



OPEN *TLR4* and *MD2* variation among horses with differential TNF α baseline concentrations and response to intravenous lipopolysaccharide infusion

Abhijit Mukhopadhyay¹, Shawna R. Cook², Phillip SanMiguel³, Kari J. Ekenstedt^{2,4} & Sandra D. Taylor^{1,4}✉

Gram-negative bacterial septicemia is mediated through binding of lipopolysaccharide (LPS) to mammalian toll-like receptor protein 4 (TLR4). TLR4 and its cognate protein, myeloid differentiation factor 2 (MD2) form a heterodimeric complex after binding LPS. This complex induces a cascade of reactions that results in increased proinflammatory cytokine gene expression, including TNF α , which leads to activation of innate immunity. In horses, the immune response to LPS varies widely. To determine if this variation is due to differences in TLR4 or MD2, DNA from 15 healthy adult horses with different TNF α dynamics after experimental intravenous LPS infusion was sequenced across exons of *TLR4* and *MD2*. Haplotypes were constructed for both genes using all identified variants. Four haplotypes were observed for each gene. No significant associations were found between either TNF α baseline concentrations or response to LPS and haplotype; however, there was a significant association (P value = 0.0460) between the baseline TNF α concentration and one *MD2* missense variant. Three-dimensional structures of the equine TLR4-MD2-LPS complex were built according to haplotype combinations observed in the study horses, and the implications of missense variants on LPS binding were modeled. Although the sample size was small, there was no evidence that variation in TLR4 or MD2 explains the variability in TNF α response observed after LPS exposure in horses.

Despite substantial advances in medical management, gram negative septicemia continues to be a significant cause of morbidity and mortality in horses^{1,2}. During gram negative septicemia, lipopolysaccharide binding protein binds the lipid A moiety of lipopolysaccharide (LPS; endotoxin) and transfers LPS to soluble (plasma) CD14 or membrane-bound CD14 on macrophages. This complex then binds and activates the toll-like receptor-4 (TLR4) and myeloid differentiation factor-2 (MD2) complex, which initiates signal transduction to ultimately induce production and release of pro-inflammatory cytokines such as TNF α ³. Compared to other species, horses are particularly sensitive to LPS^{4,5}, the deleterious effects of which are mediated through this TLR4 cell signaling. Given the relative sensitivity of horses to LPS, a non-lethal, low-dose (30 ng/kg) intravenous (IV) LPS administration model of gram negative septicemia has been established that can induce transient and systemic inflammation in the horse⁶. This model has been used extensively to investigate the clinical and clinicopathologic abnormalities associated with equine gram negative septicemia, and to test the efficacy of various drugs⁷⁻⁹. However, the individual equine response to IV LPS administration is variable⁹⁻¹². In horses with a strong inflammatory response to low-dose IV LPS infusion, clinical signs such as pyrexia, tachycardia, tachypnea, and colic occur in conjunction with neutropenia and a $\geq 50\%$ increase in TNF α serum concentration within one hour^{7,13}. Yet, in a study investigating the effects of ascorbic acid and hydrocortisone as a treatment protocol after infusion of low-dose IV LPS, eight of 40 horses were excluded due to complete lack of clinical and clinicopathological response to LPS⁹. Reports of naturally occurring cases of equine gram negative septicemia also demonstrate this

¹Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN, USA. ²Department of Basic Medical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN, USA. ³Bindley Bioscience Center, Purdue University, West Lafayette, IN, USA. ⁴These authors jointly supervised this work: Kari J. Ekenstedt and Sandra D. Taylor. ✉email: taylo248@purdue.edu

variation in response¹⁴. To date, factors that predict the clinical, clinicopathological, or cytokine response to LPS have not been identified in the horse.

The variation in response to LPS might be a result of molecular differences in equine *TLR4* and/or *MD2*. While *E. coli* LPS is a TLR4 agonist across all species^{15,16}, previous in vitro work has determined that equine TLR4 and MD2 are unique. For example, the TLR4/MD2 complex responds differently to LPS from *Rhodobacter sphaeroides* in equine cells, where it is an LPS agonist, versus human cells, where it is an LPS antagonist¹⁷. Indeed, discrete regions of both MD2 and TLR4 appear to be required for lipid IVa signaling, based on in vitro studies (lipid IVa being an LPS derivative that is an antagonistic in humans but an agonist in horses)¹⁸. Conversely, DiC14-amidine nanoliposomes are weak agonists in the horse, but strong agonists in human; here, two TLR4 regions were identified that modulated the human agonist activity of diC14-amidine, but both regions are outside the known LPS-binding domain¹⁵. Antagonists typically act by binding to monomeric TLR4/MD2 complexes in a non-dimerizing manner, preventing signaling¹⁶; it is reasonable to hypothesize that activity of antagonists could also be affected by differences in the TLR4 and/or MD2 proteins, even within a species.

Although the intracellular domains of human and murine TLR4 are highly conserved across species, the extracellular domains that contain the LPS binding site exhibit considerable sequence divergence in humans and mice^{19,20}. Variations in this extracellular domain can ultimately change receptor affinity and specificity to LPS⁴. Human and equine *TLR4* genes are similar (86.1% similarity [amino acid sequence homology]), comparing human NP_612564.1 and horse NP_001093239.2 in NCBI's BLAST [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>] using the Needleman-Wunsch global alignment²¹; however, divergence between species in the TLR4 extracellular domain is expected²². In a small study of five horses, four expressed single nucleotide polymorphisms (SNPs) in the *TLR4* gene but to date, the significance of these are unknown²³. In fact, the specific LPS binding site on TLR4 has not been defined in the horse, although, as mentioned above, the interactions of the equine TLR4/MD2 complex with various other agonists and antagonists have been defined. As with *TLR4*, the human and equine *MD2* genes are similar (81% similarity, comparing human NP_056179.4 and horse NP_001075367.1 as above)²¹, but sequence variations in *MD2* and the potential effects on LPS response among individual horses is unknown. In humans and mice, the three-dimensional (3D) crystal structure of the TLR4-MD2-LPS complex has been determined^{24,25}, and these can be used to predict the equine TLR4-MD2-LPS 3D structure. Sequence differences in *TLR4* and *MD2* resulting in changes to their respective proteins might lead to conformational changes that affect LPS binding and subsequent clinical responses to gram negative infections^{19,26}.

The first objective of this study was to sequence the exons of *TLR4* and *MD2* from 15 horses that were administered IV LPS. Amino acid (AA) sequences of both proteins were derived from gene sequencing. For both genes, both individual variants and constructed haplotypes were analyzed for association with baseline TNF α concentration and TNF α response to LPS. A second objective of the study was to describe the LPS binding sites on equine TLR4 and MD2 based on molecular modeling (as observed in the study horses) and determine the extent to which variations in either protein could affect LPS binding.

Results

TNF α concentrations at, and correlation between, baseline and response. Fifteen horses (designated H1–H15 and representing five breeds) had baseline TNF concentrations measured; nine horses were low (<10,000 pg/mL), three horses were moderate (10,000–100,000 pg/mL), and three horses were high (>100,000 pg/mL) (Table 1). Within one hour of IV LPS administration, the change in plasma concentrations of TNF α varied widely among horses; the percent change in plasma TNF α concentration ranged from -9% to 100% (Fig. 1, Table 1). Four horses (H4, H6, H8 and H9) were “high responders” ($\geq 50\%$ increase in plasma TNF α concentration), seven (H1, H2, H3, H5, H7, H10, H12, H14) had a minimal response (<20% increase in plasma TNF α concentration), and three (H11, H13 and H15) had a moderate response (20–49%).

