Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

Article

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

Nuket Bilgen [1,*], Bengi Cinar Kul [1], Victoria Offord [2], Dirk Werling [2] and Okan Ertugrul [1]

[1] Department of Animal Genetics, Faculty of Veterinary Medicine, University of Ankara, Ankara 06560, Turkey; bkul@ankara.edu.tr (B.C.K.); oertugrul@ankara.edu.tr (O.E.)
[2] Department of Pathology and Infectious Diseases, Royal Veterinary College, University of London, London WC1E 7HU, UK; vofford@rvc.ac.uk (V.O.); dwerling@rvc.ac.uk (D.W.);
* Correspondence: nbilgen@ankara.edu.tr; Tel.: +90-312-317-0315 (ext: 4327)

Work was carried out in fulfilment of the requirements for a Doctoral thesis, The Graduate School of Health Sciences of Ankara University, Ankara, Turkey

Abstract: In recent years, the focus of disease resistance and susceptibility studies in cattle have been on determining patterns in the innate immune response of key proteins, such as Toll-like receptors (TLR). In the bovine genome, there are 10 TLR family members and, of these, *TLR2*, *TLR4* and *TLR6* are specialized in recognition of bacterial ligands. Indigenous cattle breeds of Anatolia have been reported to show fewer signs of clinical bacterial infections, such as bovine tuberculosis and mastitis, and it is hypothesized that this might be due to a less stringent genetic selection during breeding. In contrast, Holstein-Friesian cattle have been under strong selection for milk production, which may have resulted in greater susceptibility to diseases. To test this hypothesis, we have compared the *TLR2*, *TLR4* and *TLR6* genes of Anatolian Black (AB), East Anatolian Red (EAR), South Anatolian Red (SAR), Turkish Grey (TG), and Holstein (HOL) cattle using Next Generation Sequencing. The SAR breed had the most variations overall, followed by EAR, AB, TG and HOL. TG had the most variations for *TLR2* whereas SAR had the most variations in *TLR4* and *TLR6*. We compared these variants with those associated with disease and susceptibility traits. We used exon variants to construct haplotypes, investigated shared haplotypes within breeds and determined candidate haplotypes for disease resistance phenotype in Anatolian cattle breeds.

Keywords: Anatolian Black; East Anatolian Red; South Anatolian Red; Turkish Grey; Holstein Friesian; Innate immunity; Next Generation Sequencing; TLR2; TLR4; TLR6

---

1. Introduction

Toll-like receptor (TLR) recognize conserved patterns in diverse microbial molecules called microbial-associated molecular patterns (MAMPs). These include lipopolysaccharide recognized by TLR4 and lipopeptides recognized by the heterodimer formed by TLR2 with either TLR1 or TLR6 [1].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

In addition, TLRs also react to damage associated molecular patterns (DAMPs) [2,3] released after cellular damage, and thus play a crucial role in initiating the innate immune response [4]. Genetic variations found in the genes encoding TLRs have been associated with disease susceptibility and resistance in a variety of animal species [1,5,6]. In the bovine genome, 10 members of the TLR family (TLR1–10) have been identified and mapped to specific chromosomes [7]. Bovine *TLR2*, *TLR4* and *TLR6* genes are located on BTA17, BTA8 and BTA6 respectively [8,9]. *TLR2* has two exons and the protein transcribed from the reverse strand of the larger exon is 784 amino acids (aa) long. The *TLR6* gene has four exons, and this gene is also transcribed from the reverse strand of two large exons, resulting in a protein of 793aa. *TLR4* has three transcribed exons, resulting in a protein of 841aa. Independent of their aa length, all TLRs identified so far contain three domains: an extra cellular ligand binding domain (ECD) consisting of various numbers of leucine-rich repeats (LRR), a trans-membrane domain and an intra-cellular domain, which is also known as Toll/Interleukin-1 Receptor (TIR) Domain [10].

Several studies have suggested that disease resistance, susceptibility and severity of clinical signs in certain individuals or breeds may be affected by altered ligand binding caused by single nucleotide polymorphisms (SNPs) within TLR genes [11,12]. Given the increased use of antibiotics in food-producing animals, native cattle breeds could provide an important genetic resource to breeders. Many of these indigenous breeds show an enhanced disease resistance due to a co-evolution with specific pathogens over decades, potentially resulting in the development of genetic resistance. Indeed, due to low selective pressure, local breeds are generally not as productive as high-yielding breeds selected for specific QTLs, and thus may have preserved their genetic makeup over the years. Anatolia has five native cattle breeds; Anatolian Black (AB), South Anatolian Red (SAR), East Anatolian Red (EAR), Turkish Grey (TG) and Native Southern Yellow. In general, these breeds grow slowly, are adapted to extensive breeding conditions and present the morphologically look of semi-wild cattle [13]. Milk production in these breeds is between 700–1000 kg per lactation period per cow. The breeds are adapted to challenging environmental conditions and are known to be resistant to specific diseases. For example, SAR and EAR breeds are known to be resistant to blood parasite infections [14], whereas cows from the AB breed showed an extremely low bovine Tuberculosis incidence rate when kept with *Mycobacterium bovis* infected cattle [15] as well as a low mastitis-incidence rate [15]. As the corresponding disease-causing bacteria have been described to bind to *TLR2*, *TLR4* and *TLR6*, the aim of the study was to determine the presence of SNP in these genes in the AB, EAR, SAR and TG indigenous breeds compared to those present in the corresponding genes of HOL cattle.

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

2. Materials and Methods

2.1.DNA isolation and amplicon sequencing

To determine variations in the genomic sequences for *TLR2*, *TLR4* and *TLR6*, blood samples from AB (*n* = 20), EAR (*n* = 20), SAR (*n* = 20), TG (*n* = 10) and HOL (*n* = 10) were used. Blood samples were collected from each breed across Turkey (Figure 1).

