diversity

MDPI

Article

Haplotype Disequilibrium in the *TLR* Genes of Czech Red Pied Cattle

Kalifa Samaké [1] and Karel Novák [2,*]

[1] Department of Genetics and Microbiology, Charles University, Prague 2, 120 00 Prague, Czech Republic
[2] Institute of Animal Science, Uhříněves, Prague 22, 104 00 Prague, Czech Republic
* Correspondence: novak.karel@vuzv.cz; Tel.: +420-267009567

Abstract: Hybrid resequencing of the antibacterial innate immune genes coding for toll-like receptors, namely *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6*, using HiSeq and PacBio technologies of pooled population samples of Czech Simmental (Czech Red) cattle allowed us to determine haplotypes formed by the polymorphisms present. Directly determined haplotypes within the range of the large proximal amplicon in *TLR2* formed two clusters in the network tree graph. The distribution of the statistically reconstructed haplotypes based on individual genotyping of the present SNPs was consistent. Similarly, the statistically reconstructed haplotypes in *TLR5* and *TLR6* formed two clusters. The trend of bimodal distribution was also observed in *TLR4*, while the limited diversity of *TLR1* did not allow for any conclusion. The observed bimodal distribution is consistent with earlier reports for cattle populations worldwide. The stability of this phenomenon cannot be ascribed to historical origin but rather to a long-term effect of balancing selection. The equilibrium might be based on two different essential functions performed by the *TLR* genes or their products. The formation of two kinds of heterodimers by the *TLR2* product, namely, *TLR2*/*TLR1* and *TLR2*/*TLR6* with different ligand specificities, is considered to be a particular case. On the other hand, the better expression of the bimodal groups in the 5′-proximal SNPs supports the localization of the selection targets in the upstream regulatory regions or the functional interactions in the proximal part of the transcripts.

Keywords: innate immunity; toll-like receptors; antibacterial TLRs; infection pressure; balancing selection

Citation: Samaké, K.; Novák, K. Haplotype Disequilibrium in the *TLR* Genes of Czech Red Pied Cattle. *Diversity* **2023**, *15*, 811. https://doi.org/10.3390/d15070811

Academic Editor: Luc Legal

Received: 15 March 2023
Revised: 17 June 2023
Accepted: 22 June 2023
Published: 27 June 2023

1. Introduction

Animal toll-like receptors (TLRs) play a key role in the recognition of the conserved molecules of bacterial, fungal, and viral origin, the so-called pathogen-associated molecular patterns (PAMPs) [1], as a class of pattern-recognizing receptors (PRRs). Although originally discovered in *Drosophila* as a morphogenetic gene [2] and only later recognized as an essential factor of resistance against pathogens [3], they are distributed almost universally. Toll-like receptors evolved in the animal kingdom starting in *Porifera* [4]. Homologs of the most conserved region of the toll-like receptor, region TIR (Toll interleukin-1 receptor and resistance genes), can be found even in plant genes for antimicrobial defense [5].

Toll-like receptors are activated by ligands originating from pathogens and involved in pathogenesis (pathogen-associated molecular patterns, PAMPs). Subsequently, they mediate downstream triggering of the transcriptional response via the toll-signaling pathway [1]. In addition to their function in innate immunity in vertebrates, the pleiotropic effects of toll-like receptor coding genes (*TLR*s) have been demonstrated in many other traits, both at the molecular and organismal levels [6]. For example, effects on parturition in the mouse model have been demonstrated using a line with a *TLR4* knockout [7]. In livestock species, *TLR*s have been considered prospective targets for breeding mainly as representatives of innate immunity genes [8,9].

In cattle (*Bos taurus* L.), ten paralogues denoted as *TLR1–TLR10* are differentiated into the so-called antibacterial and antiviral series [10]. Genes *TLR1*, *TLR2*, *TLR4*, and *TLR6*

recognize polysaccharidic or glycoprotein ligands originating from the cell walls of Gram-negative and Gram-positive bacteria. The product of *TLR5* is an exception by recognizing the protein flagellin from bacterial flagella. The toll-like protein product consists of the outer region, which is extracellular in the antibacterial group, the transmembrane part, and the inner (intracellular) region.

Natural variation in bovine *TLR* genes has been repeatedly reported for a panel of world cattle breeds [11–13], and is documented in public nucleotide sequence databases, such as the European Variation Archive (EVA) of the European Bioinformatics Institute (https://www.ebi.ac.uk/eva/?Home; accessed on 21 June 2023) and ENSEMBL (https://www.ensembl.org/info/docs/tools/vep/index.html; accessed on 21 June 2023). Some of the variants in coding regions are predicted to disturb the function of the protein product [14–18]. According to expectations, associations with health traits in cattle populations have been reported for naturally occurring variants [13,16,19]. For instance, the predicted functional change in c.2021C>T in *TLR4*, affecting the transmembrane region, causes a shift in somatic cell count in milk [14,20]. In *TLR2*, the 1047G>T and 1313G>A nonsynonymous polymorphisms have been reported to be effective in increasing susceptibility to mycobacterial infections in cattle—tuberculosis and paratuberculosis (PTB) [16–18]. Nevertheless, in many cases, synonymous mutations without a change to the protein structure or even mutations in the noncoding regions of bovine *TLR* genes demonstrated associations with phenotypic traits. For example, noncoding 1313G>A in *TLR2* is part of a haplotype that increases susceptibility to PTB [19].

Consequently, the effects of haplotypes should be considered in place of the effects of the causal SNPs alone. In contrast to simple SNPs, haplotypes comprising particular blocks of SNPs and identified via tagSNPs integrate causal polymorphisms in the coding and regulatory regions [21]. Therefore, knowledge of haplotype structure is a prerequisite for objective association studies with consequences for breeding.

With respect to this assumption, we decided to explore the haplotype structure of antibacterial *TLR* species in the population of Czech Red Pied (CRP) cattle. The choice of the Czech Red Pied cattle is due to its role among the cattle breeds in the agriculture of the Czech Republic. CRP cattle are the second most abundant breed, with a tradition extending back to the middle of the 19th century. Moreover, the availability of an almost complete set of archived samples characterizing full genetic variability (insemination doses in the breeding firm CHD Impuls, blood samples in the Institute of Animal Science) was taken into account. The Czech Red Pied cattle is a member of the Simmental breed type group. CRP cattle were formed during the 19th and 20th centuries from the import of the original Simmental cattle, presumably by including the features of local adaptation from local historical breeds like Czech Red cattle. The state corresponding to the year 2000 is now conserved as a nucleus herd in the context of the Czech National Programme for the Conservation of Genetic Resources.

Therefore, the historical changes in the variation of innate immune genes and the impact of the recent genomic selection cannot be excluded. Moreover, this breed is unrelated to the breeds used in previous studies on *TLR* diversity in cattle [11–13,18].

The data originated from a screening of the variability in *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6* that was performed in the CRP cattle population. A pooled DNA sample was resequenced with two NGS technologies: targeted PacBio RSII was combined with whole-genome HiSeq X Ten to reduce the incidence of false positives. The haplotypes were directly determined from long PacBio reads on the one hand and statistically reconstructed from the results of subsequent individual genotyping on the other hand. The distribution of haplotypes was visualized with a tree graph that allowed us to demonstrate the difference from the expected equilibrium corresponding to random recombination. The found departure from equilibrium, which could be observed as a bimodal distribution of haplotype clusters, was interpreted with respect to the assumed balancing selection.