When both baseline and response TNF α concentrations were categorized as high, moderate, or low, and subjected to a Kendall's tau test for ordinal-ordinal data, the correlation was -0.519, with $P=0.033$, suggesting that a correlation does exist between baseline and response.

Chr	EquCab3.0 Position	Ref	Alt	Consequence	Protein Position	SNP ID
25	22404058	G	T	Synonymous	260 L	rs782825011
25	22404206	G	A	Missense	D310N	rs782857623
25	22404251	A	C	Missense	K325Q	rs782902089
25	22404355	C	T	Synonymous	359 F	rs1144637357
25	22404541	C	T	Synonymous	421 G	rs1145524760
25	22404891	T	C	Missense	M538T	rs782839327
25	22405051	T	C	Synonymous	591 F	rs782871951
25	22405084	G	A	Synonymous	602 L	rs782895582
25	22405211	A	G	Missense	M645V	rs782844813

Table 1. Effect of IV LPS infusion on plasma TNF α concentrations in 15 horses. High, moderate, and low responders are defined as $\geq 50\%$, 20–49%, and <20% increase in plasma TNF α concentration, respectively. Baseline categories indicate pre-LPS administration TNF α concentrations as high (>100,000 pg/mL), moderate (10,000–100,000 pg/mL), and low (<10,000 pg/mL).

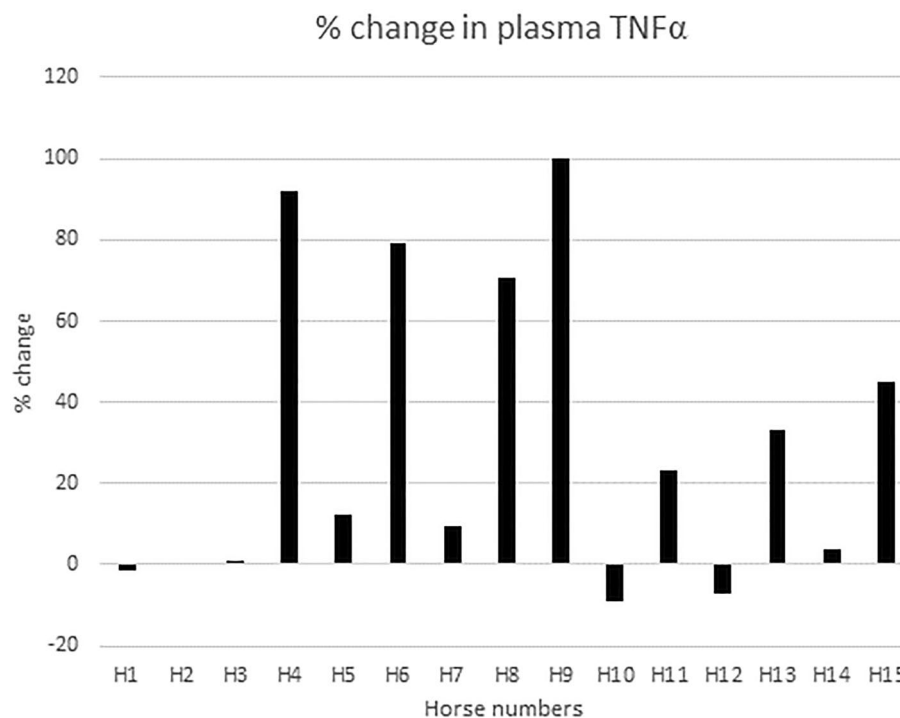


Figure 1. Percent change in plasma TNF α concentration. The percent change in plasma TNF α concentration one hour after IV LPS infusion compared to baseline values in 15 horses (H1–H15).

Identification of TLR4 and MD2 variants, haplotype construction, and association analysis. Sequencing of the *TLR4* gene exons revealed nine different variants, all previously reported (<https://www.ebi.ac.uk/eva/>; Table 2)²⁷, comprising four haplotypes among the 15 study horses (Supplementary Table S2). Four single nucleotide polymorphisms were missense changes, resulting in a predicted AA change; the remaining five variants were synonymous (no AA change). Of the four haplotypes, three shared an AA sequence, i.e.,

Horse, grouped by TNF α response (Breed)	Baseline TNF α (pg/mL) [Baseline category]	TNF α (pg/mL) one hour after LPS infusion	% increase in TNF α concentration
High responders			
H4 (Standardbred)	25.2 [Low]	310	1,129
H6 (Saddlebred)	115 [Low]	549	377
H8 (Thoroughbred)	322 [Low]	1089	238
H9 (Paint)	0 [Low]	93	N/A
Moderate responders			
H11 (Thoroughbred)	66,606 [Moderate]	86,792	30.3
H13 (Standardbred)	277 [Low]	415	50
H15 (Quarter Horse)	145 [Low]	264	82.1
Low responders			
H1 (Paint)	834,000 [High]	821,000	-1.6
H2 (Quarter Horse)	0 [Low]	0	N/A
H3 (Paint)	155,147 [High]	156,362	0.80
H5 (Thoroughbred)	18,225 [Moderate]	20,520	12.6
H7 (Standardbred)	508 [Low]	560	10.2
H10 (Thoroughbred)	32,440 [Moderate]	29,760	-8.3
H12 (Thoroughbred)	5,528 [Low]	5120	-7.4
H14 (Quarter Horse)	199,959 [High]	207,868	4.0

Table 2. *TLR4* variants identified in 15 horses with variable responses to intravenous LPS infusion. Chr: equine chromosome. EquCab3.0 Position: base pair position using EquCab3.0 genome build. SNP ID: drawn from European Variation Archive. Note that “Ref” and “Alt” alleles from EquCab3.0 are occasionally flipped compared to the GenBank sequence (NP_001093239.2).

the haplotypes differed at the base pair level, but not at the AA level. The two predicted AA sequences are shown in Supplementary Figures S1 and S2.

Three exonic *MD2* variants were identified among 11 horses compared to the GenBank reference sequence (Table 3); all three variants were previously reported (<https://www.ebi.ac.uk/eva/>)²⁷. Only one of the *MD2* variants was missense; the other two were synonymous. Four haplotypes were observed for *MD2* in this horse population (Supplementary Table S3); the wild type sequence was most prevalent (accounting for of 18/22 total haplotypes). Of the four haplotypes, three shared an AA sequence. Predicted AA sequences are shown in Supplementary Figure S3 and S4.

When categorized by TNF α response to IV LPS infusion (i.e. high, moderate, or low response), there was no association between either *TLR4* or *MD2* haplotype and response (P value = 0.814 and 0.602, respectively). Additionally, no individual variants within *TLR4* or *MD2* were associated with response (P values all > 0.058). Baseline TNF α (when categorized as high, moderate, or low) was associated with *TLR4* haplotype (P value = 0.018), but not *MD2* haplotype (P value = 0.316). However, follow-up pairwise comparisons of baseline TNF α categories (high, moderate, low) between each pair of *TLR4* haplotype, via two-tailed Cochran Armitage tests, revealed no significant differences after applying a Bonferonni correction (Table 4). When testing associations between individual variants and all phenotypes, all results were not significant, except for the individual baseline TNF α concentration for the missense *MD2* variant, which was significantly associated with phenotype even following permutation (P value = 0.046). This significant result was driven by horse H1, which had very high baseline TNF α concentration and was the only horse with *MD2* haplotype #4 (the missense variant).

