: SE, standard error.

bottlenecks and a small population size and genetic variability within the breed remain on a comparable level with other breeds (Janova et al. 2013). To encompass to the maximum extent of the diversity in these genes, a group of representatives of 14 other horse breeds was studied as well. The origins of these breeds cover geographically and ecologically distinct biotopes, such as the harsh Icelandic and Welsh highlands and the warm Iberian Peninsula, thus, increasing the expected variability in both neutral- and immune-related markers.

Our previous studies focused on the history of the Old Kladruber horse showed that the black and grey varieties of this breed are two groups with different molecular variability (Horín et al. 1998; Janova et al. 2013). Therefore, the black variety of the Old Kladruber horse was considered as an independent group and was included in the diverse breed group.

Altogether, 25 SNPs were identified across the *TLR6-1-10* gene cluster in the 4 equine breed groups studied; out of these, 7 were synonymous and 18 non-synonymous (Table 1). With a single exception (*TLR6*: g.27104 T>A), they have all been recorded already in the EVA. Breed-specific SNPs were found in *TLR6* and *TLR1* in each of the groups except the Old Kladruber group. The polymorphism rates observed ranged between 0.29 SNP per 100 bp in *TLR10* and 0.38 SNP per 100 bp in *TLR6* and *TLR1*. In *TLR6*, there were 9 SNPs, of which 5 were non-synonymous; in *TLR1*, there were 9 SNPs with just one synonymous; and in *TLR10*, there were 5 non-synonymous out of 10 total SNPs (Tables S1 and S2). Such a level of polymorphism is rather low compared to observations in domestic pigs (Bergman, Edman et al. 2012; Bergman, Rosengren et al. 2010), where altogether 55 SNPs were identified in the *TLR6-1-10* gene cluster. In porcine *TLR6*, there were 9/9 (syn/non-syn) SNPs, 13/7 in *TLR1*, and the