2. Materials and Methods

2.1. Animals

Thanks to the support of the breeding firm CHD Impuls (Bohdalec, Czech Republic) and the Association of the Breeders of the Czech Red Pied Cattle, an almost complete set of archived bull samples characterizing full variability in this breed was available. The modern population of Czech Red Pied cattle was represented by a set of 164 bulls originating from the portfolio of CHD Impuls.

2.2. DNA Samples

Genomic DNA was prepared from cryopreserved insemination doses using affinity binding on paramagnetic particles with the MagSep Tissue Kit (Eppendorf, Hamburg, Germany). A normalized gDNA set containing 20 ng/µL DNA from each animal was prepared according to the concentrations determined spectrophotometrically and additionally purified with the AMPure XP magnetic bead procedure (Beckman Coulter, Brea, CA, USA).

2.3. Next Generation Sequencing (NGS)

The conditions of the NGS methods applied have already been published elsewhere (Novák et al., 2019). In brief, screening for polymorphisms in the bovine antibacterial *TLR* series, i.e., *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6*, was performed by hybrid resequencing of pooled DNA samples of 164 bulls. The amplicon panel (Supplementary Table S1) was prepared in standard PCR mixtures [22], and the products were sequenced with consensus circular sequencing (CCS) technology in the PacBio RSII sequencer (Pacific Biosciences, Menlo Park, CA, USA) in the core laboratory of Eurofins (GATC Services, Constance, Germany). The amplicons covering coding sequences and flanking regions basically followed a previous publication [22] and were adapted according to the amplicons designed by White et al. [23] and Seabury et al. [24]. To reduce the error rate of sequencing, these PacBio sequencing results were completed with whole-genome sequencing (WGS) using HiSeq X Ten technology (Illumina, San Diego, CA, USA) in the laboratory of Novogene (London, UK). The obtained 60× coverage in the HiSeq sequencing was sufficient to detect polymorphisms above the 0.05 frequency at 95% efficiency.

2.4. Sequencing Data Processing

The read assemblies were built on the Fasta files of reads using the Geneious Mapper algorithm implemented in the Geneious program package (Biomatters, Auckland, New Zealand) and the UMD 3.1.1 bovine genome sequence (https://www.ncbi.nlm.nih.gov/assembly/GCF_000003055.6/) (accessed on 21 June 2023) issued by the Centre for Bioinformatics and Computational Biology, University of Maryland. The mapping to the reference sequence was iterated 5 times to reduce the error rate. The reliability of the single nucleotide polymorphism (SNP) detection was enhanced by comparing the results obtained with the two sequencing platforms. Additionally, the detected polymorphisms were verified by matching with the variants present in the EVA database. For compactness and continuity with previous publications, the individual gene reference sequences FJ147090 (*TLR1*), EU746465 (*TLR2*), AC000135.1 (*TLR4*), EU006635 (*TLR5*), and AJ618974 (*TLR6*) were used for SNP identification.

2.5. Characterisation of SNPs

The functional consequences of the validated mutations were estimated using the prediction software Variant Effect Predictor (VEP) distributed by the ENSEMBL database. These results were confronted with an independent evaluation of the calculated impact using the SIFT (Sorting Intolerant from Tolerant) program [25].

2.6. Genotyping

Subsequent genotyping of individual animals was performed using the primer extension method [26] in the commercial modification SNaPshot (Thermo Fisher Scientific,

Waltham, USA). The genotyped SNPs in coding or presumed regulatory regions partially overlapped with those used in previous association studies in this cattle population [27]. The individually genotyped SNPs, along with the corresponding SNaPshot reactions, are specified in Table 1.

2.7. Haplotype Determination

When the density of SNPs was sufficient, the haplotypes present in the populations and their frequencies were determined directly from the long-read PacBio amplicon sequences. In practice, this approach was applied to the proximal amplicon 1 of *TLR2*.

In parallel, the haplotypes present in the population were statistically reconstructed from the SNPs determined by individual genotyping. The reconstruction was performed using the PHASE program [28] version 2.1. The program is based on the Bayesian algorithm for haplotype phasing and is available at http://www.stat.washington.edu/stephens/software.html (accessed on 21 June 2023). The statistical reconstruction allowed for the determination of haplotypes on a whole-gene scale.

2.8. Haplotype Graph Construction

The distances among haplotypes and haplotype frequencies were visually presented as tree graphs. Data were used for the initial network graph construction using the median-joining (MJ) algorithm according to Bandelt et al. [29], as implemented in the Network program (Fluxus Technology, Colchester, UK). The algorithm combines the advantages of Kruskal's algorithm for finding minimum spanning trees and Farris's maximum-parsimony (MP) heuristic. The algorithm favors short connections and adds new virtual vertices called "median vectors". In the second step, the superfluous (nonparsimonial) links are reduced using the maximum parsimony (MP) method, according to Polzin and Daneschmand [30].

The most likely tree of the set of the minimum spanning trees obtained upon the reduction step was used for presentation. The remaining tree variants did not differ significantly, and the differences mostly included only the positions of leaves within the tree graph. The node positions were adjusted by editing for the sake of graph comprehension.

3. Results

3.1. Directly Read Haplotypes

The SNPs present in the population were established within the range of the designed PacBio amplicons at 200–300× coverage. The existence of SNPs was verified by the consistency with the SNPs revealed in short reads provided by the HiSeq technology at 60× coverage. In the next step, the realness of the found SNPs was supported by the match with the known SNPs present in the EVA database. Data on the variability in antibacterial *TLR*s in the studied population have been partially published elsewhere [22,27].

Although most of the PacBio amplicons used (according to Supplementary Table S1) did not yield a sufficient number of SNPs to create an informative haplotype graph, the proximal amplicon of *TLR2* (amplicon T2_1 in Supplementary Table S1) contained nine SNPs (Table 2) that allowed for direct haplotype determination. The description of directly read haplotypes present in this region of amplicon T2_1 is summarized in Supplementary Table S2.

The directly read haplotypes were used as an input in the Network program. The haplotype frequencies in the population and the haplotype relatedness were visualized as nodes and edges of a created network graph, subsequently reduced to the most likely tree from among the set of the minimum spanning trees, i.e., graphs without cyclic structures, connecting all vertices and with a minimized summary weight of the edges. The node size represents the particular haplotype frequency in the population. However, the connecting path within the tree does not necessarily correspond to the shortest path between two nodes in the full network. The finding of the minimum tree is complex and includes the sum of the weights of all edges across the tree graph. This approach is illustrated in Figure 1A.

Table 1. Individual genotyping reactions in *TLR1*, *-2*, *-4*, *-5*, and *-6* and basic characteristics of genotyped polymorphisms in the bull population of the CRP cattle, partially following Bjelka and Novák [27].