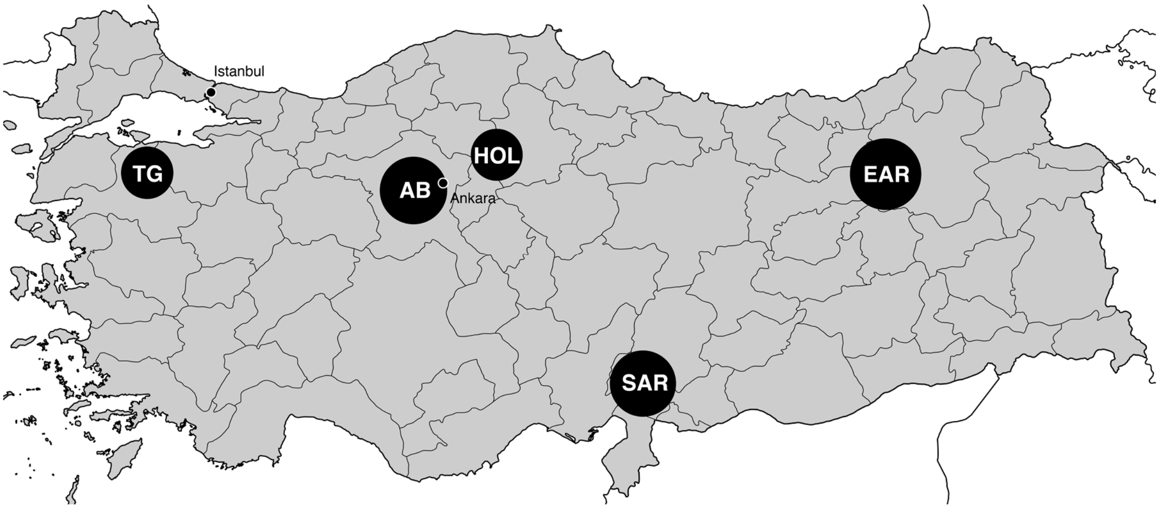

Figure 1. Geographical regions of the sampled breeds: Anatolian Black (AB), East Anatolian Red (EAR), South Anatolian Red (SAR), Turkish Grey (TG), and Holstein (HOL). Size of circle denotes relative population sizes.

Genomic DNA (gDNA) extraction performed using a commercially available gDNA extraction kit (Qiagen Blood and Tissue, Germany). Extracted DNA was measured by spectrophotometry (Nanodrop 2000, Thermo) and integrity visualised by agarose gel electrophoresis. To obtain reliable variant information, sequencing was performed at 50x coverage to analyze *TLR2*, *TLR4*, and *TLR6*, which spanned 13, 11 and 19 kb regions, respectively. Different primer pair combinations were used for each gene (for details see: Supplementary Table 1). Primers were designed using the Primer3 software package [16] and were spaced out over 2500–3500 bp intervals. Hot start taq DNA polymerase (Phire Hot Start II, Thermo Fisher, Germany) was used in PCR applications. PCR conditions were as followed; 95°C 60 sec, for 45 cycles; 95°C 10 sec, 60°C 120 sec, 72°C 20 sec and final elongation at 72°C 60 sec. For GC rich regions 5% DMSO was added to the reaction to enhance PCR results (for details see: Supplementary Table 2). PCR amplicons were visualized using SybrSafe (Invitrogen, UK) stained agarose gels. NGS library preparation was performed following the manufacturer's instructions (Nextera Library Preparation Kit, Illumina inc.). Amplicons were sequenced using paired end sequencing on the MiSeq platfrom (Illumina in Intergen Lab Inc.; Ankara, Turkey), and reads were aligned to the present bovine genome version available Btau 4.6.1 (Btau7) using the MiSeq software. The resulting alignment was used to construct binary aligned map (.bam) files.

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

2.2. Sequence Analysis, variant verification and Protein modelling

TLR2, *TLR4* and *TLR6* sequence data sets from individuals were analysed using Picard [17], BAM tools, SAM tools [18], GATK [19], SnpSift [20] and SnpEff analysis tools [21] to generate Variation Call Files (VCF) [22] for each individual. Thereafter, VCFs were aggregated at breed level to determine novel SNP and InDel variants, and the results were stored in distinct .vcf files. Positions of both SNPs and InDels were lifted over to a newer assembly version (Btau8) in UCSC [8] and new positions were uploaded to Ensembl Variant Effect Predictor (VEP) [23] database. The accuracy and the annotation of the identified genetic variations were assessed using VEP [23] and SnpSift [20], respectively.

Although VEP and SnpSift are useful tools to assess accuracy and to annotate novel candidate SNPs, we also screened variants using their coverage and frequency to eliminate false variant calls. Potentially important non-synonymous and novel SNP variants located in exons were validated subsequently by using Sanger sequencing. New sets of primers covering the regions including variants were designed by using Primer3 [16] (Supplementary Table 1). The PCR products were sequenced using BigDye Terminator v3.1 cycle sequencing kit and ABI310 automatic sequencer (Applied Biosystems, Foster City, CA, USA). Amplification primers were used for bidirectional sequencing. Obtained data was analyzed using BioEdit software [24].

After analyzing the variants at the amino acid level, a protein model was constructed for the most important variant identified in TLR2. Models were constructed using the Modeller [25] software package, and validated using ProCheck [26], Verify 3D [27], ERRAT [28] and ProQ [29].

3. Results

Obtained .bam files were visualized using Integrative Genomics Viewer (IGV) [30]. The average read length was determined to be 125bp. Some .bam files contained ambiguous alignments due to presence of nonspecific amplicons produced during the amplification of gene regions. Thus, all ambiguous alignments and reads with less than 125bp were removed and filtered according to their mapping quality before calling variants using MQ>50. Across all analyzed sequences from all breeds, five intronic regions in 13 individuals could not be amplified, and therefore NGS results are missing for these individuals. Furthermore, one individual from HOL breed appeared to have variations found only in Anatolian breeds and was subsequently excluded from the analysis.

Within the three *TLR* genes analyzed, a total of 360, 463, 520, 423 and 274 SNP variants were determined on breed level in AB, EAR, SAR, TG and HOL individuals used in this study, including 26, 23, 68, 22 and four novel SNPs (Table 1). According to average data on determined SNPs and genes, the highest SNP variation was found in the *TLR6* gene, whereas the lowest one was found in *TLR4* (Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

Table 1. SNP variants identified in *TLR2*, *TLR4* and *TLR6* genes. In each cell of the table, the total/novel number of SNPs is shown, or only the total if no novel SNPs were found.

gene	breed				
	AB	**EAR**	**SAR**	**TG**	**HOL**
TLR2	**133/11**	**153/17**	**155/13**	**187/16**	**45/3**
Exon	23	27	28	26	6
missense	13	16	17/1	15/1	4
synonymous	10	11/1	11/1	11	2
Intron	93/10	111/15	110/10	146/15	34/3
3'UTR	15/1	13/1	15/1	14	4
5'UTR	2	2	2	1	1
TLR6	**209/14**	**230/5**	**285/52**	**179/5**	**155**
Exon	26	23	26	24	9
missense	10/2	9/1	9/1	8/1	3
synonymous	16/1	14	17/2	16/1	6
Intron	177/8	203/4	252/46	153/2	103
3'UTR	2	1	3/2	2	1
5'UTR	4/3	4	5/1	1/1	2
TLR4	**18/1**	**80/1**	**80/3**	**57/1**	**74/1**
Exon	2	16	13	10	15
missense	1	6	2	3/1	5
synonymous	1	10	11	7	10
Intron	14	59/1	62/3	45	55/1
3'UTR	1	2	2	1	1
5'UTR	1	3	5	1	5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