Structure of equine TLR4-MD2-LPS complex. The structure of equine TLR4-MD2-LPS was constructed to determine if the differing AA residues identified in study horses were close the LPS binding site. While the human TLR4-MD2-LPS structure is known (Fig. 2A; PDB ID 3FXI), the corresponding equine structures, as determined by haplotypes in the present study, have not been determined to date. Sequence homology between the human and equine proteins allowed the construction of an equine TLR4-MD2-LPS model. One unit of the equine TLR4-MD2-LPS structure was built using the human TLR4 and MD2 as templates, and the equine structure was then merged with the human structure (Fig. 2B). As expected, the merged structures showed similarities between the two species. Next, the complete model of equine TLR4-MD2-LPS was built (Fig. 2C).

Haplotype analysis demonstrated that there were two TLR4 AA sequences and two MD2 AA sequences among the study horses. These AA sequences were designated TLR4-1 (Suppl Fig. S1), TLR4-2 (Supplementary Figure S2), MD2-1 (Supplementary Figure S3) and MD2-2 (Supplementary Figure S4). Each polypeptide combination (pair of haplotypes) observed among study horses were built (Fig. 3 and Supplementary Figures S5, S6, and S7). In TLR4-1, the relevant AA variations were Asp310, Lys325, and Met538. In TLR4-2, the relevant AA variations were Asn310, Gln325, and Thr538. MD2-1 had one AA variation, Arg106, whereas MD2-2 had Thr106. Four different TLR4-MD2-LPS complex formations were possible, given the haplotypes of the study horses: TLR4-1:TLR4-1:MD2-1:MD2-1 (designated as TLR4-MD2-LPS Complex 1), TLR4-2:TLR4-2:MD2-1:MD2-1 (designated as TLR4-MD2-LPS Complex 2), TLR4-2:TLR4-2:MD2-1:MD2-2 (designated as TLR4-MD2-LPS Complex 3), and TLR4-1:TLR4-2:MD2-1:MD2-1 (designated as TLR4-MD2-LPS Complex 4). These four structures were modeled to investigate possible effects of the variations on LPS binding.

The simulated model of Complex 1 (Supplementary Figure S5A) showed that: (1) the TLR4 Asp310 residues were not close to the LPS molecule, the TLR4-MD2 interface, or the dimer-dimer interface; (2) the TLR4 Lys325 residues were not close to the LPS molecule, and the distance between these two lysine residues from each subunit was 5.25 Å (Supplementary Figure S5B); (3) the TLR4 Met538 residues were located distant from relevant

Chr	EquCab3 Position	Ref	Alt	Consequence	Protein Position	SNP ID
9	12923460	G	A	Synonymous	55 K	rs3432833610
9	12923430	C	T	Synonymous	65 F	rs1146887989
9	12919464	G	C	Missense	R106T	rs1147734285

Table 3. *MD2* variants identified in 11 horses with variable responses to LPS infusion. Chr: equine chromosome. EquCab3.0 Position: base pair position using EquCab3.0 genome build. SNP ID: drawn from European Variation Archive. Note that “Ref” and “Alt” alleles from EquCab3.0 are occasionally flipped compared to the GenBank sequence (NP_001075367.1).

<i>TLR4</i> Haplotype	1	2	3
2	0.0494		
3	0.0275	0.6001	
4	0.1596	0.0675	0.0533

Table 4. Pairwise two-tailed Cochran Armitage test p-values for categorical baseline TNF α values (high, moderate, low) and *TLR4* haplotype. Following Bonferonni correction, a significant P would be ≤ 0.0083 .

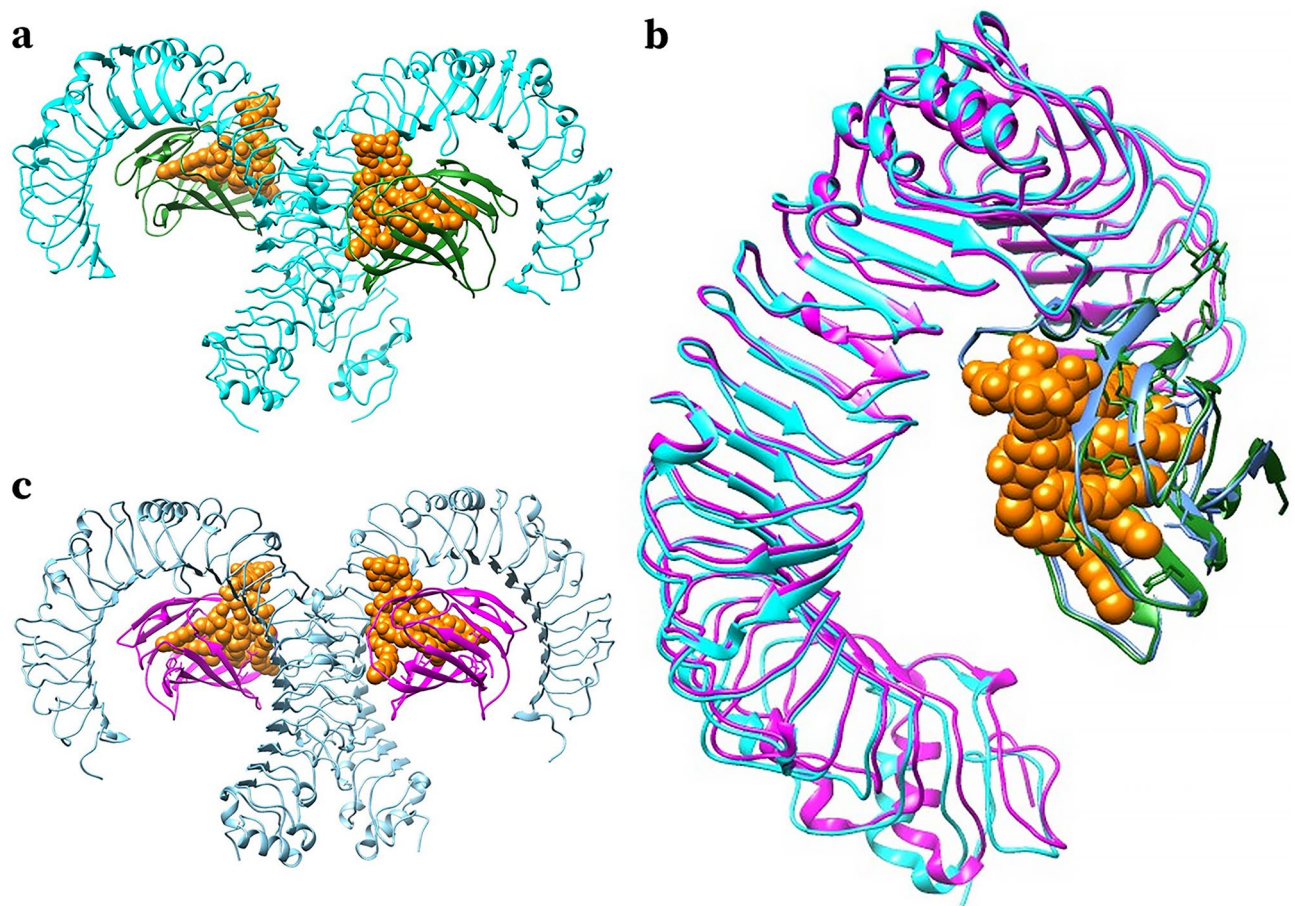


Figure 2. Construction of Equine TLR4-MD2-LP structure. (a) Structure of the human TLR4-MD2-LPS complex (PDB ID 3FXI)¹⁹, with TLR4 proteins (cyan blue), MD2 proteins (green), and LPS endotoxin (orange). (b) Merged equine (magenta) and human (cyan blue) TLR4 subunit, and merged equine (cornflower blue) and human (green) MD2 protein; LPS endotoxin (orange). (c) The complete structure of the equine TLR4-MD2-LPS complex, with TLR4 (sky blue), MD2 (magenta), and LPS endotoxin (orange).

sites; (4) the MD2 Arg106 was physically close to AA in the neighboring TLR4 chain; specifically, two TLR4 AA (Asp209 and Phe263) were both within 3.0 Å (H-bonding distance) of MD2 Arg106 (Supplementary Figure S5C).