Gene	SNP Identifier	Chromosomal Position	SNP [a]	SNaPshot Primer (Sequence 5′→3′)	Multiplex	SNP Type	Predicted Effect [b]	Variant Frequency
TLR1	rs43702940	6_59688857	798C>T	CGCCAAACCAACTGGAGGATCGT	A1	synonymous	low	0.473
TLR1	rs210538093	6_59687893	1762G>A	CAGCCCAGGGACCACAATGGTGA	C1	Val523Ile	moderate	0.380
TLR1	rs109456287	6_59687558	2097T>C	(T4)CTCTGGACAAAGTTGGGAGACAAGAC	C2	synonymous	low	0.780
TLR2	rs68343162	17_3953930	115T>C	(T10)GAAATAACCAAGAGGGAAATGGAAT	A2	intron	modifier	0.827
TLR2		17_3953036	1009A>G	ACACACCTCTGCAGGTCTCTGTTGC	B1	Ala151Thr	tolerated	0.260
TLR2	rs68268249	17_3953001	1044T>C	(T9)GTGAAAAGCCTTGACCTGTCCAACAA	B3	synonymous	low	0.847
TLR2	rs55617172	17_3952998	1047G>T	(T12)GCAGGTCTCTGTTGCYGACATAGGTGAT	B4	missense, 63Glu>Asp	moderate	0.490
TLR2	rs68268250	17_3952985	1060G>A	(T5)CACACCTCTGCAGGTCTCTGTTGC	B2	Gly68Ser	moderate	0.700
TLR2	rs43706434	17_3952732	1313G>A	(T18)CTGTTACTATTTCCTACTTTTAGGGTC	B5	Arg152Gln in LRR5 [c]	moderate	0.787
TLR2	rs68268260	17_3951499	2546G>A	(T12)GGTCGACTGGCCCGATGACTACC	C3	Arg563His in LRR20c	moderate	0.877
TLR2	rs41830058	17_3951480	2565T>C	(T16)CTACCRCTGTGACTCTCCCTCCCA	C4	synonymous	low	0.688
TLR2	rs68343171	17_3951162	2883T>C	(T18)CCACTTGCCAGGAATGAAGTCTCGCTT	C5	synonymous	low	0.118
TLR2	rs68268268	17_3950839	3206G>A	(T17)GGTTAAATTTGAGAGCTGCAATAA	F1	Arg782Lys	tolerated	0.833
TLR4	rs29017188	8_108829143	245G>C	CTTCTTCTTCCTCTAACTTCCCCTC	D1	5′-UTR	changing expression	0.401
TLR4	rs43578094	8_108829508	610C>T	(T5)GGGCCCAGCACAGGGAAACTGAGCA	D2	intron	modifier	0.931
TLR4	rs8193046	8_108833985	5087A>G	(T10)GCTAAGGTGCATGCAGGAAGACACC	D3	intron	modifier	0.575
TLR4	rs8193047	8_108834032	5134G>A	(T13)GATTTTGTAGAGATTCAGCTCCATGCA	D4	synonymous	low	0.849
TLR4	rs43578100	8_108836897	7999A>G	(T21)GGTTTCCTATTCAGCAGAAATATT	D5	intron	modifier	0.463
TLR4	rs8193060	8_108838320	9422C>T	ACTCGCTCCGGATCCTAGACTGCAG	E1	synonymous	low	0.356
TLR4	rs8193072	8_108839208	10310T>G	(T25)CCACCTGAGGAGGAGAATCCCCTGA	E6	3′-UTR	low	0.839
TLR5		16_27307966	305A>G	GCATGGTAAC TCGTGTACAC CATCAGA	F1	upstream	modifier	0.385
TLR5		16_27307783	488C>G	(T6)CCAGGGATGAAACCCRTGTCTCCTG	E2	upstream	modifier	0.383
TLR5	rs55617223	16_27307726	545C>T	(T10)CCAGGGAAGTCTTGCTGGCCTACTG	E3	upstream	modifier	0.407
TLR5	ss73689429	16_27307652	619T>G	CCACAGCACCTTTGAGGCTGTGAC	G1	upstream	modifier	0.853

Table 1. *Cont.*

Gene	SNP Identifier	Chromosomal Position	SNP [a]	SNaPshot Primer (Sequence 5′→3′)	Multiplex	SNP Type	Predicted Effect [b]	Variant Frequency
TLR5	ss73689443	16_27306535	1736C>T	(T2)GTACTTACAAYCATGCTTGCTATTTTT	G2	upstream	modifier	0.617
TLR5	rs55617187	16_27304557	3714T>C	(T15)GATTGAGCCAATGGATAAAAGCACT	E4	synonymous	low	0.481
TLR5	rs55617178	16_27304380	3891C>T	(T18)CACGAGGAACAGAGTCAAGGTGACAGT	E5	synonymous	low	0.907
TLR5	rs55617288	16_27303645	4626C>T	(T9)GGGGTCGCAAAGAGTAGGACATGACC	G3	downstream	modifier	0.690
TLR5		14_27302910	5144A>G	(T16)CGTTTCCAGAGGGGCTGGTCAGTG	H4	downstream	modifier	0.10
TLR6	rs43702941	6_59706074	855G>A	(T3)CCCAAATAGCTTTTTCTCTGTCCAAGTG	F2	Asp214Asn	moderate	0.714
TLR6		6_59706064	865G>C	(T10)GTCAGTTGTAAGCACSCTAAACTATTC	F3	Gly217Ala	moderate	0.316
TLR6		6_59705939	990G>A	(T15)CCTTACTAAATTTTACCCTCAACCAC	F4	Val259Met	moderate	0.484
TLR6	rs68268274	6_59705592	1337T>C	(T18)GATAAGTGTCTCCAATCTAGCTAAAGT	F5	Asp374Glu	moderate	0.333

[a] Positions of SNPs are determined in the reference sequences FJ147090, EU746465, AC000135.1, EU006635, and AJ618974 for *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6*, respectively. [b] The functional effect of the change predicted using VEP software. [c] LRR—leucine-rich repeat domains located in the extracellular part of receptors and essential for ligand binding.

Table 2. SNPs in directly read haplotypes in the proximal amplicon of *TLR2* in the bull population of the CRP cattle.

Gene	SNP Identifier	Chromosomal Position	SNP [a]	Multiplex	Effect Prediction	Variant Frequency
TLR2	rs68343162	17_3953930	115T>C	intron	low	0.163
TLR2	rs68268242	17_3953823	222A>T	intron	moderate	0.168
TLR2	rs68343164	17_3953633	412G>A	intron	low	0.159
TLR2	rs378996667	17_3953622	423G>A	intron	modifier	0.176
TLR2	rs68268244	17_3953586	459T>C	intron	modifier	0.150
TLR2	rs68343166	17_3953517	528G>A	intron	tolerated	0.173
TLR2	rs68268246	17_3953507	538A>C	intron	low	0.176
TLR2	rs68268247	17_3953421	624G>T	intron	moderate	0.19
TLR2	rs68268248	17_3953388	657C>T	intron	moderate	0.168

[a] Positions of SNPs are determined in the reference sequence EU746465.