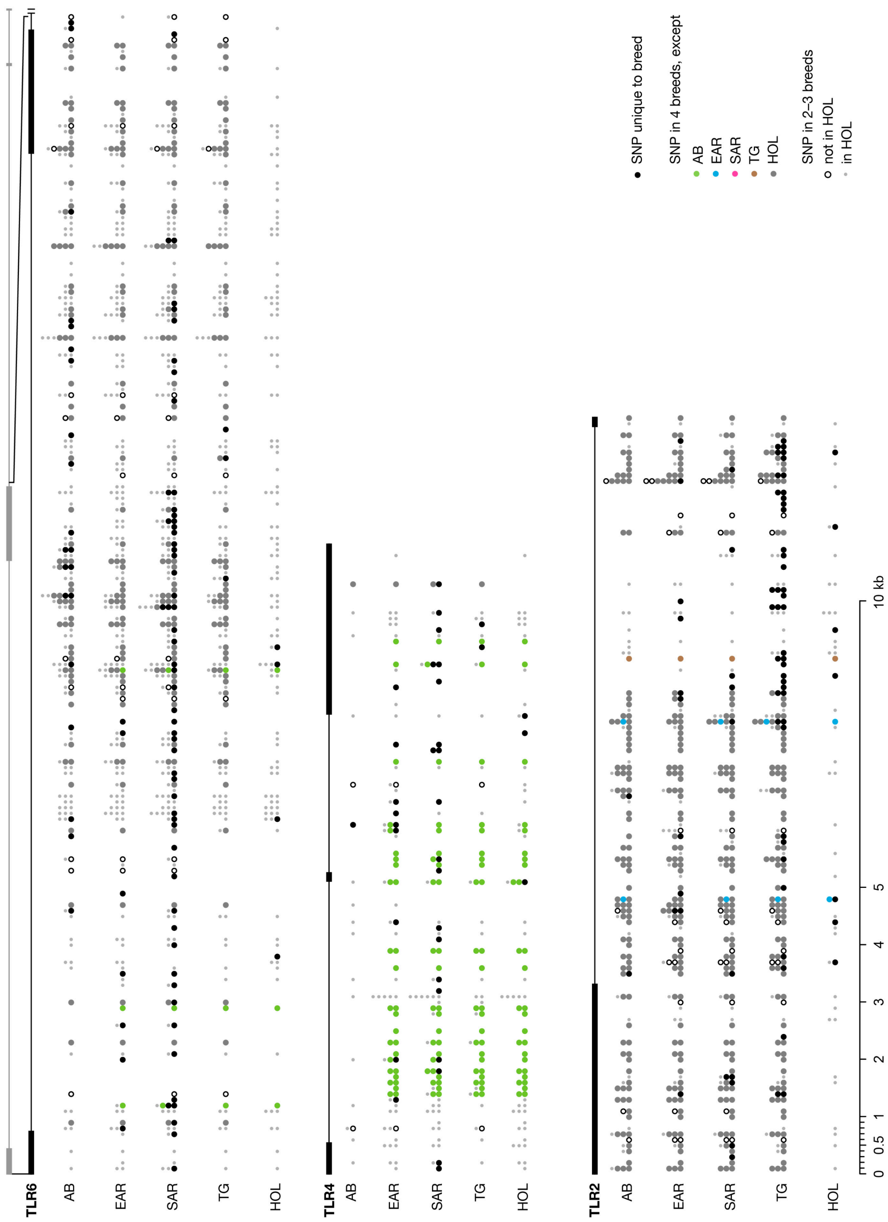

Figure 2. The location of SNPs in TLR6, TLR4 and TLR2 in five breeds of cattle: AB, EAR, SAR, TG and HOL. SNPs are shown by circles formatted to emphasize their potential interest. In decreasing order of interest: SNPs unique to a breed (black), SNPs found in all breeds except one (colored by breed), SNPs found in 2–3 breeds and not in HOL (hollow), SNPs found in 2–3 breeds including HOL (small grey circle). Only the first two exons and first intron for TLR6 is shown.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

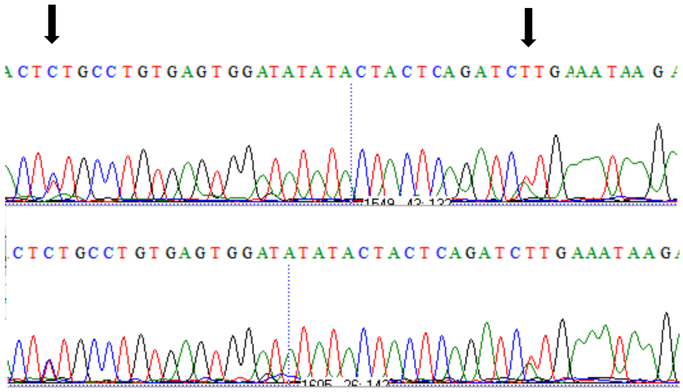

Figure 3. An example of the electropherogram, arrows showing R337Q and H326Q, respectively, in AB breed.

The SNP variants spread in the whole genes; however InDel variants were identified only in intronic regions, suggesting no effect on the gene function (Table 2). Novel variants and variants potentially impacting on molecule structure as identified by NGS results were subsequently confirmed using bidirectional Sanger sequencing (Figure 3).

Table 2. InDel variants identified in *TLR2*, *TLR4* and *TLR6* genes. In each cell of the table, the total/novel number of variants is shown.

gene	breed				
	AB	**EAR**	**SAR**	**TG**	**HOL**
TLR2	12/12	16/13	15/12	16/13	6/6
TLR6	23/9	28/14	38/24	21/13	17/3
TLR4	3/1	5/3	8/3	6/4	5/1

By just analyzing SNPs occurring in the exons of *TLR2*, *TLR4* and *TLR6*, a total of 33, 24, 46 SNP variants covering the privates and the shared among breeds, respectively were determined and according to these SNPs 36, 25 and 98 different haplotypes per corresponding TLR were constructed using ShapeIT software [31]. Obtained haplotypes (Supplementary Table 3) were visualised in a Median Joining (MJ) tree using Network program [32] (Figure 4, 5), except TLR6 due to high number of haplotype. Within all haplotypes, two main haplotypes were identified for *TLR2* and *TLR6*, whereas three major haplotypes were identified for *TLR4*. Analyzing haplotypes by breed, we identified 36 haplotypes in *TLR2* for Anatolian breeds, but only four within HOL. Interestingly however, we were unable to identify any breed specific haplotypes for *TLR4* and *TLR6*, but we able to identify 8 and 12 haplotypes, respectively, that were shared between Anatolian breeds and HOL (Supplementary Table 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