Complex 2 (Supplementary Figure S6A), just as in Complex 1, demonstrated that the TLR4 Asn310 residues were not near either the LPS molecule, the TLR4-MD2 interface, or the dimer-dimer interface. Also similar to Complex 1, the MD2 Arg106 residue was physically close to AA in the neighboring TLR4 chain. Complex 2 differed from Complex 1 in that the TLR4 Gln325 residues were 11.5 Å apart (Supplementary Figure S6B).

Complex 3 (Fig. 3) was observed in only one horse (H1), where its individual baseline TNF α concentration was significantly associated with the missense MD2 variant (Arg106 versus Thr106; Fig. 3A). Complex 3 was identical to Complex 2 for TLR4 polypeptides, but MD2 existed in this horse as two different haplotypes. One copy of MD2, included Arg106 that was stabilized by interactions with Asp209, as shown in Supplementary Figure S5C. However, the other copy of MD2 included Thr106, to which the closest TLR4 residue was 4.42 Å away (Fig. 3B).

Complex 4 (Supplementary Figure S7) exhibited matched MD2 chains, as in Complexes 1 and 2. Complex 4 had two different TLR4 polypeptides, varying at three AA residues (Supplementary Figure S7A). The Lys325 from one TLR4 chain was located 6.74 Å from the Gln325 on the other TLR4 chain (Supplementary Figure S7B).

Surface charges for each TLR4 and MD2 haplotype observed in the study population were also modeled (Supplementary Figure S8A–D). The varying 325 residue of TLR4 (Lys325 versus Gln 325) and the varying 106 residue of MD2 (Arg106 versus Thr106) demonstrate surface charge differences that may affect binding.

Discussion

The low-dose IV LPS infusion model has been used extensively in equine research to study the pathophysiology of early gram negative septicemia and test the efficacy of potential drug candidates^{8,9,13,28–34}. A strong inflammatory response to low-dose IV LPS is typically associated with clinical signs of systemic inflammation as well as increased TNF α serum concentrations by $\geq 50\%$ ^{13,28,29}. However, up to 20% of horses fail to respond to this model⁹. One possible explanation is endotoxin tolerance, which is characterized by a diminished pro-inflammatory response to endotoxin that is recognized in animals and humans^{35–37}. However, endotoxin tolerance typically

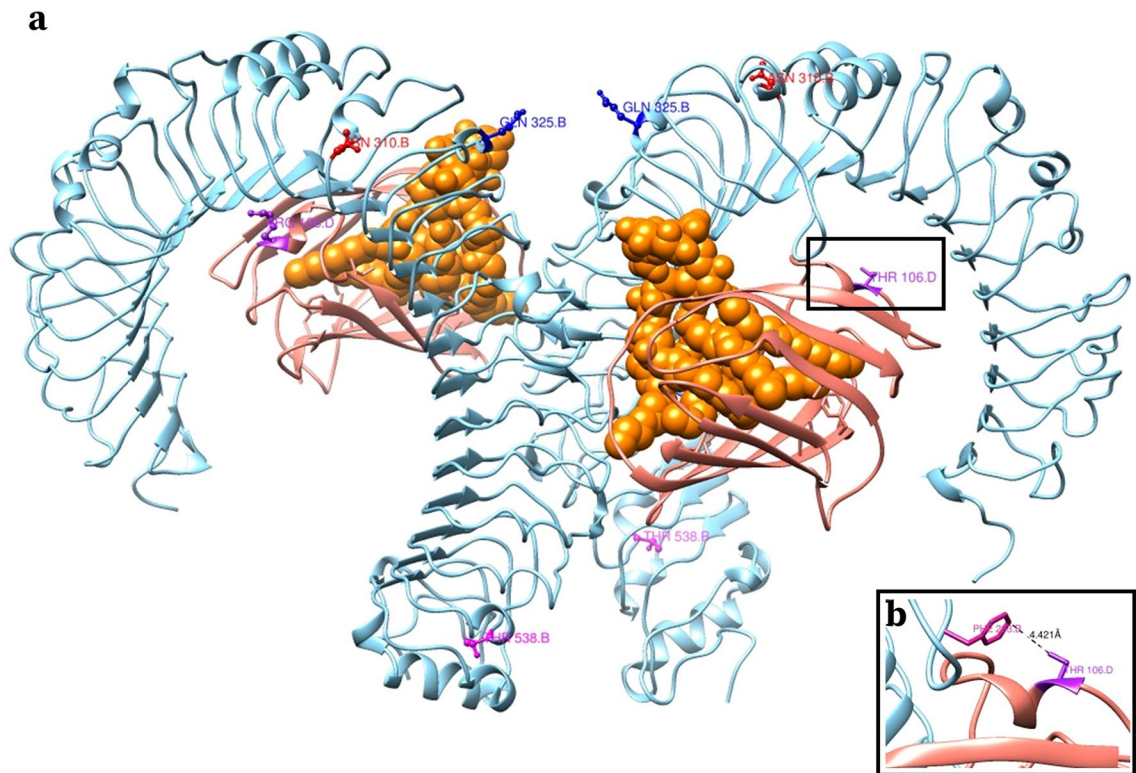


Figure 3. Complex 3: equine TLR4-MD2-LPS with TLR4-2:TLR4-2:MD2-1:MD2-2. **a.** The relevant amino acids are shown in a ball-and-stick model. **b.** The distance between MD2's Thr106 and TLR4's Phe263 (4.42 Å) is shown.

subsides within 3 weeks of endotoxin exposure in horses³⁸, making it an unlikely explanation in clinically normal horses. This study aimed to investigate TLR4 and MD2 variations among horses with different responses to LPS.

DNA sequencing of *TLR4* and *MD2* from horses representing multiple breeds with different TNF α dynamics after experimental IV LPS infusion revealed four haplotypes and two AA sequences for each gene. There were no associations between any haplotypes or single variant and TNF α response categories (high, moderate, low) for either gene. Similarly, when TNF α baseline concentrations were categorized as high, moderate, or low, there were no associations with haplotypes for either gene. The only significant association was between the individual baseline TNF α concentration and a missense *MD2* variant. However, this result was based on a single horse (H1) which had the highest baseline TNF α concentration and was the only horse to carry the *MD2* missense variant (the MD2-2 AA sequence). Baseline TNF α concentrations varied widely among study horses, in spite of meeting normal clinical inclusion criteria. This variability is discussed further below. H1 was a Paint, and while one other Paint horse (H3) also had a higher baseline TNF α concentration, the third Paint (H9) in the study had an undetectable baseline TNF α concentration; indeed, H9 had the lowest baseline value in the study. H1 may have been acutely exposed to gram negative bacteria prior to the LPS infusion in this study, accounting for the very high baseline TNF α concentration; however, this horse had no evidence of an ongoing inflammatory response and met the inclusion criteria as “normal” before the study began. Given the small sample size, conclusions regarding the significance of this single *MD2* missense variant association cannot be made. Previous *in vitro/ex-vivo* work evaluated responsiveness to LPS in a whole blood assay ($n = 10$ horses)³⁹; the same four *TLR4* missense mutations were detected among these horses and, identical to the current findings, no relationship was observed between any variant and altered LPS response. Unlike the present study no *MD2* variants were reported from this previous small group of horses; however, their sample population consisted of only one breed (Dutch Warmblood)³⁹.