Surprisingly, 15 haplotypes formed by the SNPs from the first amplicon of *TLR2* (amplicon located in the proximal part of the *TLR2* transcript) formed two distinct clusters (Figure 1A), which can be denoted as a bimodal distribution of frequencies.

3.2. Reconstructed Haplotypes

To verify this result in an independent way, the haplotypes in all five investigated genes were obtained using statistical reconstruction. First, polymorphisms were genotyped in 164 bulls with SNaPshot (primer extension) assays, according to Table 1. The allele frequencies found in the population are also summarized in Table 1. The variants present in the *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6* genes in the individual animals were processed using the PHASE program to determine the conceivable haplotypes and their assumed frequencies (Supplementary Table S3). The group of the most realistic haplotypes was selected using the thresholds in frequency (>0.5 predicted copy/population) and the standard error/frequency ratio (<1), as well as the drop in the calculated frequency distribution plot. The drop in the predicted haplotype frequency graph is supposed to indicate the boundary between the real haplotypes and the artifacts of statistical reconstruction or haplotypes arising from genotyping errors.

The filtered haplotypes were used as an input in the reconstruction of the population structure in the Network program. In the case of all 10 SNPs in *TLR2* used to reconstruct the haplotypes, the observed pattern (Figure 1B) matched the pattern observed in directly read haplotypes in Figure 1A. Two main nodes of two clusters were separated by five resulting substitutions (in fact, by nine substitutions, including reversions) and were radially surrounded by minor haplotypes. This pattern became more distinct when only six SNPs proximal to the 5′ end of *TLR2* were used for statistical haplotype reconstruction (Supplementary Table S3) and subsequently involved in graph building (Figure 1C). This variant corresponds even more to the bimodal distribution of haplotypes observed in the direct reading of *TLR2*, as shown in Figure 1A, than the graph employing all SNPs in Figure 2B.

Similar to the case of *TLR2*, the bimodal distribution was detectable in the statistically reconstructed haplotypes of *TLR5* (Figure 2A), *TLR4* (Figure 2B), and *TLR6* (Figure 2C) after graphical clustering. Analogously to *TLR2*, the bimodal distribution in *TLR5* was clearer in the 5′-proximal region of the transcript (not shown). In *TLR1*, an analogical structure was not found, since the low number of three SNPs that were present in the population of Czech Red Pied cattle created only rudimentary haplotypes forming a simple tree graph presented in Figure 2D.

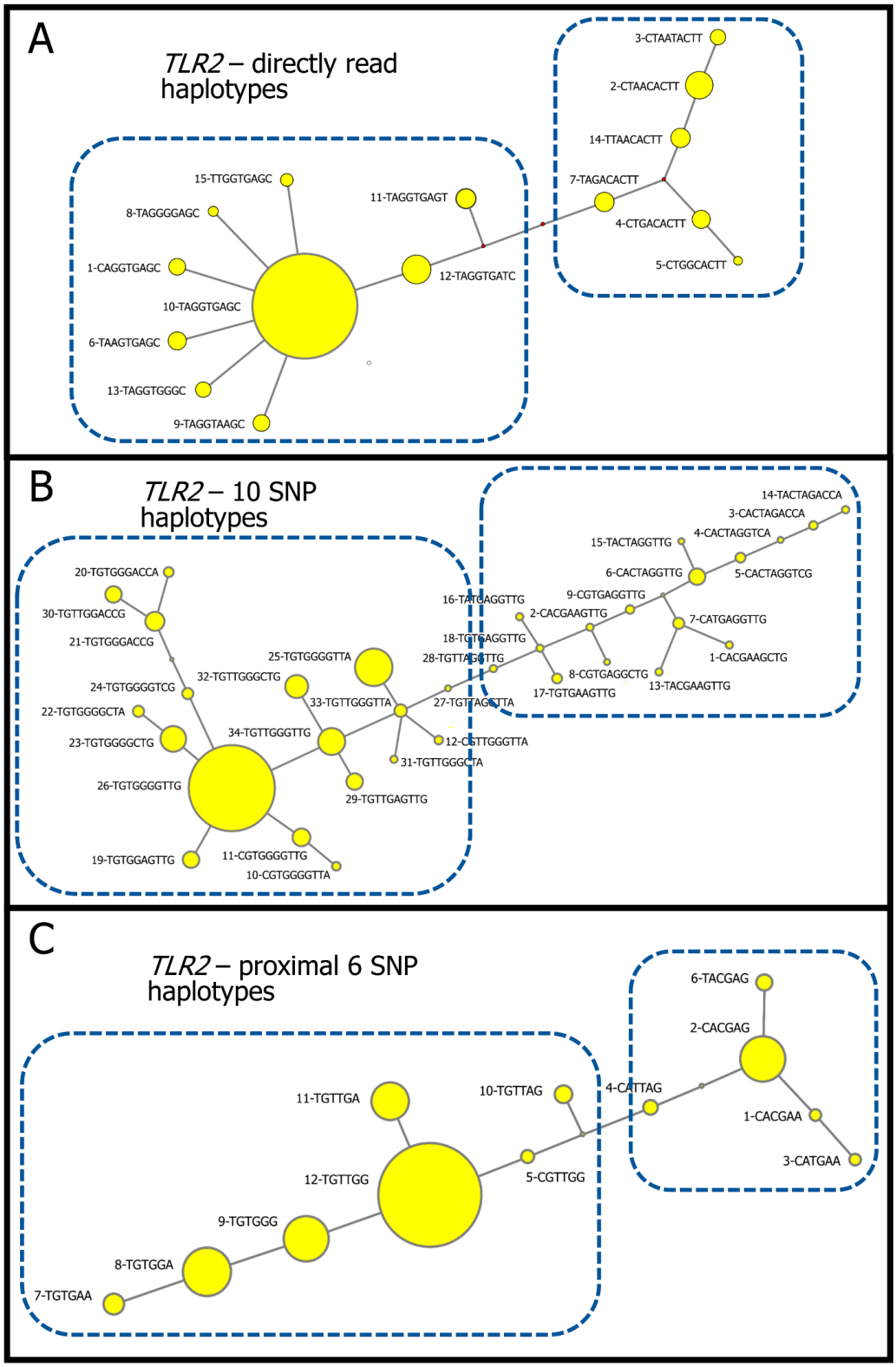

Figure 1. Distribution of directly determined vs. statistically reconstructed haplotypes in *TLR2*. (**A**)—network tree graph for directly read haplotypes in proximal amplicon 1 of *TLR2* (comprising 9 SNPs), (**B**)—network tree graph for statistically reconstructed haplotypes (from 10 SNPs), and (**C**)—network tree graph for statistically reconstructed haplotypes (built on 5′-proximal 6 SNPs). The node size represents the haplotype frequency.