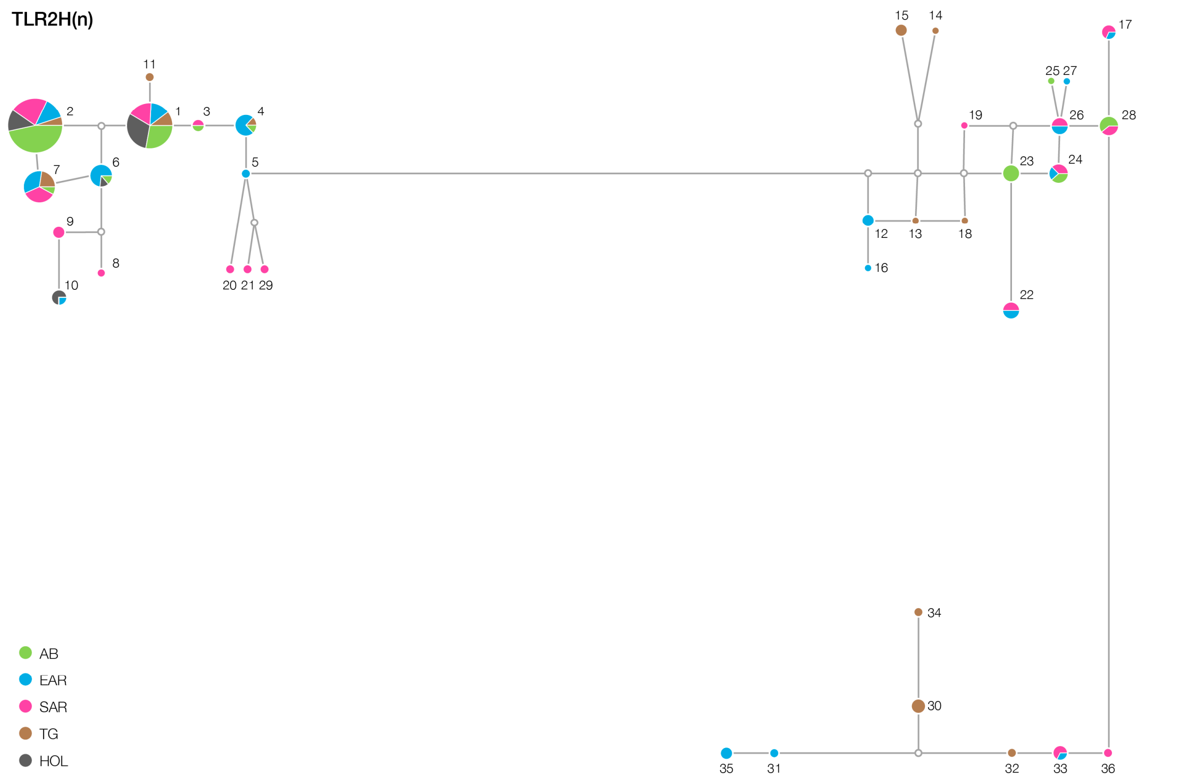

Figure 4. Median joining network constructed on the basis of 36 haplotype of *TLR2*. Filled circles represent haplotypes, areas within circles are proportional to the number of individuals. The length of the lines connecting haplotypes and branching points correlates approximately with the number of base substitution.

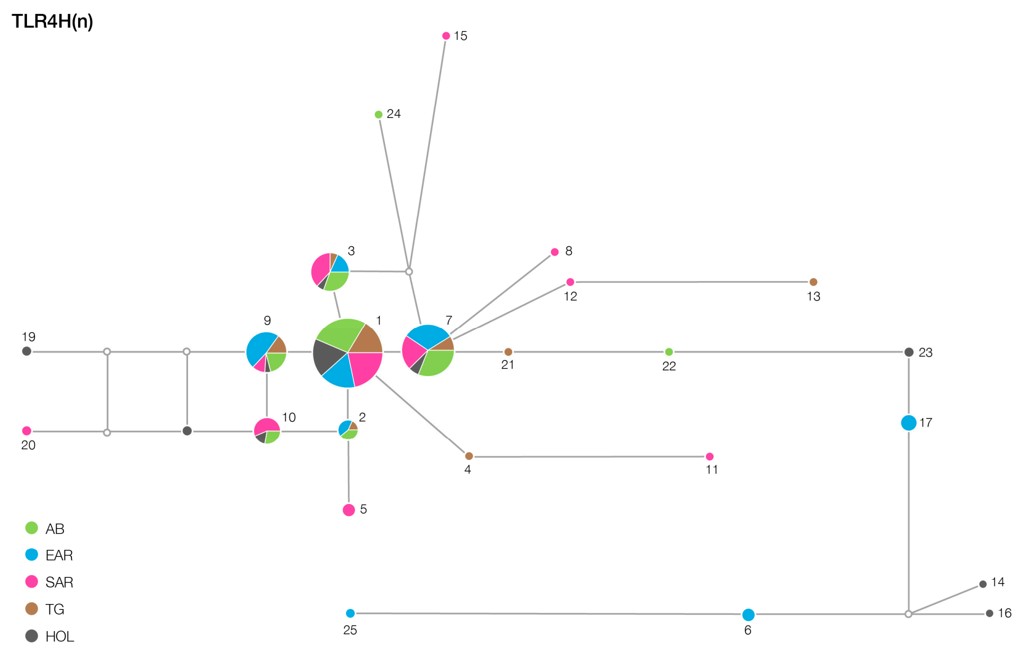

Figure 5. Median joining network constructed on the basis 25 haplotype of *TLR4*. Filled circles represent haplotypes, areas within circles are proportional to the number of individuals. The length of the lines connecting haplotypes and branching points correlates approximately with the number of base substitution.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1
Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

Missense and synonymous variations were analysed according to reference proteins for all genes and several changes were found in the analysed genes affecting the amino acids characteristics (Table 3-5).

Table 3. Missense and synonymous variations on protein domain level of *TLR2* (according to reference sequence, NP_776622.1).