Each haplotype combination was modeled as a 3D structure, using the known human TLR4 and MD2 structures as templates. This allowed visualization of each AA-changing variant's position in relation to the LPS binding site, the TLR4-MD2 interface, and the dimer-dimer interfaces, with the aim of elucidating whether any of the AA changes might alter function and explain the variable TNF α values. There are six lipid chains in each LPS molecule, five of which are buried inside the hydrophobic pocket. The remaining chain is exposed to the surface of MD2 and forms hydrophobic interactions with TLR4²⁴. Since LPS binds both TLR4 and MD2 proteins, variations in their critical LPS-binding domains or in overall protein structural integrity could lead to a significant effect. Complexes 1, 2, and 4 were not associated with any significant differences in TNF α values (baseline or response), although some of the AA changes could potentially result in alteration to attraction/repulsion at the quaternary protein structure level. For example, the variant Lys325 residues in Complex 1, which has two identical TLR4-1 subunits, are located just 5.25 Å apart. Lysine is a positively charged residue, and this small distance could create charge repulsion not observed in the wild type proteins. Similarly, also in Complex 1, the variant Arg106 residues in MD2-1 is close to TLR4's Asp209, which likely creates a strong ionic attraction

between the positively charged Arg106 and the negatively charged Asp209. Such small changes, though not statistically significant in the present work, may still contribute to overall protein stability.

Complex 3 contained one copy of the MD2-2 haplotype and therefore the MD2 missense variant, which was the only variant significantly associated with TNF α baseline concentration. In the dimer-dimer complex, one MD2 protein had Arg106 and the other had Thr106. One monomer (consisting of one TLR4 subunit and one MD2 subunit) having Arg106 is predicted to be stabilized by this residue interacting with TLR4's Asp209 (as shown in Supplementary Figure S5C), while the other monomer having Thr106 is predicted to have no substantial interactions with the neighboring TLR4 protein (Supplementary Figure S5B). Threonine, a polar, uncharged AA is smaller than arginine, which is a basic AA residue, and the closest TLR4 residue (Phe263) was 4.42 Å distant. It is unclear how this single missense change translates into the high baseline TNF α concentration in this single horse (H1). It is possible that the significant association is spurious and this missense change is well-tolerated. In fact, the European Variant Archive (<https://www.ebi.ac.uk/eva/>), under accession datasets PRJEB28306 and PRJEB9799, has whole genome sequence from 94 horses representing over two dozen breeds; nine of these horses are heterozygous for the same MD2 missense variant (minor allele frequency = 0.049). This demonstrates that the variant is not overly rare among horses, although the impact of this variant on MD2 protein function, response to LPS, and resulting TNF α concentration response, are still unknown. Ultimately, it is important to remember that H1 was clinically normal at baseline for all other parameters, yet very high for baseline TNF α concentration; the present protein modeling work has unfortunately not provided any meaningful additional explanation for this finding. Previous work has demonstrated that the MD2 122 locus, positioned on the outer lip of the MD2 pocket⁴⁰, varies between species (mouse, human, and horse), and that the R122E change in equine MD2 (arginine and glutamic acid, amino acids with opposite charge), can impair activation of equine TLR4⁴¹. Importantly, previous work examining TLR4 and MD2 sequence differences between four species (human, mouse, horse, and hamster), did not include the currently described MD2 106 variant⁴². More in-depth studies of the MD2 106 locus, ideally in an in vitro model, and potentially with additional agonist ligands, may identify different outcomes with the variant compared to wild type.

TLR4-MD2 heterodimer binding is mediated by a primary contact surface of TLR4 with both negatively charged and positively charged regions, which are complementary to the positive and negative charges on the surface of MD2, respectively^{43,44}. Mutations in these primary interfaces can disrupt TLR4-MD2 binding. Previous reports have demonstrated different surface charges between species; for example, species differences were noted in TLR4 surface charge distribution when comparing human to bovine¹⁹. Other work demonstrated differences in electrostatic potential based on MD2 surface charge between different species, including horse, and also suggested that TLR4 missense variants could change surface charge distribution in addition to other structural changes¹⁸. Therefore, surface charge differences likely influence ligand (LPS) binding. The surface charges for each TLR4 and MD2 subunit observed in the equine study population were modeled (Supplementary Figure S8), and differences were observed. The varying 325 residue of TLR4 (Lys325 versus Gln 325) does affect surface charge, which may affect the dimer-dimer stability at that location. Similarly, the varying MD2 106 residue (Arg106, a basic residue, versus Thr106, a neutral residue) also affects surface charge; because this residue is in close proximity to TLR4's Asp209 residue (acidic), the ionic interactions at this location may affect stability of the entire complex. While neither of these loci appears to directly affect LPS binding, even TLR4 and MD2 residues that do not directly participate in ligand binding can determine the signaling outcome of a given ligand⁴⁵.

Baseline TNF α concentrations varied widely among the horses in this study, despite their normal clinical examinations, CBCs, and biochemical profiles at the time of TNF α measurements. While a Kendall's tau test comparing baseline TNF α concentrations with response TNF α concentrations (each categorized as high, moderate, or low) suggested a significant negative correlation ($P=0.033$), the data set was very small. Further, all horses ($n=3$) that had a high baseline TNF α concentration were categorized as low responders, and every horse ($n=4$) that was categorized as a high responder started with a low baseline TNF α concentration. It is possible that subclinical inflammation was present in horses with "moderate" or "high" TNF α concentrations, but to date, large-scale screening studies of TNF α concentrations in healthy horses have not been performed, and there is a paucity in the literature of correlations between TNF α concentrations and clinical status in this species⁴⁶. It is also possible that initial high TNF α concentrations might prevent a further measurable response, rendering the response automatically to the "low" TNF α response category. Unexpectedly, two horses (H10 and H12) had decreased TNF α concentrations one hour after LPS infusion, the meaning of which is uncertain. Until more is known, it seems reasonable to continue to screen horses for TNF α concentrations, but perhaps not to use this metric as either inclusion or exclusion criteria; it may be prudent, though, to exclude horses with very high baseline cytokine levels in future studies. Future work could also include more controlled evaluation of the TLR4 and MD2 variants via in vitro stimulation of whole blood or PBMCs with increasing LPS doses and regular measurement of TNF α .