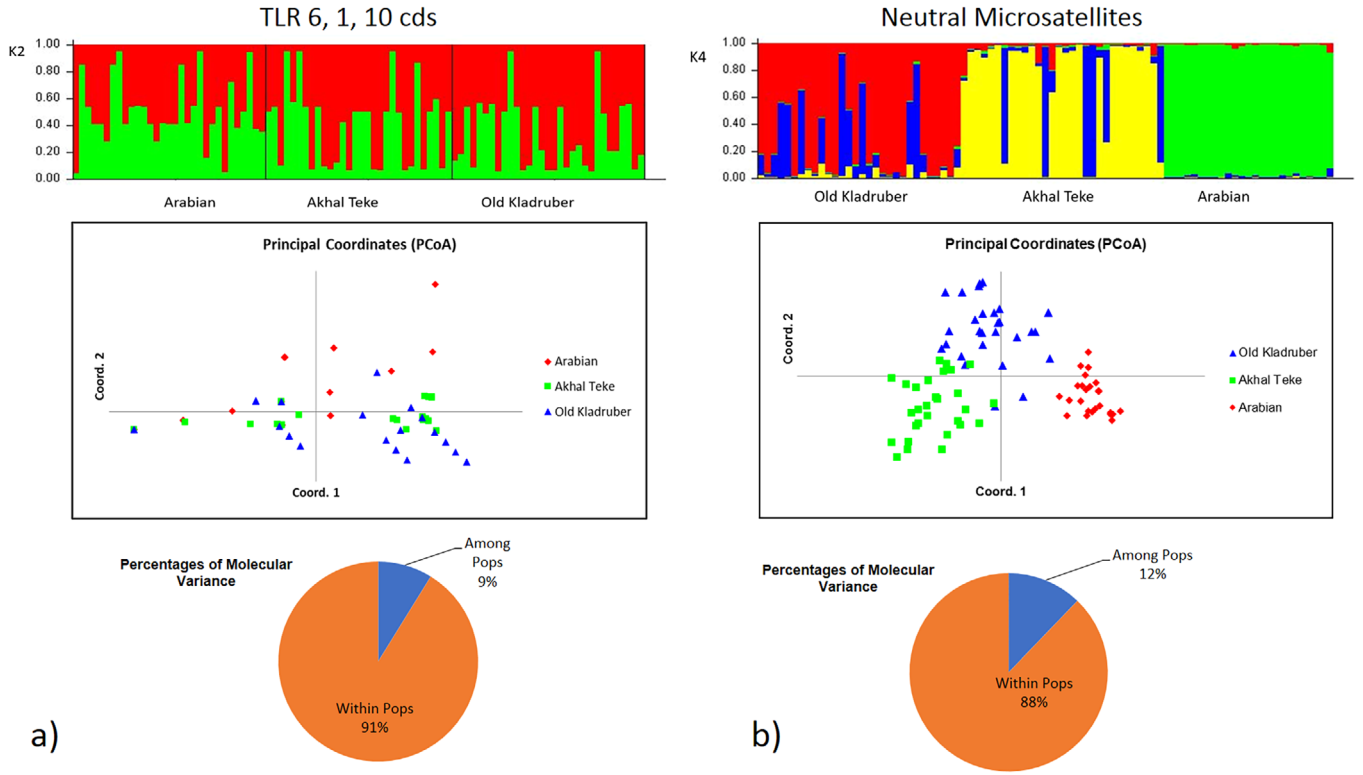

FIGURE 3 | Principal coordinates (PCoA), percentages of molecular variance and STRUCTURE analyses of population structure in three solid-breed horse groups. (a) Twenty-five immune-related *TLR6-1-10* markers. (b) Eight neutral microsatellite markers. While most of the molecular variance remains within breed groups (populations) in both sets of markers, there is no breed-specific clustering recognized in *TLR6-1-10* marker set.

TABLE 4 | An estimation of genetic distance between horse breed groups based on FST values. Immune-related markers (grouped as the TLR6-1-10 cluster or as separate genes) and neutral microsatellites were evaluated. Significant values are shown ($p < 0.01$).

		FST ($p < 0.01$)				
Breed Group 1	**Breed Group 2**	**TLR6-1-10**	**TLR6**	**TLR1**	**TLR10**	**Microsatellites**
Old Kladruber	Akhal-Teke	0.068	0.053	0.068	0.076	0.097
Old Kladruber	Arabian	0.134	0.139	0.172	0.097	0.121
Akhal-Teke	Arabian	0.061	0.072	0.099	—	0.178

most extensive polymorphism was detected in *TLR10* with 13/20 SNPs. Studies in cattle reported 5 non-synonymous SNPs in *TLR1* and 12 in *TLR10* (Seabury et al. 2007). Small group size is a possible but unlikely explanation for decreased SNP variability observed in horses, as our study involved 131 individuals representing 17 breeds. In contrast to the high SNP counts, only 13 *TLR6-1-10* SNP haplotypes were derived for domestic pigs (Bergman, Edman et al. 2012), compared to the 28 haplotypes that we inferred in horses. The authors also reported several haplotypes determined in wild boars and noted a group- and environment-driven specificity of most of the inferred haplotypes. The higher number of haplotypes observed in the horse cohort can be attributed to its genetic heterogeneity in terms of breed representation and their very different origins. Nevertheless, we identified one markedly predominant major haplotype present in all breed groups (Figure 1; Tables S3 and S4), along with several group-specific haplotypes. The presence of the major haplotype suggests it confers a possible selective advantage across the breed groups, whereas the less frequent haplotypes may represent variability required for responses to local pathogen pressures.