Domain	TLR2(aa)	AB	EAR	SAR	TG	HOL	SNP ID
LRR1	54..77	62 N/N	62 N/N	62 N/N	62 N/N		rs68268249
		63 E/D	63 E/D	63 E/D	63 E/D	63 E/D	rs55617172
		68 G/S	68 G/S	68 G/S	68 G/S		rs68268250
LRR3	102..125		119W/L	119W/L	119W/L		rs211243949
LRR5	151..175		152 R/Q	152 R/Q		152 R/Q	rs43706434
LRR7	200..223		201 S/N			201 S/N	rs110491977
		211 I/V	211 I/V	211 I/V	211 I/V	211 I/V	rs43706433
LRR8	224..250	227 F/L	227 F/L	227 F/L	227 F/L		rs68268251
LRR9-10	251..308						
LRR11	309..337				315 R/R		rs68268253
		326 H/Q	326 H/Q	326 H/Q	326 H/Q		rs68343167
LRR12	338..361	337 R/Q	337 R/Q	337 R/Q	337 R/Q		rs68343168
LRR13	362..388						
LRR14	389..414	405 T/M	405 T/M	405 T/M	405 T/M		rs68268255
LRR15	415..437	417 N/S	417 N/S	417 N/S	417 N/S		rs68268256
		436 G/G	436 G/G	436 G/G	436 G/G		rs68268257
LRR16-18	438..500						
LRR19	501..524	502 S/A	502 S/A	502 S/A	502 S/A		rs68268258
				530 A/A			novel
				531 G/S			novel
LRR20	533..586	544 F/F	545 F/F	546 F/F	547 F/F	544 F/F	rs68268259
		563 R/H	563 R/H	563 R/H	563 R/H		rs68268260
		569 H/H	569 H/H	569 H/H	569 H/H	569 H/H	rs41830058
Trans Membrane	588..608			574 R/W			novel
		593 A/A	593 A/A	593 A/A	593 A/A		rs68268261
		594 A/A	594 A/A	594 A/A	594 A/A		rs68343169
		605 T/M	605 T/M	605 T/M	605 T/M		rs68343170
					634A/V		novel
TIR	640..784		644 F/F				novel
					655 V/A		
		665 H/Q	665 H/Q	665 H/Q	665 H/Q		rs68268263
		675 H/H	675 H/H	675 H/H	675 H/H		rs68343171
		685 I/I	685 I/I	685 I/I	685 I/I		rs68268264
		738 E/Q	738 E/Q	738 E/Q			rs207552166
		738 E/E	738 E/E	738 E/E	738 E/E		rs68268266
ATG16Lmotif	761..778	765 P/P	765 P/P	765 P/P	765 P/P		rs68268267

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

Table 4. Missense and synonymous variations on protein domain level of *TLR4* (according to reference sequence, NP_776623.5).

Domain	TLR4(aa)	AB	EAR	SAR	TG	HOL	SNP ID
LRR1	55..76						
LRR2	79..100						
LRR3	103..124						
LRR4	127..148						
LRR5	151..172		151 N/T			151 N/T	rs8193049
LRR6	176..197						
LRR7	205..225						
			238 N/K				rs8193050
			276 F/F				rs8193051
			347 A/E			347 A/E	rs8193053
LRR8	352..373						
LRR9	374..394		374 P/P	374 P/P	374 P/P	374 P/P	rs8193054
				381 K/R			rs8193055
				385 L/L			rs8193056
				389 G/G	389 G/G		rs8193057
LRR10	400..422						
LRR11	423..444						
LRR12	448..469						
LRR13	472..495				482 S/Y		novel
LRR14	497..518		507 Q/Q	507 Q/Q	507 Q/Q	507 Q/Q	rs8193059
LRR15	521..542						
LRR16	545..568	552 S/S	552 S/S	552 S/S	552 S/S	552 S/S	rs8193060
LRR17	-						
LRR18	-						
LRR19	-						
LRRCT	579..626			589 S/S			rs8193061
			609 C/C			609 C/C	rs8193062
					622 S/S		rs8193063
			625 N/N			625 N/N	rs8193064
			649 G/G		649 G/G	649 G/G	rs8193065
			640 V/I			640 V/I	rs8193066
Trans membrane	633..653						
			664 G/G			664 G/G	rs8193067
		674 T/I	674 T/I	674 T/I	674 T/I	674 T/I	rs8193069
				676 D/D			rs8193070
TIR	677..815						

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

Table 5. Missense and synonymous variations on protein domain level of TLR6 (according to reference sequence, NP_001001159.1).

Domain	TLR6 (aa)	AB	EAR	SAR	TG	HOL	SNP ID
LRR_RI	<43..164	37 D/N		37 D/N	37 D/N		
LRR1	54..77	61 Q/Q	61 Q/Q	61 Q/Q	61 Q/Q		rs68268271
LRR2	78..101	87 R/G	87 R/G	87 R/G	87 R/G		rs68268272
LRR3	102..122	116 S/P	116 S/P				
LRR4	123..147	135 D/H	135 D/H	135 D/H	135 D/H		rs520121582
LRR5- LRR6	148..196						
LRR7	197..219	214 D/N	214 D/N	214 D/N	214 D/N	214 D/N	rs43702941
		217 A/A	217 A/A	217 A/A	217 A/A		rs68268273
LRR8-LRR12	220.354						
LRR13	355..378	374 D/D			374 D/D	374 D/D	rs68268274
LRR14	379..404	395 T/A	395 T/A	395 T/A	395 T/A		rs68268275
		400 K/K	400 K/K	400 K/K	400 K/K		rs211657505
LRR15	405..428	425 S/S	425 S/S	425 S/S	425 S/S		rs68268276
LRR16	429..449						
LRR17	450..473		458 H/H	458 H/H			rs68268277
LRR18	474..495						
LRR19	496..519	505 N/N	505 N/N	505 N/N	505 N/N		rs55617146
		526 V/A	526 V/A	526 V/A	526 V/A	526 V/A	rs68343174, rs133754378
		526 V	526 V	526 V	526 V	526 V	rs68343175, rs136574510
LRRCT	529..582	539 D/D	539 D/D	539 D/D			rs68343176
			544 V/I	544 V/I			rs55617465, rs68268279
		573 K/K	573 K/K	573 K/K	573 K/K		rs55617193
Trans membrane	585..605	589 V/I	589 V/I	589 V/I	589 V/I	589 V/I	rs55617317, rs207882984
		605 L/L	605 L/L	605 L/L	605 L/L	605 L/L	rs68268280, rs378853146
TIR	641..784	642 F/F	642 F/F	642 F/F	642 F/F		rs438448894
		669 I/V	669 I/V	669 I/V	669 I/V		rs210580164
		674 H/H	674 H/H	674 H/H	674 H/H	674 H/H	rs209572763
		676 R/R	676 R/R	676 R/R	676 R/R		rs68343178
		680 A/A		680 A/A	680 A/A		novel
		684 I/I	684 I/I	684 I/I	684 I/I		rs68343179
		700 F/F	700 F/F	700 F/F	700 F/F		rs55617339, rs211454671
		701 V/V	701 V/V	701 V/V	701 V/V	701 V/V	rs207586910
			709 S/S	709 S/S			rs55617335
				E/E 710			rs68268282