The major limitation to this study is the small sample size, with only 15 and 11 horses sequenced for TLR4 and MD2, respectively. While horses are an excellent clinical and research model for endotoxemia, they remain an expensive animal to house, decreasing the available sample size. Another limitation is the lack of measurement of other cytokines; TNF α concentrations may not be the best marker for LPS efficacy. In addition, the absence of generated sequences in the study horses for introns or untranslated regions of TLR4 or MD2 poses another limitation. Many non-coding variants can still impact gene and protein expression; these were not assessed in the present study. Furthermore, the present study did not sequence the equine TNF α gene, which has variants in its promoter/5'-UTR and throughout its structure (<https://www.ebi.ac.uk/eva/> and <https://www.ensembl.org/>)^{47,48}. Neither did this work examine gene sequence from LPS binding protein or CD14 in the study horses; both of these proteins are also involved in LPS binding. Future work would benefit from incorporation of sequence variant analysis in these genes as well. Finally, the possibility must be considered that some of the

in vivo response may be due to TLR2-mediated signals (in addition to TLR4), owing to signaling contaminants in the LPS preparation. This would obviously only affect TNF α response values and not baseline concentrations.

The overall goal of this study was to determine the relationship between gene variants and TNF α profile. Future work should examine a larger cohort of horses and investigate associations between numerous inflammatory cytokines (e.g. TNF α , IL1 β , IL6, and IL8) and TLR4 and MD2 variation. It is also possible that variants in other related genes (e.g., LPS binding protein, CD14, and TNF α) or other genetic predispositions are involved. Comparison among different breeds of horse might also show varying TLR4-MD2 complexes that could affect differences in LPS binding. This work combines for the first time *TLR4* and *MD2* genotyping and protein structural predictions with equine TNF α phenotypes and informs future work to further elucidate these molecular relationships.

Methods

Equine samples. Blood samples from 15 horses used in a previous study⁹ were evaluated for TNF α concentrations at the time of the original study; samples from both baseline and one hour after low-dose LPS (*E. coli* (0111:B4), Sigma Aldrich) administration were tested. Although details of LPS administration in these horses can be found elsewhere⁹, endotoxemia was induced by IV infusion of approximately 30 ng/kg (15 μ g) LPS in 500 mL 0.9% sodium chloride over a 30 min period. Baseline TNF α concentrations were categorized as low (< 10,000 pg/mL), moderate (10,000–100,000 pg/mL), and high (> 100,000 pg/mL) (Table 1). The change in plasma concentrations of TNF α was measured one hour after IV LPS administration and reported as a percent change. (Fig. 1, Table 1). Concentrations of TNF α were measured using a commercially available enzyme-linked immunosorbent assay (ELISA) kit validated for use in horses (Horse ELISA Kits, R&D Systems, Minneapolis, MN)⁴⁹. The mean of duplicate samples was recorded. The TNF α concentration data was generated in the context of the original study⁹, but is only being published now, together with sequencing data.

TLR4 and MD2 gene sequencing. Three mL of whole blood was collected once each from every horse still available from the original study (n = 15) for DNA sequencing of the *TLR4* gene (approved by Purdue University's Institutional Animal Care and Use Committee, #1803001719)⁹. Blood from 11 of the same 15 horses was available for *MD2* gene sequencing; four horses had inadequate sample for the *MD2* sequencing. For the *TLR4* gene sequencing, genomic DNA was isolated using the Qiagen DNeasy Blood and Tissue Kit following manufacturer's instructions. Equine-specific primers were designed using equine *TLR4* mRNA sequence (GenBank accession # NM_001099769.2) (Supplementary Table S1) and Primer3 software (<https://primer3.ut.ee/>);⁵⁰ the 1800 bp region that corresponded to the known extracellular domain coding region of the human *TLR4* gene was targeted. PCR was performed using New England BioLabs NEBNext Ultra II Q5 Master Mix, with PCR conditions following manufacturer instructions. PCR products were purified (Qiagen PCR purification kit) and sequenced at the Purdue University Bindley Bioscience Genomics Core Facility using the "WideSeq" service which leverages the Illumina DNA prep library construction kit and an Illumina MiSeq 500 cycle reagent cassette⁵¹.

The equine *MD2* coding sequence (GenBank accession number NM_001081898) is 482 bp long. Due to *MD2* having many introns, the direct genomic PCR methodology was passed over in preference for RT-PCR of mRNA, which was isolated from fresh whole blood samples (Qiagen Qiasm RNA Blood Mini Kit). PCR primers were designed based on equine reference *MD2* sequence (Supplementary Table S1) using Primer3. PCR products were generated using Invitrogen Superscript III One-step RT-PCR with Platinum Taq. The DNA sequencing of *MD2* PCR products was performed as described above. Gene variations in *TLR4* and *MD2* were defined as deviations from established sequences published as GenBank accession numbers NM_001099769.2 and NM_001081898, respectively.

Variant phasing and haplotype prediction. Identified genetic variants within *TLR4* and *MD2* were phased with SHAPEIT2 software using the default recombination rate⁵². Since this population was not composed of a singular breed, an effective population size of 341 was used; this is the average of the estimated effective population sizes across 36 breeds⁵³. Haplotypes were assessed using Haploview software⁵⁴.

Statistical analysis. The number of horses chosen for the study was based on available (live) horses with measured TNF α responses to LPS administration in a previous study⁹. All methods were performed in accordance with the relevant institutional guidelines and regulations. A Kendall's tau test (Kendall rank correlation coefficient) for ordinal-ordinal data was used to evaluate correlation between TNF α baseline and response, each as categorical data (high, moderate, low)⁵⁵. Fisher's exact tests were performed to identify any significant relationships between all *TLR4* or all *MD2* haplotypes and TNF α serum concentration, for both categorical response (high, moderate, low) and categorical baseline (high, moderate, low). Follow-up pairwise comparisons were made between individual haplotypes, with categorical baseline TNF α serum concentrations as phenotypes, using Cochran Armitage testing; a Bonferroni multiple-testing correction was applied to these tests. Finally, PLINK software⁵⁶ was used to test associations between individual variants and phenotypes (including categorical response, categorical baseline, and individual TNF α serum concentration; the latter used the measured TNF α concentrations as phenotypes), with an adaptive permutation applied to correct for multiple testing.

Construction of three dimensional (3D) structure of equine TLR4, MD2, and complexes. Amino acid sequences of TLR4 and MD2 were determined in SnapGene software (Insightful Science; snapgene.com). Equine TLR4 and equine MD2 AA sequences were aligned with human TLR4 and human MD2, respectively, using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) to determine the identity and similarity

between the two species⁵⁷. Equine and human TLR4 were found to be 73% identical and 82% similar, while MD2 was 65% identical and 81% similar. Because of the homology between species, it was possible to build 3D structures of equine TLR4 and MD2 proteins using human proteins as templates (Swiss-Model PDB viewer, <http://www.expasy.org/spdbv/>)^{58,59}. The Protein databank ID of the human TLR4-MD2-LPS structure used in the present study was 3FXI; a simulated equine TLR4-MD2-LPS complex model was created using 3FXI as a template. The human structure included AA from N-terminal 27 to C-terminal 631 for TLR4 and AA 19 to 158 for MD2. The corresponding horse structure included AA from N-terminal 23 to C-terminal 625 for TLR4 and from N-terminal 21 to C-terminal 155 for MD2. The 3D structures were visualized using the UCSF Chimera program (<https://www.cgl.ucsf.edu/chimera/>),⁶⁰ a structural biology tool that generates 3D images based on AA sequence. Distances between AA residues of interest (e.g., missense variants resulting in AA changes compared to reference) were measured in Angstroms (Å) based on the 3D models in UCSF Chimera. Lastly, surface charges for each TLR4 and MD2 monomer as observed in the study population were modeled in UCSF Chimera.