Effects of overall selection were identified only for one (*TLR1*) of the three genes. Considering the structure of TLR proteins and the different functions of their domains, it is not surprising that effects of purifying and/or diversifying selection were identified in codons, encoding these domains with specific functions rather than at the level of entire genes. While signalling pathways and relevant TIR domains are highly conserved both between TLR types and between species, the need to recognize different pathogens can result in adaptive evolutionary changes in the LRR region. Here, we noticed that the synonymous SNP *TLR6*: g.26397 g/a at AA site 536, located in the LRR-C terminal domain

and acting under the effect of purifying selection, was in strong LD with all non-synonymous *TLR1* SNPs and most of the other SNPs detected in *TLR6* and *TLR10* (Figure 2). Similar LD was also observed for *TLR6*: g.26143 a/g (I452V), but no selection forces were detected at this site. Five sites were found to be under the effect of diversifying selection in *TLR1*; of these, AA site H290R (*TLR1*: g.3200 g/a) showed medium to strong LD with a majority of the tested SNPs (Figure 2). LD in the *TLR6-1-10* cluster has been described before, supporting the observation that not only independent SNPs but also across-block haplotypes are associated with specific phenotypes (Dannemann et al. 2016; Stevens et al. 2008). AAs at sites under diversifying selection in TLR1 in the major haplotype are also common in other haplotypes, whereas at other sites (*TLR1*: 290 and *TLR10*:117, 492) the AAs in the major haplotype are rather uncommon (Figure 1).

Phylogenetic analyses have revealed effects of positive selection in *TLR* genes in both humans and other mammals (Areal et al. 2011), including in family Suidae (Darfour-Oduro et al. 2015) and in Artiodactyla (Ishengoma and Agaba 2017). In our previous study (Stejskalova, Janova, Splichalova et al. 2023), we identified two AA sites under positive selection in the *TLR6-1-10* cluster in the entire family Equidae (TLR1: I678V and TLR10: T785A,M), but neither of them was found to be under selection in any of the four horse breed groups studied here. None of the five positively selected sites we identified in TLR1 matched those described by Liu et al. (2020) in vertebrates as a whole. The differences between these findings may at least partly be due to sampling bias and limited numbers of horses/equids available for the analyses.

Population structure analyses using *TLR6-1-10* genotypes did not distinguish between breeds, as no breed-specific clusters were detected (Figure 3a). This is not surprising considering that one major haplotype was identified across the entire cohort, as well as within each of the breed groups studied. Pathogen pressure tends to reduce the population structure (Schierup et al. 2000). Our data suggest that all breed groups analysed may be confronted with common selective pressures acting on immune responses.

This assumption is supported by the finding that analyses of neutral microsatellite loci showed the three breed groups to be genetically separated (Figure 3b), even though there were hints of mixed contributions from ancestral populations in some clusters. Microsatellites interspersed throughout the genome provide a tool for population structure analyses and have been used for breed characterization previously (Bjørnstad and Røed 2001; Parker et al. 2004; Rahal et al. 2021). Although we have genotyped only a limited number of loci covering eight equine chromosomes, the cohort is more structured than for the *TLR* loci (Figure 3a,b). The biallelic nature of *TLR6-1-10* SNP markers may reduce the information content and limit the resolution of the detection of population structure based on them. However, the existence of a major haplotype common to all breed groups and the comparison of frequencies of other haplotypes illustrate that real inter-breed differences are limited. This interpretation of our data is also supported by a study comparing microsatellites and SNPs in red deer, which found correlations between estimates of parameters of genetic diversity and genetic differentiation for both types of markers, with microsatellites showing lower accuracy in representing the distribution of genetic diversity among individuals (Pérez-González et al. 2023).

5 | Conclusion

The extent of variability in the *TLR6-1-10* gene cluster in three geographically and historically distinct breeds of horses and in a genetically diverse group of representatives of 14 other horse breeds was analysed. Altogether, 25 SNPs (72% non-synonymous) were identified in the *TLR6-1-10* gene cluster and were assembled into 28 inferred SNPs and 22 translated AA haplotypes. A predominant major haplotype was present in all breed groups, and several group-specific haplotypes were identified. The existence of a major haplotype suggests that this haplotype may present a selective advantage across breeds, whereas less frequent breed-specific haplotypes may represent variability required for responses to local pathogen pressures. Traces of both purifying and diversifying selection were found. Population structure based on the immune-related *TLR6-1-10* SNP markers did not distinguish between the breed groups, whereas in the structure based on neutral microsatellite markers, breed groups clustered separately. The observed diversity of the *TLR6-1-10* cluster may therefore have an adaptive value. The *TLR6-1-10* alleles and haplotypes identified represent suitable candidate markers for disease association studies.