3.1 Impact of identified SNPs on protein model

Within the identified SNPs, we identified an important amino acid characteristic change a charged residue into an uncharged residue in LRR11 of TLR2. This area is part of the TLR2 ligand binding domain, spanning LRRS 9-12.We visualized the importance of this change by constructing a hybrid protein model for one AB individual, carrying the H326Q change, and other variants with one HOL individual. The resulting model was visualized in PyMol [33] (Figure 6).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1
Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

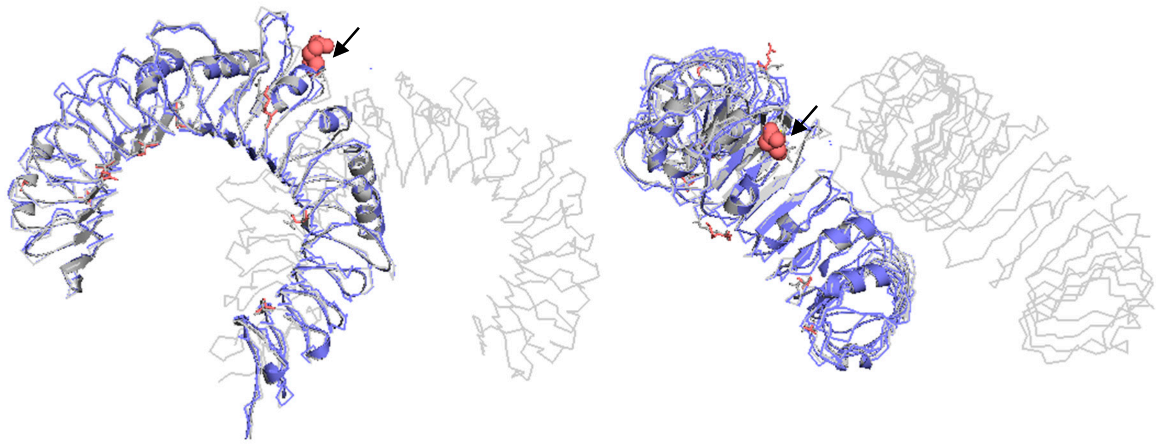

Figure 6. Predicted three-dimensional structures of homodimer of TLR2 representing the H326Q (arrows) variant. Blue color: AB breed H326Q carrying individual, dark grey color: one of HOL breed.

4. Discussion

Within the current study, our goal was to determine whether genetic variations in the sequences of innate immune receptor between indigenous cattle breeds may explain some of the observed resistance and susceptibility to specific bacterial infections.

Several association studies have been conducted to find susceptibility related alleles on *TLR2*, *TLR4* and *TLR6* genes [34-37]. None of the previously identified susceptibility associated alleles or haplotypes were found in the present study in indigenous cattle breeds, however, all of them were identified in HOL.

Anatolian breeds seem to be more genetically resistant to infection with bacteria such as *Mycobacterium bovis* and mastitis-causing bacteria, and when analyzing TLRs involved in the recognition of these bacterial pathogens, we indeed identified breed specific SNPs within the genes for the receptors investigated which differ significantly from those present in Holstein-Friesian cows.

Interestingly, when grouping the identified SNPs into haplotypes, Anatolian breeds grouped into different haplotypes for TLR2, but not for the other TLRs. This is an important observation, as those bacterial diseases to which Anatolian breeds have been described to be more resistant are mainly caused by bacteria binding to TLR2 [38]. Analyzing the identified SNPs in TLR2 in more detail, one aa change was identified in LRR1-10, four between LRR11-LRR20, one in the TM, and four in the TIR domains, respectively, in native cattle breeds whereas only one was determined in LRR5 in HOL. Considering that the ligand-binding region of TLR2 encompasses LRR9-12, the most important change causing amino acid characteristic change (H326Q) was found in LRR11. In addition, the aa changes identified in the TIR domain (H665Q and E738Q) need to be further investigated, as these might impact on subsequent intracellular signaling events, similar as described recently for bovine TLR5 [39].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

In previous studies, aa changes L227P, H305P and H326Q in the bovine TLR2 gene have been described to be under positive selection [10]. In our study, aa at position 227 was L, whereas aa at position 326 was Q and no change was determined for aa at position 305. Furthermore, aa changes L227P, H326Q, N417S and H665Q have been identified as being specific to *Bos indicus* cattle breeds [10], similar to T405M [10], which we also identified in our study. It is currently assumed that these variations may represent geographical differences, being driven by different microbiological environment. It has also been suggested that the *Bos indicus* specific aa changes H326Q and R563H might also be found in cattle breed that originated in a similar geographical and microbial environment [10].

Indeed, blood parasites have been described to cause substantial economical losses in terms of production [40]. Genetic variations in *TLR2* of *Bos indicus* cattle breeds are assumed to impact on blood parasites infection [41]. When EAR hybrid cattle population breed in Diyarbakir were screened for *Theileria* blood parasite, only 24 out of 100 samples taken from physiologically healthy cows tested positive for *Theileria* [42]. Similarly, when 300 EAR hybrid cattle from Erzurum were screened for the presence of *Babesia spp.*; only 45 samples tested positive [43]. Taken further into account that previous phylogenetic studies concluded that Anatolian cattle breeds were *Bos indicus* and *Bos taurus* hybrids [44,45], we believe that either similar selective pressure may exists in Anatolian cattle breeds due to the similar geographic/microbial environment compared to pure *Bos indicus* breeds, or that *Bos indicus* breeds may have been crossed in at an earlier time due to increase the resistance of the local breeds to various infectious diseases.