Data availability

All data generated or analyzed during this study are included in this published article and its Supplementary files.

Received: 15 September 2022; Accepted: 10 January 2023

Published online: 27 January 2023

References

- Moore, J. N. & Vandenplas, M. L. Is it the systemic inflammatory response syndrome or endotoxemia in horses with colic?. *Vet. Clin. North Am. Equine Pract.* **30**, 337–351. <https://doi.org/10.1016/j.cveq.2014.04.003> (2014).
- Taylor, S. A review of equine sepsis. *Equine Vet. Educ.* **27**, 99–109. <https://doi.org/10.1111/eve.12290> (2015).
- Akira, S. & Takeda, K. Toll-like receptor signalling. *Nat. Rev. Immunol.* **4**, 499–511. <https://doi.org/10.1038/nri1391> (2004).
- Vaure, C. & Liu, Y. A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Front. Immunol.* **5**, 316. <https://doi.org/10.3389/fimmu.2014.00316> (2014).
- Morris, D. D., Crowe, N. & Moore, J. N. Correlation of clinical and laboratory data with serum tumor necrosis factor activity in horses with experimentally induced endotoxemia. *Am. J. Vet. Res.* **51**, 1935–1940 (1990).
- MacKay, R. J., Merritt, A. M., Zertuche, J. M., Whittington, M. & Skelley, L. A. Tumor necrosis factor activity in the circulation of horses given endotoxin. *Am. J. Vet. Res.* **52**, 533–538 (1991).
- Jacobs, C. C. *et al.* Ethyl pyruvate diminishes the inflammatory response to lipopolysaccharide infusion in horses. *Equine Vet. J.* <https://doi.org/10.1111/j.2042-3306.2012.00634.x> (2012).
- Bonelli, F. *et al.* Kinetics of plasma procalcitonin, soluble CD14, CCL2 and IL-10 after a sublethal infusion of lipopolysaccharide in horses. *Vet. Immunol. Immunopathol.* **184**, 29–35. <https://doi.org/10.1016/j.vetimm.2016.12.010> (2017).
- Anderson, M. J. *et al.* Effects of administration of ascorbic acid and low-dose hydrocortisone after infusion of sublethal doses of lipopolysaccharide to horses. *J. Vet. Intern. Med.* **34**, 2710–2718. <https://doi.org/10.1111/jvim.15896> (2020).
- Werners, A. H., Bull, S. & Fink-Gremmels, J. Endotoxaemia: a review with implications for the horse. *Equine Vet. J.* **37**, 371–383 (2005).
- Vinther, A. M., Skovgaard, K., Heegaard, P. M. & Andersen, P. H. Dynamic expression of leukocyte innate immune genes in whole blood from horses with lipopolysaccharide-induced acute systemic inflammation. *BMC Vet. Res.* **11**, 134. <https://doi.org/10.1186/s12917-015-0450-5> (2015).
- Fessler, J. F. *et al.* Plasma endotoxin concentrations in experimental and clinical equine subjects. *Equine Vet. J. Suppl.* **21**, 24–28 (1989).
- Forbes, G., Church, S., Savage, C. J. & Bailey, S. R. Effects of hyperimmune equine plasma on clinical and cellular responses in a low-dose endotoxaemia model in horses. *Res. Vet. Sci.* **92**, 40–44. <https://doi.org/10.1016/j.rvsc.2010.10.020> (2012).
- Fogle, J. *et al.* Comparison of lipopolysaccharides and soluble CD14 measurement between clinically endotoxaemic and nonendotoxaemic horses. *Equine Vet. J.* **49**, 155–159. <https://doi.org/10.1111/evj.12582> (2017).
- Lopez, C. *et al.* Critical residues involved in Toll-like receptor 4 activation by cationic lipid nanocarriers are not located at the lipopolysaccharide-binding interface. *Cell Mol. Life Sci.* **72**, 3971–3982. <https://doi.org/10.1007/s00018-015-1915-1> (2015).
- Scior, T., Alexander, C. & Zaehring, U. Reviewing and identifying amino acids of human, murine, canine and equine TLR4 / MD-2 receptor complexes conferring endotoxic innate immunity activation by LPS/lipid A, or antagonistic effects by Eritoran, in contrast to species-dependent modulation by lipid IVa. *Comput. Struct. Biotechnol. J.* **5**, e201302012. <https://doi.org/10.5936/csbj.201302012> (2013).
- Lohmann, K. L., Vandenplas, M. L., Barton, M. H., Bryant, C. E. & Moore, J. N. The equine TLR4/MD-2 complex mediates recognition of lipopolysaccharide from *Rhodobacter sphaeroides* as an agonist. *J. Endotoxin Res.* **13**, 235–242. <https://doi.org/10.1177/0968051907083193> (2007).
- Walsh, C. *et al.* Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J. Immunol.* **181**, 1245–1254. <https://doi.org/10.4049/jimmunol.181.2.1245> (2008).
- Werling, D., Jann, O. C., Offord, V., Glass, E. J. & Coffey, T. J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends Immunol.* **30**, 124–130. <https://doi.org/10.1016/j.it.2008.12.001> (2009).
- Rehli, M. Of mice and men: species variations of Toll-like receptor expression. *Trends Immunol.* **23**, 375–378. [https://doi.org/10.1016/S1471-4906\(02\)02259-7](https://doi.org/10.1016/S1471-4906(02)02259-7) (2002).
- Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453. [https://doi.org/10.1016/0022-2836\(70\)90057-4](https://doi.org/10.1016/0022-2836(70)90057-4) (1970).
- EMBL-EBL. Clustal Omega, Multiple Sequence Alignment. (2021).
- Vychodilova-Krenkova, L., Matiasovic, J. & Horin, P. Single nucleotide polymorphisms in four functionally related immune response genes in the horse: CD14, TLR4, Cepsilon, and Fepsilon R1 alpha. *Int. J. Immunogenet.* **32**, 277–283. <https://doi.org/10.1111/j.1744-313X.2005.00522.x> (2005).
- Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **458**, 1191–1195. <https://doi.org/10.1038/nature07830> (2009).
- Kim, H. M. *et al.* Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130**, 906–917. <https://doi.org/10.1016/j.cell.2007.08.002> (2007).
- Hamann, L. *et al.* A coding mutation within the first exon of the human MD-2 gene results in decreased lipopolysaccharide-induced signaling. *Genes Immun.* **5**, 283–288. <https://doi.org/10.1038/sj.gene.6364068> (2004).
- Cezard, T. *et al.* The European Variation Archive: a FAIR resource of genomic variation for all species. *Nucleic Acids Res.* **50**, D1216–D1220. <https://doi.org/10.1093/nar/gkab960> (2022).