Author Contributions

Karla Stejskalova: investigation, data curation, visualization, writing–original draft. **Leona Vychodilova**: formal analysis, writing–review and editing. **Eva Janova**: formal analysis, writing–review and editing. **Jan Oppelt**: software, data curation, resources. **Petr Horin**: conceptualization, supervision, validation, writing–review and editing.

Ethics Statement

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to. No ethical approval was required as all blood samples used in this study were archived samples originally collected for the purposes of other projects (Futas and Horin 2013; Bayerova et al. 2016; Klumplerova et al. 2020). Based on the legislation at the time, all samples were collected by licensed veterinarians in agreement with all ethical, welfare and professional standards.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Areal, H., J. Abrantes, and P. J. Esteves. 2011. "Signatures of Positive Selection in Toll-Like Receptor (TLR) Genes in Mammals." *BMC Evolutionary Biology* 11: 368. https://doi.org/10.1186/1471-2148-11-368.

Bagheri, M., and A. Zahmatkesh. 2018. "Evolution and Species-Specific Conservation of Toll-Like Receptors in Terrestrial Vertebrates." *International Reviews of Immunology* 37, no. 5: 217–228. https://doi.org/10.1080/08830185.2018.1506780.

Bayerova, Z., E. Janova, J. Matiasovic, L. Orlando, and P. Horin. 2016. "Positive Selection in the SLC11A1 Gene in the Family Equidae." *Immunogenetics* 68, no. 5: 353–364. https://doi.org/10.1007/s00251-016-0905-2.

Behzadi, P., H. A. García-Perdomo, and T. M. Karpiński. 2021. "Toll-Like Receptors: General Molecular and Structural Biology." *Journal of Immunology Research* 2021: 9914854. https://doi.org/10.1155/2021/9914854.

Bergman, I.-M., K. Edman, K. N. Ekdahl, K. J. Rosengren, and I. Edfors. 2012. "Extensive Polymorphism in the Porcine Toll-Like Receptor 10 Gene." *International Journal of Immunogenetics* 39, no. 1: 68–76. https://doi.org/10.1111/j.1744-313X.2011.01057.x.

Bergman, I.-M., J. K. Rosengren, K. Edman, and I. Edfors. 2010. "European Wild Boars and Domestic Pigs Display Different Polymorphic Patterns in the Toll-Like Receptor (TLR) 1, TLR2, and TLR6 Genes." *Immunogenetics* 62, no. 1: 49–58. https://doi.org/10.1007/s00251-009-0409-4.

Bielefeldt-Ohmann, H., A. Bosco-Lauth, A.-E. Hartwig, et al. 2017. "Characterization of Non-Lethal West Nile Virus (WNV) Infection in Horses: Subclinical Pathology and Innate Immune Response." *Microbial Pathogenesis* 103: 71–79. https://doi.org/10.1016/j.micpath.2016.12.018.

Bjørnstad, G., and K. H. Røed. 2001. "Breed Demarcation and Potential for Breed Allocation of Horses Assessed by Microsatellite Markers." *Animal Genetics* 32, no. 2: 59–65. https://doi.org/10.1046/j.1365-2052.2001.00705.x.

Brennan, J. J., and T. D. Gilmore. 2018. "Evolutionary Origins of Toll-Like Receptor Signaling." *Molecular Biology and Evolution* 35, no. 7: 1576–1587. https://doi.org/10.1093/molbev/msy050.

Bryant, C. E., and T. P. Monie. 2012. "Mice, Men and the Relatives: Cross-Species Studies Underpin Innate Immunity." *Open Biology* 2, no. 4: 120015. https://doi.org/10.1098/rsob.120015.