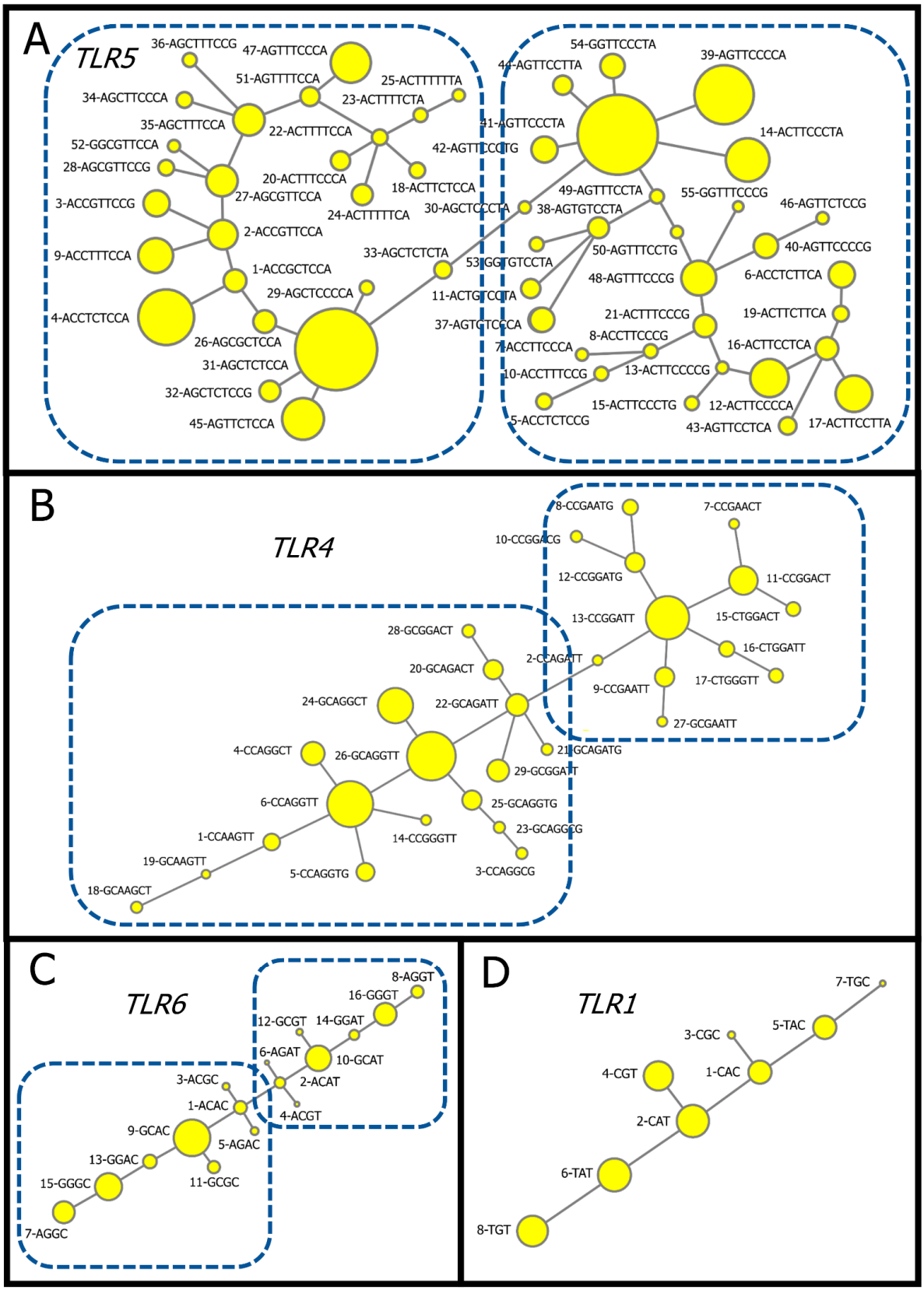

Figure 2. Network tree graph for statistically reconstructed haplotypes in additional *TLR*s. (**A**)—*TLR5* (haplotypes built from 9 SNPs); (**B**)—*TLR4* (from 7 SNPs); (**C**)—*TLR6* (from 4 SNPs); and (**D**)—*TLR1* (from 3 SNPs). The node size represents the haplotype frequency.

4. Discussion

4.1. Context of the Previous Knowledge

Bimodal clustering of the SNP haplotypes in the antibacterial genes *TLR2*, *TLR4*, *TLR5*, and *TLR6* was independently detected in the population of Czech Red Pied cattle. This trend was expressed mainly in *TLR2* and to different extents in the remaining genes. The low diversity found in *TLR1* did not allow us to confirm this phenomenon for this gene. Direct haplotype reading from long sequencing reads, and haplotype statistical reconstruction from the genotyping results led to a consistent model in *TLR2*.

A match with the previously noted clustering of bovine TLR haplotypes in unrelated cattle breeds can be determined. This phenomenon was first reported and analyzed in a panel of cattle breeds, namely, for the *TLR3* and *TLR8* genes [13]. Subsequently, haplotype clustering was reported for *TLR2* and as a tendency for *TLR4* in Turkish historical cattle breeds and the Holstein population in Turkey by Bilgen et al. [18]. Despite thorough statistical processing in the work by Fisher et al. [13], the origin of this disequilibrium is still unexplained. The unrelatedness of the breed sets studied to date, present work [13,18] indicates that haplotype clustering is neither of a common historical origin nor associated with current introgression into modern breeds. Nevertheless, balancing selection was suggested as a reason for the observed disequilibrium by Fisher et al. [13].

It must be mentioned that balancing selection is considered a reason for polymorphisms in the *TLR2* gene in other species, namely in species that serve as models in population genetics. This might be exemplified in the bank vole (*Myodes glareolus*), where balancing selection has been demonstrated by Kloch et al. [31]. One of the three *TLR2* haplotype clusters observed in this model was assigned to resistance against the bacterium *Borrelia afzelii* [32]. Using the polymorphism pattern distribution across Europe, the stable polymorphism in this species was dated back before the split of the mitochondrial lineages 0.19–0.56 Mya [33]. In contrast, no consequences of supposedly balancing selection were observed in a related species, yellow-necked mice. This difference was assigned to different strategies to combat *Borrelia* infections in these two rodent species [34].

4.2. Possible Reasons for the Observed Haplotype Clustering

Therefore, the yet unanswered question is how the presence of clusters of haplotypes in the *TLR* genes of cattle, observed independently in several breeds from different geographical locations, is stably supported in the world population.

The existence of balancing selection has been postulated by Fisher et al. [13] for *TLR3* and *TLR8*. However, in the simplest case of the heterozygote advantage, the mechanism of recombination should lead to merging both haplotype groups with time.

On the other hand, fluctuating selection by alternating infectious agents has been considered by Bilgen et al. [18] to explain this *TLR* haplotype disequilibrium in Turkish cattle populations. Analogically, in roe deer (*Capreolus capreolus*), the antagonistic selection exerted by *Toxoplasma* and *Chlamydia* infections alternating across years and landscape features was suggested as a reason for a stable *TLR2* polymorphism [35]. Consistently, only limited support for the *TLR* heterozygote advantage was found in the studied set of cervid *TLR2*, *TLR4*, and *TLR5*, i.e., for otherwise the most common reason of balancing selection leading to polymorphism. Moreover, the case of balancing selection in *TLR*s in rodents, declared to be caused by balancing selection presumably due to either frequency-dependent selection or a rare allele advantage during *Borrelia* infection [31], might also be the result of alternating infections regularly occurring in natural habitats [32].