28. Taylor, S. D. *et al.* Effects of intravenous administration of peripheral blood-derived mesenchymal stromal cells after infusion of lipopolysaccharide in horses. *J. Vet. Intern. Med.* **36**, 1491–1501. <https://doi.org/10.1111/jvim.16447> (2022).
29. Jacobs, C. C. *et al.* Ethyl pyruvate diminishes the inflammatory response to lipopolysaccharide infusion in horses. *Equine Vet. J.* **45**, 333–339. <https://doi.org/10.1111/j.2042-3306.2012.00634.x> (2013).
30. Moore, J. N. *et al.* Rapid infusion of a phospholipid emulsion attenuates the effects of endotoxaemia in horses. *Equine Vet. J.* **39**, 243–248. <https://doi.org/10.2746/042516407x173343> (2007).
31. Alcott, C. J. *et al.* Clinical and immunomodulating effects of ketamine in horses with experimental endotoxemia. *J. Vet. Intern. Med.* **25**, 934–943. <https://doi.org/10.1111/j.1939-1676.2011.0749.x> (2011).
32. MacKay, R. J., Clark, C. K., Logdberg, L. & Lake, P. Effect of a conjugate of polymyxin B-dextran 70 in horses with experimentally induced endotoxemia. *Am. J. Vet. Res.* **60**, 68–75 (1999).
33. Cudmore, L. A., Muurlink, T., Whittem, T. & Bailey, S. R. Effects of oral clenbuterol on the clinical and inflammatory response to endotoxaemia in the horse. *Res. Vet. Sci.* **94**, 682–686. <https://doi.org/10.1016/j.rvsc.2013.01.003> (2013).
34. Kimura, S. *et al.* Effects of a single dose of orally and rectally administered misoprostol in an in vivo endotoxemia model in healthy adult horses. *Am. J. Vet. Res.* <https://doi.org/10.2460/ajvr.21.12.0206> (2022).
35. Tadros, E. M. & Frank, N. Effects of continuous or intermittent lipopolysaccharide administration for 48 hours on the systemic inflammatory response in horses. *Am. J. Vet. Res.* **73**, 1394–1402. <https://doi.org/10.2460/ajvr.73.9.1394> (2012).
36. Greisman, S. E. & Hornick, R. B. Mechanisms of endotoxin tolerance with special reference to man. *J. Infect. Dis.* **128**(Suppl), 265–276. https://doi.org/10.1093/infdis/128.supplement_1.s265 (1973).
37. Huber, R., Bikker, R., Welz, B., Christmann, M. & Brand, K. TNF tolerance in monocytes and macrophages: characteristics and molecular mechanisms. *J. Immunol. Res.* **2017**, 9570129. <https://doi.org/10.1155/2017/9570129> (2017).
38. Holcombe, S. J. *et al.* Duration of in vivo endotoxin tolerance in horses. *Vet. Immunol. Immunopathol.* **173**, 10–16. <https://doi.org/10.1016/j.vetimm.2016.03.016> (2016).
39. Werners, A. H. *et al.* Genotyping of Toll-like receptor 4, myeloid differentiation factor 2 and CD-14 in the horse: an investigation into the influence of genetic polymorphisms on the LPS induced TNF-alpha response in equine whole blood. *Vet. Immunol. Immunopathol.* **111**, 165–173. <https://doi.org/10.1016/j.vetimm.2005.12.003> (2006).
40. Scior, T. *et al.* Three-dimensional mapping of differential amino acids of human, murine, canine and equine TLR4/MD-2 receptor complexes conferring endotoxic activation by lipid A, antagonism by Eritoran and species-dependent activities of Lipid IVA in the mammalian LPS sensor system. *Comput. Struct. Biotechnol. J.* **7**, e201305003. <https://doi.org/10.5936/csbi.201305003> (2013).
41. Oblak, A. & Jerala, R. The molecular mechanism of species-specific recognition of lipopolysaccharides by the MD-2/TLR4 receptor complex. *Mol. Immunol.* **63**, 134–142. <https://doi.org/10.1016/j.molimm.2014.06.034> (2015).
42. Anwar, M. A., Panneerselvam, S., Shah, M. & Choi, S. Insights into the species-specific TLR4 signaling mechanism in response to Rhodobacter sphaeroides lipid A detection. *Sci. Rep.* **5**, 7657. <https://doi.org/10.1038/srep07657> (2015).
43. Ain, Q. U., Batool, M. & Choi, S. TLR4-targeting therapeutics: structural basis and computer-aided drug discovery approaches. *Molecules* <https://doi.org/10.3390/molecules25030627> (2020).
44. Jin, M. S. & Lee, J. O. Structures of the toll-like receptor family and its ligand complexes. *Immunity* **29**, 182–191. <https://doi.org/10.1016/j.immuni.2008.07.007> (2008).
45. Irvine, K. L. *et al.* Identification of key residues that confer Rhodobacter sphaeroides LPS activity at horse TLR4/MD-2. *PLoS ONE* **9**, e98776. <https://doi.org/10.1371/journal.pone.0098776> (2014).
46. Cacciolatti, C. *et al.* In vitro effects of poly(ADP-ribose) polymerase inhibitors on the production of tumor necrosis factor-alpha by interferon-gamma—and lipopolysaccharide-stimulated peripheral blood mononuclear cells of horses. *Am. J. Vet. Res.* **80**, 663–669. <https://doi.org/10.2460/ajvr.80.7.663> (2019).
47. Horin, P. *et al.* Polymorphisms in equine immune response genes and their associations with infections. *Mamm Genome* **15**, 843–850. <https://doi.org/10.1007/s00335-004-2356-6> (2004).
48. Brown, J. J. *et al.* TNF-alpha SNP haplotype frequencies in equidae. *Tissue Antigens* **67**, 377–382. <https://doi.org/10.1111/j.1399-0039.2006.00581.x> (2006).
49. Lavoie-Lamoureux, A., Maghni, K. & Lavoie, J. P. Optimization of a procedure to accurately detect equine TNFalpha in serum samples. *Vet. Immunol. Immunopathol.* **138**, 118–123. <https://doi.org/10.1016/j.vetimm.2010.06.018> (2010).
50. Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115. <https://doi.org/10.1093/nar/gks596> (2012).
51. https://support.illumina.com/content/dam/illumina/support/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prep-reference-guide-1000000025416-09.pdf Vol. 1000000025416 (2020).
52. Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6. <https://doi.org/10.1038/nmeth.2307> (2013).
53. Petersen, J. L. *et al.* Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS ONE* **8**, e54997. <https://doi.org/10.1371/journal.pone.0054997> (2013).
54. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265. <https://doi.org/10.1093/bioinformatics/bth457> (2005).
55. Khamis, H. Measures of association how to choose?. *J. Diagn. Med. Sonogr.* **24**, 155–162. <https://doi.org/10.1177/8756479308317006> (2008).
56. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7. <https://doi.org/10.1186/s13742-015-0047-8> (2015).
57. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkac240> (2022).
58. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303. <https://doi.org/10.1093/nar/gky427> (2018).
59. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723. <https://doi.org/10.1002/elps.1150181505> (1997).
60. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).

Acknowledgements

The authors thank Theresa Jeffrey for her technical assistance.

Author contributions

Conceptualization: A.M., K.J.E., S.T. Data curation: A.M., S.C. Formal analysis: A.M., S.C. Methodology: A.M., P.S.M., S.C. Resources: A.M., K.J.E., S.T. Supervision: K.J.E., S.T. Visualization: A.M., S.C. Writing—original draft: A.M., S.C., K.J.E., S.T. Writing—review and editing: A.M., S.C., K.J.E., P.S.M., S.T. All authors have read and agreed to the submitted version of the manuscript.

Funding

This study was supported by the Purdue University College of Veterinary Medicine, Veterinary Clinical Sciences Faculty Pilot Research grant and the USDA National Institute of Food and Agriculture (Hatch Project # 1016369). Partial support for KJE was provided by the Office of the Director, National Institutes of Health under award number K01-OD027051.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-27956-y>.

Correspondence and requests for materials should be addressed to S.D.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023