Cieslak, M., M. Pruvost, N. Benecke, et al. 2010. "Origin and History of Mitochondrial DNA Lineages in Domestic Horses." *PLoS One* 5, no. 12: e15311. https://doi.org/10.1371/journal.pone.0015311.

Dannemann, M., A. M. Andrés, and J. Kelso. 2016. "Introgression of Neandertal- and Denisovan-Like Haplotypes Contributes to Adaptive Variation in Human Toll-Like Receptors." *American Journal of Human Genetics* 98, no. 1: 22–33. https://doi.org/10.1016/j.ajhg.2015.11.015.

Darfour-Oduro, K. A., H.-J. Megens, A. L. Roca, M. A. M. Groenen, and L. B. Schook. 2015. "Adaptive Evolution of Toll-Like Receptors (TLRs) in the Family Suidae." *PLoS One* 10, no. 4: e0124069. https://doi.org/10.1371/journal.pone.0124069.

Futas, J., and P. Horin. 2013. "Natural Killer Cell Receptor Genes in the Family Equidae: Not Only Ly49." *PloS One* 8, no. 5: e64736. https://doi.org/10.1371/journal.pone.0064736.

Gasteiger, G., A. D'Osualdo, D. A. Schubert, A. Weber, E. M. Bruscia, and D. Hartl. 2017. "Cellular Innate Immunity: An Old Game With New Players." *Journal of Innate Immunity* 9, no. 2: 111–125. https://doi.org/10.1159/000453397.

Głażewska, I. 2010. "Speculations on the Origin of the Arabian Horse Breed." *Livestock Science* 129, no. 1: 49–55. https://doi.org/10.1016/j.livsci.2009.12.009.

Horín, P., E. G. Cothran, K. Trtková, et al. 1998. "Polymorphism of Old Kladruber Horses, a Surviving but Endangered Baroque Breed." *European Journal of Immunogenetics* 25, no. 5: 357–363. https://doi.org/10.1046/j.1365-2370.1998.00117.x.

Imler, J.-L., and L. Zheng. 2004. "Biology of Toll Receptors: Lessons From Insects and Mammals." *Journal of Leukocyte Biology* 75, no. 1: 18–26. https://doi.org/10.1189/jlb.0403160.

Ishengoma, E., and M. Agaba. 2017. "Evolution of Toll-Like Receptors in the Context of Terrestrial Ungulates and Cetaceans Diversification." *BMC Evolutionary Biology* 17, no. 1: 54. https://doi.org/10.1186/s12862-017-0901-7.

Janova, E., J. Futas, M. Klumplerova, et al. 2013. "Genetic Diversity and Conservation in a Small Endangered Horse Population." *Journal of Applied Genetics* 54, no. 3: 285–292. https://doi.org/10.1007/s13353-013-0151-3.

Johnson, C. M., E. A. Lyle, K. O. Omueti, et al. 2007. "Cutting Edge: A Common Polymorphism Impairs Cell Surface Trafficking and Functional Responses of TLR1 but Protects Against Leprosy." *Journal of Immunology* 178, no. 12: 7520–7524. https://doi.org/10.4049/jimmunol.178.12.7520.

Kawai, T., and S. Akira. 2011. "Toll-Like Receptors and Their Crosstalk With Other Innate Receptors in Infection and Immunity." *Immunity* 34, no. 5: 637–650. https://doi.org/10.1016/j.immuni.2011.05.006.

Klumplerova, M., P. Splichalova, J. Oppelt, et al. 2020. "Genetic Diversity, Evolution and Selection in the Major Histocompatibility Complex DRB and DQB Loci in the Family Equidae." *BMC Genomics* 21: 677. https://doi.org/10.1186/s12864-020-07089-6.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. "CLUMPAK: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across *K*." *Molecular Ecology Resources* 15, no. 5: 1179–1191. https://doi.org/10.1111/1755-0998.12387.