However, one of the haplotype groups should be, sooner or later, lost in this fluctuating selection model either by the temporarily prevailing unilateral infection pressure or as a result of stochastic processes leading to gene erosion in limited populations. To preserve distinct haplotype groups over a long time period on the worldwide cattle population scale, the mechanism must be more stable and independent of the unpredictable timing of the waves of epizooties.

The alternating infection pressure might also be interpreted as infections occurring during the life cycle of animals and representing, for instance, calfhood diseases and adult animal threats. This might be a stable factor of frequency-dependent balancing selection favoring certain groups of haplotypes.

4.3. A Dual Function Hypothesis

However, the most likely explanation is that the balancing selection originates from two different essential functions performed by *TLR* genes or their products simultaneously. The haplotype groups might then represent protein isoforms adapted to one of these functions. An example of a dual function might be the formation of two kinds of products differing in ligand specificity, such as the interactions of the *TLR2* product with TLR1 and TLR6. This leads to the formation of two functional dimers with different recognized ligands [36]. While the TLR2/TLR1 protein dimer recognizes triacylated lipoproteins, TLR2/TLR6 detects diacylated lipoproteins, both of bacterial origin. The insufficiency in one of these two functions, provided by a specialized group of haplotypes, might lead to frequency-dependent selection in favor of the depleted haplotype group. The discrimination of two bacterial antigen groups might also be mediated by the TLR2/TLR10 heterodimers that have been documented [37], although their function is not clear until now.

This model matches the permanent selection pressures performed by two groups of pathogens differing in surface antigens. In particular, mycobacterial infections and borreliosis are caused by bacteria forming triacylated lipoproteins that are recognized by the heterodimer TLR2/TLR1 [38]. On the other hand, the diacylated lipoprotein of the mastitis causal agent, *Staphylococcus aureus*, can be recognized by the heterodimer TLR2/TLR6. Nevertheless, the combination of some conditions, such as acidic pH and a post-logarithmic growth phase, is required for the expression of this trait. Also, high temperatures and high salt concentrations additively accelerated the accumulation of the diacyl lipoprotein form of *Staphylococcus* [39].

While mycobacterial infections are considered to be ubiquitous, the role of borreliosis is not so perceived, although it is also a factor acting in cattle almost worldwide [40]. However, *Borrelia afzelii* is still considered to be a causal agent of the *TLR2* haplotype disequilibrium in rodents [32]. This hypothesis does not comply with the fact that none of the polymorphic sites are located directly in the TLR1–TLR2 interface in the bank vole, as determined by molecular modeling [32].

In cattle, some of the SNPs known for *TLR2* were linked to the recognition of the considered bacterial groups: 1313G>A (Arg152Gln) was reported to increase susceptibility to bovine tuberculosis [17], while 385T>G was reported to affect mastitis incidence in Holstein, Simmental, and Sanhe cattle [15]. Unfortunately, no correlation of the former polymorphism with the reconstructed haplotype groups can be seen in our work, as documented in Supplementary Table S3 and in Figure 2A, where 1313G>A is represented by the sixth nucleotide in the haplotype sequence. The latter polymorphism, 385T>G, was not found in the studied population of CRP cattle at all. Therefore, the correlation of these SNPs with haplotype groups could not be used to corroborate the dual infection pressure hypothesis as a reason for the observed balancing selection.

Two factors exerting the selective advantage of the individual groups of haplotypes might be other pairs of ligands or proteins interacting with toll-like receptors. For example, exogenous vs. endogenous groups of TLR ligands of different functions might preferentially direct the selection towards different haplotype groups. Exogenous and endogenous ligands have been associated with the nonimmune functions of TLRs [41]. Multiple roles of TLRs have been demonstrated, including a role in female reproductive functions, as reported for mice as a model species and strongly evidenced with the TLR4-deficient genotype [7]. The role of TLR4 in perinatal signaling is so strong that it is used for the development of drugs reducing perinatal risk in human medicine [42]. If both the immune and nonimmune signaling functions are essential, none of the specialized haplotype groups will be lost.

Additionally, non-TLR protein interactors with TLR molecules in the standard signaling pathway might counteract the loss of specialized haplotype clusters. These factors might include myeloid differentiation protein-2 (MD-2), which forms a functional complex with TLR4, and bacterial lipopolysaccharide [43].

Another line of evidence for the historical selection of *TLR* gene variants can be obtained from the study of conserved breeds in the programs of genetic resources. For example, elevated frequency of *TLR1*, *-2*, and *-4* alleles associated with anti-mycobacterial resistance was found in the historical Marchigiana breed [44].

4.4. Approaches to the Discrimination of the Causative Mechanisms

Since the better expression of bimodal clustering in the 5′-proximal SNPs supports the localization of the selection targets in the upstream regulatory regions or the functional interactions in the proximal part of the transcripts, additional analysis of the extragenic regions might represent an approach to solving the existing questions. A low-error NGS over newly designed large amplicons would help produce more comprehensive data.

An insight into the stable divergence of the haplotypes in the immune receptor genes might be provided by research on the bimodal distribution of haplotypes in the genes for killer-cell immunoglobin-like receptors (KIRs), as reported by Roe et al. [45].

The observed disequilibrium in the *TLR* haplotypes should also be interpreted in the context of the haplotype-resolved genomes being constructed for several cattle populations. The first case was reported by Low et al. [46] (2021) for the Holstein cattle breed. Also, the in silico modeling of interactions of the TLR protein variants with known ligands and adaptors might provide additional evidence for the dual function hypothesis.

5. Conclusions

The formation of two clusters of haplotypes in the genes for toll-like receptors, namely, *TLR2*, *TLR4*, *TLR5*, and *TLR6*, was observed in the population of Czech Red Pied cattle. Since this finding is consistent with previous reports in other cattle breeds, the phenomenon seems to be general. To explain the stability of haplotype disequilibrium, simultaneous adaptation to two different functions can be considered. In the particular case of *TLR2*, it might be the fitting of the SNP series to the interactions in the functional heterodimers formed with the TLR1 and TLR6 proteins. Additional evidence for this hypothesis can be provided by extending the haplotypes over the adjacent regulatory regions and by interaction studies of the TLR protein variants. Both the surveyed haplotypic variations and the anticipated phenotypic effects indicate that the present findings should be considered in breeding for the health and reproductive traits of cattle.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/d15070811/s1. Supplementary Table S1: Reference sequences and amplification primers for bovine *TLR*s; Supplementary Table S2: Haplotypes of *TLR2* directly determined by PacBio resequencing; and Supplementary Table S3: Reconstructed haplotypes for the *TLR1*, *-2*, *-4*, *-5*, and *-6* genes.

Author Contributions: Conceptualization, K.N. and K.S.; methodology, K.N.; formal analysis, K.N. and K.S.; investigation, K.S. and K.N.; resources, K.N.; writing—original draft preparation, K.N. and K.S.; writing—review and editing, K.S. and K.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Long-Term Concept of the Research Institution Development MZE-RO0723 of the Ministry of Agriculture of the Czech Republic and the National Agency of Agricultural Research of CR project No. QJ1610489.