In comparison with previous studies [34, 46, 47] highest variation number found in the analyzed genes in Anatolian breeds except for *TLR4*, where HOL breed has more variations than AB and TG breeds. In these studies, researchers sampled minimum one individual of each breed that went under artificial selection and analyzed innate immunity related genes partially. It's a known fact that the selection pressure for the quantitative traits associated with productions has a negative effect on immunity traits [48,49]. The cited studies showed that cattle breeds were under strong selection pressure, which might lead to decrease in the variation on the gene regions. In the present study Anatolian cattle breeds were indigenous breeds, which evolved under natural selection over the years. With natural selection animals which cannot resist to diseases and environmental conditions cannot find chance for reproduction thus they are eliminated from the population [50]. It can be assumed that the determined polymorphism and haplotype have a potential positive effect on immunity traits. Phylogenetic trees are useful tools for visualization the evolutionary relationships of species and breed. Determination of the lowest haplotype number in HOL breed, might be due to low sampling size and/or inbreeding for years. However TG breed, have the same sample size with HOL breed, 12 haplotypes were determined for TLR2 (Figure 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

In addition to *TLR2*, we also assessed the occurrence of SNPs in *TLR4* and *TLR6*, known to form heterodimers with *TLR2*. In the *TLR6* gene, synonymous SNPs were positively associated with the susceptibility for bovine tuberculosis [51]. In the analysed individuals of the presented study there were not any breed specific differences for the 4 synonymous variations. Beside that a SNP array study associated protein tyrosine phosphatase receptor T (PTPRT) and myosin IIIB (MYO3B) with bovine tuberculosis resistance [51]. The highest variation were identified in *TLR6* but non-synonymous variations was seen in the LRR1-LRR14 in native cattle whereas only one seen in HOL cattle. One variation found in LRR14 may impact on resistance and subsequently might be under positive selection as it was only detected in native cattle breeds [47].

With regards to *TLR4*, we identified two non-synonymous SNP in LRR7 (N238K and A347E), which does not contribute to the MD2 binding region. The remaining variants identified all represented synonymous variations. *TLR4* region analyzed for somatic cell score (SCS) and 3 SNPs were associated with high and low SCS values (intron1 rs8193046 A/G, exon3 rs8193060 C/T, 5'UTR rs29017188 C/G) [52]. ACC haplotype was associated with low SCS whereas GTG haplotype was associated with high SCS. High allele frequencies were found for ACC haplotype in TG cattle breed. *TLR4* is also located in one of the QTL loci for milk production and mastitis [5]. Within the analyzed breeds, AB breed has the lowest milk production and variation for *TLR4*, whereas highest variation number determined in the high milk yield breeds, SAR, EAR and HOL as shown in Figure 5.

5. Conclusions

Given global warming due to climate change and increased anti-microbial resistance due to the overuse of antibiotics, new approaches are needed to manage infectious diseases in farm animals. A key step towards these is identification of genotypes conferring resistance to both disease and adverse environmental conditions. Our identification of non-synonymous SNPs and novel variants of *TLR2*, *TLR4* and *TLR6* genes can provide one such set of variants for future association studies and the validation of the candidate haplotypes at a cellular level.

Supplementary Materials: The following are available online at www.mdpi.com/link, Table S1: title,

Acknowledgments: This research has been supported by Ankara University Scientific Research Projects Coordination Unit. Project Number: 13B3338011, 2013; 14H0239004, 2014 and The Scientific and Technological Research Council of Turkey (TUBITAK): 1059B141400991, 2015. Authors would like to thank Martin Krzywinski for assistance in preparation of Figures 1-5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

Author Contributions: N.B.; B.C.K. and O.E. conceived and designed the experiments; N.B. and B.C.K. performed the experiments; N.B.; V.O.; and D.W. analyzed and interpret the data; N.B. and D.W. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kumar, H.; Kawai, T.; Akira, S. Pathogen recognition by the innate immune system. *Intl Rev Immunol* **2011,** 30, 16–34.
2. Medzhitov, R.; Janeway, C. A. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* **1997**, 91, 295-8.
3. Tizard, I. R. Veterinary Immunology. 3rd Ed. Saunders Elsevier, Missouri, USA, 2009; pp. 13-28
4. Takeda, K.; Akira, S. TLR signaling pathways. *Semin Immunol* **2004**,16, 3-9.
5. Ogorevc, J.; Kunej, T.; Razpet, A.; Dovc, P. Database of cattle candidate genes and genetic markers for milk production and mastitis. *Anim Genet* **2009**,40, 832-851.
6. Huang, T. H.; Uthe, J. J.; Bearson, S. M.; Demirkale, C. Y.; Nettleton, D.; Knetter, S.; Tuggle, C. K. Distinct peripheral blood RNA responses to Salmonella in pigs differing in Salmonella shedding levels: intersection of IFNG, TLR and miRNA pathways. *PloS one* **2011**, 6, e28768.
7. McGuire, K.; Jones, M.; Werling, D.; Williams, J. L.; Glass, E. J.; Jann, O. Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors. *Anim Genet* **2006**, 37, 47-50.
8. Kent, W. J.; Sugnet, C. W.; Furey, T. S.; Roskin, K. M.; Pringle, T. H.; Zahler, A. M.; Haussler, D. The human genome browser at UCSC. *Genome Res* **2002**, 12, 996-1006.
9. Cunningham, F.; Amode, M. R.; Barrell, D.; Beal, K.; Billis, K.; Brent, S.; Gil, L. Ensembl 2015. *Nucleic Acids Res* **2015**, 43, D662-D669.
10. 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**, 8,1.
11. Schröder, N.W.; Schumann, R. R. Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *Lancet Infect Dis* **2005**, 5, 156-164.
12. Werling, D.; Jann, O. C.; Offord, V.; Glass, E. J.; Coffey, T. J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends Immunol* **2009**, 30, 124-130.
13. Alpan O.; Arpacik, R. Sigir Yetiştiriciliği, 2nd Ed. Sahin Matbaası, Ankara, Turkey, 1996; pp. 39-46.
14. Yilmaz, O.; Akin, O.; Yener, S. M.; Ertugrul, M.; Wilson, R. T. The domestic livestock resources of Turkey: cattle local breeds and types and their conservation status. *Anim Genet Resour* **2012**, 50, 65-73.
15. Özbeyaz C. Sigir Yetiştiriciliği Ders Notlari. Ankara,Turkey, 2015; pp. 10-26
16. Rozen, S.; Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. M*ethods Mol Biol* **1999**, 365-386.
17. Picard: http://broadinstitute.github.io/picard. 11 May 2015
18. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Durbin, R. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, 25, 2078-2079,.