Kruithof, E. K. O., N. Satta, J. W. Liu, S. Dunoyer-Geindre, and R. J. Fish. 2007. "Gene Conversion Limits Divergence of Mammalian TLR1 and TLR6." *BMC Evolutionary Biology* 7, no. 1: 148. https://doi.org/10.1186/1471-2148-7-148.

Kumar, V. 2022. "Toll-Like Receptors in Adaptive Immunity." *Toll-like Receptors in Health and Disease: Handbook of Experimental Pharmacology*, edited by V. Kumar, 95–131. Springer. https://doi.org/10.1007/164_2021_543.

Lazarus, R., B. A. Raby, C. Lange, et al. 2004. "TOLL-Like Receptor 10 Genetic Variation Is Associated With Asthma in Two Independent Samples." *American Journal of Respiratory and Critical Care Medicine* 170, no. 6: 594–600. https://doi.org/10.1164/rccm.200404-491OC.

Lee, S., and G. Cho. 2006. "Parentage Testing of Thoroughbred Horse in Korea Using Microsatellite DNA Typing." *Journal of Veterinary Science* 7, no. 1: 63–67. https://doi.org/10.4142/jvs.2006.7.1.63.

Lindsay, S. A., and S. A. Wasserman. 2014. "Conventional and Non-Conventional *Drosophila* Toll Signaling." *Developmental and Comparative Immunology* 42, no. 1: 16–24. https://doi.org/10.1016/j.dci.2013.04.011.

Liu, G., H. Zhang, C. Zhao, and H. Zhang. 2020. "Evolutionary History of the Toll-Like Receptor Gene Family Across Vertebrates." *Genome Biology and Evolution* 12, no. 1: 3615–3634. https://doi.org/10.1093/gbe/evz266.

Ma, X., L. Yuhua, B. B. Gowen, E. A. Graviss, A. G. Clark, and J. M. Musser. 2007. "Full-Exon Resequencing Reveals Toll-Like Receptor Variants Contribute to Human Susceptibility to Tuberculosis Disease." *PLoS One* 2, no. 12: e1318. https://doi.org/10.1371/journal.pone.0001318.

Meyer, C. G., N. Reiling, C. Ehmen, et al. 2016. "TLR1 Variant H305L Associated With Protection From Pulmonary Tuberculosis." *PLoS One* 11, no. 5: e0156046. https://doi.org/10.1371/journal.pone.0156046.

Mukherjee, S., S. Huda, and S. P. Sinha Babu. 2019. "Toll-Like Receptor Polymorphism in Host Immune Response to Infectious Diseases: A Review." *Scandinavian Journal of Immunology* 90, no. 1: e12771. https://doi.org/10.1111/sji.12771.

Opsal, M. A., D. I. Våge, B. Hayes, I. Berget, and S. Lien. 2006. "Genomic Organization and Transcript Profiling of the Bovine Toll-Like Receptor Gene Cluster *TLR6-TLR1-TLR10*." *Gene* 384: 45–50. https://doi.org/10.1016/j.gene.2006.06.027.

Parker, H. G., L. V. Kim, N. B. Sutter, et al. 2004. "Genetic Structure of the Purebred Domestic Dog." *Science* 304, no. 5674: 1160–1164. https://doi.org/10.1126/science.1097406.

Peakall, R., and P. E. Smouse. 2012. "GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—An Update." *Bioinformatics* 28, no. 19: 2537–2539. https://doi.org/10.1093/bioinformatics/bts460.

Pérez-González, J., J. Carranza, G. Anaya, et al. 2023. "Comparative Analysis of Microsatellite and SNP Markers for Genetic Management of Red Deer." *Animals* 13, no. 21: 3374. https://doi.org/10.3390/ani13213374.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155, no. 2: 945–959. https://doi.org/10.1093/genetics/155.2.945.

Rahal, O., C. Aissaoui, N. Ata, et al. 2021. "Genetic Characterization of Four Algerian Cattle Breeds Using Microsatellite Markers." *Animal Biotechnology* 32, no. 6: 699–707. https://doi.org/10.1080/10495398.2020.1746321.