Institutional Review Board Statement: Ethical review and approval were waived for this study since the DNA samples used were isolated from the archived commercial insemination doses of the involved bull population. The research project, in general, was supervised by the Ethics Committee of the Institute of Animal Science.

Data Availability Statement: The primary sequencing and genotyping data are archived by the authors of the study. Data are available on request at the breeding firm CHD Impuls (Bohdalec, Czech Republic). The Czech Red Pied cattle population data are stored at the Czech-Moravian Breeders' Corporation, Hradištko, Czech Republic.

Acknowledgments: The authors thank Marek Bjelka and the breeding company CHD Impuls for providing access to the archive of the Czech Red Pied cattle.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Kawai, T.; Akira, S. The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors. *Nat. Immunol.* **2010**, *11*, 373–384. [CrossRef] [PubMed]
2. Stein, D.; Roth, S.; Vogelsang, E.; Nüsslein-Volhard, C. The polarity of the dorsoventral axis in the *Drosophila* embryo is defined by an extracellular signal. *Cell* **1991**, *65*, 725–735. [CrossRef] [PubMed]
3. Lemaitre, B.; Nicolas, E.; Michaut, L.; Reichhart, J.M.; Hoffman, J.A. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **1996**, *86*, 973–983. [CrossRef] [PubMed]
4. Wiens, M.; Korzhev, M.; Perovic-Ottstadt, S.; Luthringer, B.; Brandt, D.; Klein, S.; Mueller, W.E.G. Toll-like receptors are part of the innate immune defense system of sponges (*Demospongiae*: *Porifera*). *Mol. Biol. Evol.* **2007**, *24*, 792–804. [CrossRef]
5. Zipfel, C.; Robatzek, S.; Navarro, L.; Oakeley, E.J.; Jones, J.D.G.; Felix, G.; Boller, T. Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **2004**, *428*, 764–767. [CrossRef]
6. Kannaki, T.R.; Shanmugam, M.; Verma, P.C. Toll-like receptors and their role in animal reproduction. *Anim. Reprod. Sci.* **2011**, *125*, 1–12. [CrossRef]
7. Wahid, H.H.; Dorian, C.L.; Chin, P.Y.; Hutchinson, M.R.; Rice, K.C.; Olson, D.M.; Moldenhauer, L.M.; Robertson, S.A. Toll-like receptor 4 is an essential upstream regulator of on-time parturition and perinatal viability in mice. *Endocrinology* **2015**, *156*, 3828–3841. [CrossRef]
8. Boichard, D.; Ducrocq, V.; Fritz, S. Sustainable dairy cattle selection in the genomic era. *J. Anim. Breed. Genet.* **2015**, *132*, 135–143. [CrossRef]
9. Maurić Maljković, M.; Vlahek, I.; Piplica, A.; Ekert Kabalin, A.; Sušić, V.; Stevanović, V. Prospects of toll-like receptors in dairy cattle breeding. *Anim. Genet.* **2023**, *2023*, 1–10. [CrossRef]
10. Jungi, T.W.; Farhat, K.; Burgener, I.A.; Werling, D. Toll-like receptors in domestic animals. *Cell Tissue Res.* **2011**, *343*, 107–120. [CrossRef]
11. Jann, O.C.; Werling, D.; Chang, J.S.; Haig, D.; Glass, E.J. Molecular evolution of bovine Toll-like receptor 2 suggests substitutions of functional relevance. *BMC Evol. Biol.* **2008**, *8*, 288. [CrossRef]
12. Seabury, C.M.; Seabury, P.M.; Decker, J.E.; Schnabel, R.D.; Taylor, J.F.; Womack, J.E. Diversity and evolution of 11 innate immune genes in *Bos taurus taurus* and *Bos taurus indicus* cattle. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 151–156. [CrossRef]
13. Fisher, C.A.; Bhattarai, E.K.; Osterstock, J.B.; Dowd, S.E.; Seabury, P.M.; Vikram, M.; Whitlock, R.H.; Schukken, Y.H.; Schnabel, R.D.; Taylor, J.F.; et al. Evolution of the bovine *TLR* gene family and member associations with *Mycobacterium avium* subspecies *paratuberculosis* infection. *PLoS ONE* **2011**, *6*, 11. [CrossRef]
14. Sharma, B.S.; Leyva, I.; Schenkel, F.; Karrow, N.A. Association of Toll-Like receptor 4 polymorphisms with somatic cell score and lactation persistency in Holstein bulls. *Dairy Sci.* **2006**, *89*, 3626–3635. [CrossRef]
15. Zhang, L.P.; Gan, Q.F.; Ma, T.H.; Li, H.D.; Wang, X.P.; Li, J.Y.; Gao, X.; Chen, J.B.; Ren, H.Y.; Xu, S.Z. Toll-like receptor 2 gene polymorphism and its relationship with SCS in dairy cattle. *Anim. Biotech.* **2009**, *20*, 87–95. [CrossRef]
16. Koets, A.; Santema, W.; Mertens, H.; Oostenrijk, D.; Keestra, M.; Overdijk, M.; Labouriau, R.; Franken, P.; Frijters, A.; Nielen, M.; et al. Susceptibility to paratuberculosis infection in cattle is associated with single nucleotide polymorphisms in Toll-like receptor 2 which modulate immune responses against *Mycobacterium avium* subspecies *paratuberculosis*. *Prev. Vet. Med.* **2010**, *93*, 305–315. [CrossRef]
17. Bhaladhare, A.; Sharma, D.; Kumar, A.; Sonwane, A.; Chauhan, A.; Singh, R.; Kumar, P.; Yadav, R.; Baqir, M.; Bhushan, B.; et al. Single nucleotide polymorphisms in toll-like receptor genes and case-control association studies with bovine tuberculosis. *Vet. World* **2016**, *9*, 458–464. [CrossRef]
18. Bilgen, N.; Kul, B.C.; Offord, V.; Werling, D.; Ertugrul, O. Determination of genetic variations of Toll-like receptor (TLR) 2, 4, and 6 with next-generation sequencing in native cattle breeds of Anatolia and Holstein Friesian. *Diversity* **2016**, *8*, 23. [CrossRef]
19. Ruiz-Larrañaga, O.; Manzano, C.; Iriondo, M.; Garrido, J.M.; Molina, E.; Vazquez, P.; Juste, R.A.; Estonba, A. Genetic variation of toll-like receptor genes and infection by *Mycobacterium avium* ssp. *paratuberculosis* in Holstein-Friesian cattle. *J. Dairy Sci.* **2011**, *94*, 3635–3641. [CrossRef]
20. Beecher, C.; Daly, M.; Childs, S.; Berry, D.P.; Magee, D.A.; McCarthy, T.V.; Giblin, L. Polymorphisms in bovine immune genes and their associations with somatic cell count and milk production in dairy cattle. *BMC Genet.* **2010**, *11*, 99. [CrossRef]
21. Abdel-Shafy, H.; Bortfeldt, R.H.; Tetens, J.; Brockmann, G.A. Single nucleotide polymorphism and haplotype effects associated with somatic cell score in German Holstein cattle. *Genet. Sel. Evol.* **2014**, *46*, 35. [CrossRef] [PubMed]