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

19. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; DePristo, M. A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **2010**, 20, 1297-1303.
20. Cingolani, P.; Patel, V. M.; Coon, M.; Nguyen, T.; Land, S. J.; Ruden, D. M.; & Lu, X. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. In: *Toxicogenomics in non-mammalian species*. Frontiers E-books **2012**, 35.
21. Cingolani, P.; Platts, A.; Wang, L. L.; Coon, M.; Nguyen, T.; Wang, L.; Ruden, D. M.; A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **2012**, 6, 80-92.
22. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C. A.; Banks, E.; DePristo, M. A.; McVean, G. The variant call format and VCFtools. *Bioinformatics* **2011**, 27, 2156-2158.
23. McLaren, W.; Pritchard, B.; Rios, D.; Chen, Y.; Flicek, P.; & Cunningham, F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **2010**, 26, 2069-2070.
24. Hall, T. A. (1999, January). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic Acids Symp Ser* **1999**, 41, 95-98
25. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Curr Protoc Bioinform* **2014**, 47, 1-32.
26. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **1993**, 26, 283-291.
27. Eisenberg, D.; Lüthy, R.; Bowie, J. U. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **1997**, 277, 396-404.
28. Colovos, C.; Yeates, T.O. ERRAT: an empirical atom-based method for validating protein structures. *Protein Sci* **1993**, 2, 1511-1519.
29. Wallner, B.; Elofsson, A. Can correct protein models be identified? *Protein Sci* **2003**, 12, 1073-1086.
30. Thorvaldsdottir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **2013**, 14, 178-192.
31. Delaneau, O.; Howie, B.; Cox, A.; Zagury, J.F.; Marchini, J. Haplotype estimation using sequence reads. *Am J Hum Bıol* **2013**, 93, 787-696.
32. Bandelt, H.; Forster, P.; Rohl, A. Median joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **1999**,16, 37-48.
33. DeLano, W. L. PyMOL. *DeLano Scientific*, San Carlos, CA, 700, **2002**.
34. Mariotti, M.; Williams, J. L.; Dunner, S.; Valentini, A.; Pariset, L. Polymorphisms within the toll-like receptor (TLR)-2,-4, and-6 genes in cattle. *Diversity* **2009**, 1, 7-18.
35. Garza-Gonzalez, E.; Bosques-Padilla, F. J.; Mendoza-Ibarra, S. I.; Flores-Gutierrez, J. P.; Maldonado-Garza, H. J.; Perez-Perez, G. I. Assessment of the toll-like receptor 4 Asp299Gly, Thr399Ile and interleukin-8-251 polymorphisms in the risk for the development of distal gastric cancer. *BMC Cancer* **2007**, 7, 70.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at Diversity 2016, 8, 23; doi:10.3390/d8040023

36. Mucha, R.; Bhide, M. R.; Chakurkar, E. B.; Novak, M.; Mikula, I. Toll-like receptors TLR1, TLR2 and TLR4 gene mutations and natural resistance to Mycobacterium avium subsp. paratuberculosis infection in cattle. *Vet Immunol Immunop* **2009**, 128, 381-388.
37. Ruiz-Larrañaga, O.; Manzano, C.; Iriondo, M.; Garrido, J. M.; Molina, E.; Vazquez, P.; 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.
38. Jungi, T. W.; Farhat, K.; Burgener, I. A.; Werling, D. Toll-like receptors in domestic animals. *Cell Tissue Res* **2011**, 343, 107-120.
39. Osvaldova, A.; Woodman, S.; Patterson, N.; Offord, V.; Mwangi, D.; Gibson, A. J.; ... & Werling, D. Replacement of two aminoacids in the bovine Toll-like receptor 5 TIR domain with their human counterparts partially restores functional response to flagellin. *Dev Comp Immunol* **2014**, 47, 90-94.
40. Kaiser, P. The Immune system. In Breeding for Disease Resistance in Farm Animals 1st Ed. Editor: Bishop, S.C.; Axford, R.F.E.; Nicholas, F.W.; Owen, J.B. Cambridge, MA 02139 USA, 2010; pp. 15-38.
41. Ropert, C.; Gazzinelli, R.T. Regulatory role of Toll-like receptor 2 during infection with Trypanosoma cruzi. *J Endotoxin Res* **2004**, 10, 425-430.
42. Deniz, A.; Oncel, T.; Icen, H.; Simsek, A. Detection of Theileria annulata and T. buffeli in Cattle by Multiplex PCR in Diyarbakir Area. *Kafkas Univ Vet Fak Derg* **2012**, 18, A111-A114.
43. Duzlu, O.; Yıldırım, A.; Inci, A.; Avcioglu, H.; Balkaya, I. Investigation of Babesia bovis and Babesia bigemina in cattle by Real Time PCR and molecular characterization of the isolates. *Ankara Univ Vet Fak Derg* **2015**, 62, 27-35.
44. Loftus, R. T.; Ertuğrul, O.; Harba, A.H.; El-Barody, M.A.; Machugh, D. E.; Park, S.D.; Bradley, D. G. A microsatellite survey of cattle from a center of origin: the Near East. *Mol Ecol* **1999**, 8, 2015-2022.
45. Decker, J. E.; McKay, S. D.; Rolf, M. M.; Kim, J.; Alcalá, A. M.; Sonstegard, T. S.; Hanotte, O.; Götherström, A.; Seabury, C.M.; Praharani, L.; Babar, M.E.; Correia, D.E.; Almeida Regıtano, L.; Yildiz, M.A.; Heaton, M.P.; Liu, W.S.; Lei, C.Z.; Reecy, J.M.; Saif-Ur-Rehman M.; Schnabel R.D.; Taylor J.F. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet*. **2014**, 10, e1004254.
46. 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* **2010**, 107, 151-156.
47. Fisher, C. A.; Bhattarai, E. K.; Osterstock, J. B.; Dowd, S. E.; Seabury, P. M.; Vikram, M.; Womack, J. E. Evolution of the bovine TLR gene family and member associations with Mycobacterium avium subspecies paratuberculosis infection. *PLoS One* **2011**, 6, e27744.
48. Ingvartsen, K. L.; Dewhurst, R. J.; Friggens, N. C. On the relationship between lactational performance and health: is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper. *Livest Prod Sci* **2003**, 83, 277-308.
49. Oltenacu, P. A.; Broom, D.M. The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim Welf* **2010**, 19, 39-49.
50. Haldane, J. B. S. Disease and evolution. *Curr Sci* **1992**, 63, 599-604.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2016 doi:10.20944/preprints201609.0080.v1

Peer-reviewed version available at *Diversity* **2016**, *8*, 23; doi:10.3390/d8040023

51. Bermingham, M. L.; Bishop, S. C.; Woolliams, J. A.; Pong-Wong, R.; Allen, A. R.; McBride, S. H.; Glass, E.J. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. *Heredity* **2014**, 112, 543-551.
52. Mesquita, A. Q. D.; Mesquita, A. J. D.; Jardim, E. A. G. D. V.; Kipnis, A. P. J. Association of TLR4 polymorphisms with subclinical mastitis in Brazilian holsteins. *Braz J Microbiol* **2012**, 43, 692-697.

© 2016 by the authors; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).