Roach, J. C., G. Glusman, L. Rowen, et al. 2005. "The Evolution of Vertebrate Toll-Like Receptors." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 27: 9577–9582. https://doi.org/10.1073/pnas.0502272102.

Rodrigues, C. R., Y. Balachandran, G. K. Aulakh, and B. Singh. 2024. "TLR10: An Intriguing Toll-Like Receptor With Many Unanswered Questions." *Journal of Innate Immunity* 16, no. 1: 96–104. https://doi.org/10.1159/000535523.

Schierup, M. H., X. Vekemans, and D. Charlesworth. 2000. "The Effect of Subdivision on Variation at Multi-Allelic Loci Under Balancing Selection." *Genetical Research* 76, no. 1: 51–62. https://doi.org/10.1017/s0016672300004535.

Seabury, C. M., E. J. Cargill, and J. E. Womack. 2007. "Sequence Variability and Protein Domain Architectures for Bovine Toll-Like Receptors 1, 5, and 10." *Genomics* 90, no. 4: 502–515. https://doi.org/10.1016/j.ygeno.2007.07.001.

Stejskalova, K., M. Cvanova, J. Oppelt, et al. 2019. "Genetic Susceptibility to West Nile Virus Infection in Camargue Horses." *Research in Veterinary Science* 124: 284–292. https://doi.org/10.1016/j.rvsc.2019.04.004.

Stejskalova, K., E. Janova, C. Horecky, et al. 2019. "Associations Between the Presence of Specific Antibodies to the West Nile Virus Infection and Candidate Genes in Romanian Horses From the Danube Delta." *Molecular Biology Reports* 46, no. 4: 4453–4461. https://doi.org/10.1007/s11033-019-04900-w.

Stejskalova, K., E. Janova, P. Splichalova, et al. 2023. "Twelve Toll-Like Receptor (TLR) Genes in the Family Equidae—Comparative Genomics, Selection and Evolution." *Veterinary Research Communications* 48: 725–741. https://doi.org/10.1007/s11259-023-10245-4.

Stevens, V. L., A. W. Hsing, J. T. Talbot, et al. 2008. "Genetic Variation in the Toll-Like Receptor Gene Cluster (*TLR10-TLR1-TLR6*) and Prostate Cancer Risk." *International Journal of Cancer* 123, no. 11: 2644–2650. https://doi.org/10.1002/ijc.23826.

Uddin, M. J., W. W. Suen, A. Bosco-Lauth, et al. 2016. "Kinetics of the West Nile Virus Induced Transcripts of Selected Cytokines and Toll-Like Receptors in Equine Peripheral Blood Mononuclear Cells." *Veterinary Research* 47: 61. https://doi.org/10.1186/s13567-016-0347-8.

Veltkamp, M., C. H. M. van Moorsel, G. T. Rijkers, H. J. T. Ruven, and J. C. Grutters. 2012. "Genetic Variation in the Toll-Like Receptor Gene Cluster (TLR10-TLR1-TLR6) Influences Disease Course in Sarcoidosis." *Tissue Antigens* 79, no. 1: 25–32. https://doi.org/10.1111/j.1399-0039.2011.01808.x.

Vijay, K. 2018. "Toll-Like Receptors in Immunity and Inflammatory Diseases: Past, Present, and Future." *International Immunopharmacology* 59: 391–412. https://doi.org/10.1016/j.intimp.2018.03.002.

Vychodilova, L., J. Matiasovic, O. Bobrova, et al. 2013. "Immunogenomic Analysis of Insect Bite Hypersensitivity in a Model Horse Population." *Veterinary Immunology and Immunopathology* 152, no. 3: 260–268. https://doi.org/10.1016/j.vetimm.2012.12.013.

Weaver, S., S. D. Shank, S. J. Spielman, M. Li, S. V. Muse, and S. L. Kosakovsky Pond. 2018. "Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes." *Molecular Biology and Evolution* 35, no. 3: 773–777. https://doi.org/10.1093/molbev/msx335.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.