22. Novák, K.; Bjelka, M.; Samake, K.; Valčíková, T. Potential of *TLR*-gene diversity in Czech indigenous cattle for resistance breeding as revealed by hybrid sequencing. *Arch. Anim. Breed.* **2019**, *62*, 477–490. [CrossRef] [PubMed]
23. White, S.N.; Taylor, K.H.; Abbey, C.A.; Gill, C.A.; Womack, J.E. Haplotype variation in bovine Toll-like receptor 4 and computational prediction of a positively selected ligand-binding domain. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10364–10369. [CrossRef] [PubMed]
24. Seabury, C.M.; Cargill, E.J.; Womack, J.E. Sequence variability and protein domain architectures for bovine Toll-like receptors 1, 5, and 10. *Genomics* **2007**, *90*, 502–515. [CrossRef] [PubMed]
25. Sim, N.L.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucl. Acids Res.* **2012**, *40*, W452–W457. [CrossRef]
26. Hoogendoorn, B.; Owen, M.J.; Oefner, P.J.; Williams, N.; Austin, J.; O'Donovan, M.C. Genotyping single nucleotide polymorphisms by primer extension and high performance liquid chromatography. *Human Genet.* **1999**, *104*, 89–93. [CrossRef]
27. Bjelka, M.; Novák, K. Association of TLR gene variants in a Czech Red Pied cattle population with reproductive traits. *Vet. Immunol. Immunopathol.* **2020**, *220*, 109997. [CrossRef]
28. Stephens, M.; Smith, N.J.; Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **2001**, *68*, 978–989. [CrossRef]
29. Bandelt, H.-J.; Forster, P.; Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
30. Polzin, T.; Daneschmand, S.V. On Steiner trees and minimum spanning trees in hypergraphs. *Oper. Res. Lett.* **2003**, *31*, 12–20. [CrossRef]
31. Kloch, A.; Wenzel, M.A.; Laetsch, D.R.; Michalski, O.; Bajer, A.; Behnke, J.M.; Welc-Falęciak, R.; Piertney, S.B. Signatures of balancing selection in toll-like receptor (TLRs) genes—Novel insights from a free-living rodent. *Sci. Rep.* **2018**, *8*, 8361. [CrossRef]
32. Tschirren, B.; Andersson, M.; Scherman, K.; Westerdahl, H.; Mittl, P.R.E.; Råberg, L. Polymorphisms at the innate immune receptor TLR2 are associated with *Borrelia* infection in a wild rodent population. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 20130364. [CrossRef]
33. Morger, J.; Råberg, L.; Hille, S.M.; Helsen, S.; Štefka, J.; Al-Sabi, M.M.; Kapel, C.M.O.; Mappes, T.; Essbauer, S.; Ulrich, R.G.; et al. Distinct haplotype structure at the innate immune receptor Toll-like receptor 2 across bank vole populations and lineages in Europe. *Biol. J. Linn. Soc.* **2015**, *116*, 124–133. [CrossRef]
34. Zhong, X.; Lundberg, M.; Råberg, L. Divergence in coding sequence and expression of different functional categories of immune genes between two wild rodent species. *Genome Biol. Evol.* **2021**, *13*, evab023. [CrossRef]
35. Quéméré, E.; Hessenauer, P.; Galan, M.; Fernandez, M.; Merlet, J.; Chaval, Y.; Morellet, N.; Verheyden, H.; Gilot-Fromont, E.; Charbonnel, N. Pathogen-mediated selection favours the maintenance of innate immunity gene polymorphism in a widespread wild ungulate. *J. Evol. Biol.* **2021**, *34*, 1156–1166. [CrossRef]
36. Yang, L.; Seki, E. Toll-like receptors in liver fibrosis: Cellular crosstalk and mechanisms. *Front. Physiol.* **2012**, *3*, 138. [CrossRef]
37. Regan, T.; Nally, K.; Carmody, R.; Houston, A.; Shanahan, F.; MacSharry, J.; Brint, E. Identification of *TLR10* as a key mediator of the inflammatory response to *Listeria monocytogenes* in intestinal epithelial cells and macrophages. *J. Immunol.* **2013**, *191*, 6084–6092. [CrossRef]
38. Coffey, T.J.; Werling, D. Therapeutic targeting of the innate immune system in domestic animals. *Cell Tissue Res.* **2011**, *343*, 251–261. [CrossRef]
39. Kurokawa, K.; Kim, M.-S.; Ichikawa, R.; Ryu, K.-H.; Dohmae, N.; Nakayama, H.; Lee, B.L. Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. *J. Bacteriol.* **2012**, *194*, 3299–3306. [CrossRef]
40. Lippi, C.A.; Ryan, S.J.; White, A.L.; Gaff, H.D.; Carlson, C.J. Trends and opportunities in tick-borne disease geography. *J. Med. Entomol.* **2021**, *58*, 2021–2029. [CrossRef]
41. Song, Y.; Shou, L.M.; Ai, L.-Y.; Bei, Y.; Chen, M.-T. Mini-review: The non-immune functions of Toll-like receptors. *Crit. Rev. Eukaryot. Gene Expr.* **2019**, *29*, 37–45. [CrossRef] [PubMed]
42. Kadam, L.; Gomez-Lopez, N.; Mial, T.N.; Kohan-Ghadr, H.R.; Drewlo, S. Rosiglitazone regulates *TLR4* and rescues *HO-1* and *NRF2* expression in myometrial and decidual macrophages in inflammation-induced preterm birth. *Reprod. Sci.* **2017**, *24*, 1590–1599. [CrossRef] [PubMed]
43. Francisco, S.; Billod, J.-M.; Merino, J.; Punzón, C.; Gallego, A.; Arranz, A.; Martin-Santamaria, S.; Fresno, M. Induction of TLR4/TLR2 interaction and heterodimer formation by low endotoxic atypical LPS. *Front. Immunol.* **2022**, *12*, 748303. [CrossRef] [PubMed]
44. Mazzone, P.; Di Paolo, A.; Petrucci, L.; Torricelli, M.; Corneli, S.; Sebastiani, C.; Ciullo, M.; Sebastianelli, M.; Costarelli, S.; Scoccia, E.; et al. Evaluation of single nucleotide polymorphisms (SNPs) associated with genetic resistance to bovine paratuberculosis in Marchigiana beef cattle, an Italian native breed. *Animals* **2023**, *13*, 587. [CrossRef]

45. Roe, D.; Vierra-Green, C.; Pyo, C.-W.; Eng, K.; Hall, R.; Kuang, R.; Spellman, S.; Ranade, S.; Geraghty, D.E.; Maiers, M. Revealing complete complex *KIR* haplotypes phased by long-read sequencing technology. *Genes Immun.* **2017**, *18*, 127–134. [CrossRef]
46. Low, W.Y.; Tearle, R.; Liu, R.; Koren, S.; Rhie, A.; Bickhart, D.M.; Rosen, B.D.; Kronenberg, Z.N.; Kingan, S.B.; Tseng, E.; et al. Haplotype-resolved genomes provide insights into structural variation and gene content in Angus and Brahman cattle. *Nat. Commun.* **2021**, *11*, 2071. [CrossRef]