



Genetic diversity of Toll-like receptor 5 among pig populations

Xiuqin Yang¹, Haitao Li¹, Qingzhi Guan¹ and Di Liu²

¹College of Animal Science and Technology, Northeast Agricultural University, Harbin, China.

²Heilongjiang Academy of Agricultural Sciences, Harbin, China.

Abstract

The Toll-like receptor 5 (TLR5) recognizes flagellin of Gram-positive and -negative bacteria and plays an important role in the host defense system. Here, we surveyed single nucleotide polymorphisms (SNPs) in the coding sequence of the porcine TLR5 gene in 83 individuals from five pig breeds, these including Chinese local populations and Western commercial pig breeds. A total of 19 medium polymorphic SNPs ($0.25 < PIC < 0.5$) were identified, three of which were missense mutations that clustered within the extracellular domain of TLR5. One of the non-synonymous SNPs fell within a 228-amino acid region which has been shown to be important for flagellin recognition. Four SNPs were only found with high frequencies in Oriental pig breeds. The 19 SNPs were found in 30 haplotypes, one of which segregated at high frequency in all samples. Compared with Western pig breeds, Chinese local populations had higher genetic diversity and more haplotypes. Tajima's test showed no evidence for deviation from neutrality. The data provide useful information for future genetic marker characterization by means of disease association analysis and/or stimulating the mutation carrier with relevant ligands.

Keywords: porcine, TLR5, SNPs, haplotype.

Received: October 19, 2011; Accepted: September 21, 2012.

Introduction

Innate immunity, which develops in minutes following microbial invasion and is the first line of host defense against infection, triggers and regulates the adaptive immune system (Akira *et al.*, 2001; Medzhitov and Janeway Jr, 2000). It can be activated by the recognition of microorganisms through pattern recognition receptors (PRRs) and involves the release of cytokines and chemokines, and the activation of monocytes and macrophages (Janeway Jr and Medzhitov, 2002). Toll-like receptors, which recognize structural components unique to pathogenic microbes, are important PRRs of microbial infection and play a central role in the activation of innate immunity and the development of antigen-specific adaptive immunity through their interaction with ligands (Medzhitov *et al.*, 1997; Poltorak *et al.*, 1998; Takeda *et al.*, 2002; Zhang *et al.*, 2004).

TLR5, a member of the TLR family, specifically recognizes flagellin, which is present in the flagellar structure of Gram-positive and -negative bacteria (Hayashi *et al.*, 2001). Like other TLRs, TLR5 is composed of adjacent extracellular, transmembrane and intracellular regions. The extracellular region contains 22 tandem leucine-rich repeats (LRRs) flanked by cysteine-rich capping structures at

the N- and C-terminal sides, which are known as LRRNT and LRRCT, respectively (Matsushima *et al.*, 2007). The LRRs are motifs that are 20-30 amino acids in length and are responsible for ligand-binding. The intracellular region includes a signaling portion, the Toll/IL-1 receptor (TIR) domain, which binds adaptor molecules and induces cellular immune responses through a signaling cascade (Muzio *et al.*, 2000).

TLR5 polymorphisms play a role in various complex human diseases, including Legionnaire's disease (Hawn *et al.*, 2003), Crohn's disease (Gewirtz *et al.*, 2006), ossification of the posterior longitudinal ligament of the spine (Chung *et al.*, 2011), systemic lupus erythematosus (Hawn *et al.*, 2005), and recurrent cystitis (Hawn *et al.*, 2009). In German shepherd dogs, the non-synonymous single nucleotide polymorphisms (SNPs) c.22GA (p.Ala8Thr), c.100C > T (p.Arg34Cys) and c.1844T > C (p.Leu615Ser) have been associated significantly with inflammatory bowel disease (Kathrani *et al.*, 2010). In addition, the SNPs c.1174C > T (p.Arg392stop), c.2081A > G (p.Asp694Gly) and c.2464C > T (p.Leu822Phe) in the human TLR5 gene have been found to have a significant impact on receptor function in transiently transfected CHO-K1 cells, revealing that variants carrying the SNPs were non-responsive upon stimulation with flagellin (Merx *et al.*, 2006).

Genetic variation in the porcine TLR5 gene may also affect the ligand recognition and signal transduction of the receptor and alter the susceptibility/resistance of the host to

infectious diseases. The aim of this study was to analyze TLR5 genetic diversity in Western commercial and Chinese local pig populations. A total of 83 animals from five pig breeds were used to explore the genetic variation present in the coding sequence. The results should provide useful information for future genetic marker characterization in studies using disease association analysis and/or stimulation of the mutation carrier with relevant ligands, which is important for breeding programs aiming at disease-resistance traits in pigs.

Materials and Methods

Animals

The sample panel included 83 unrelated domestic pigs comprising Chinese local pig breeds (27 Min and 23 Beijing black pigs), Western commercial pig breeds (13 Yorkshire and 15 Landrace pigs), and 5 Chinese wild boars. Genomic DNA was extracted from ear tissue by a standard phenol-chloroform method.

PCR and sequencing

The complete coding sequence of the porcine TLR5 gene, which is 2571 bp long and composed of a single exon, was amplified by a pair of primers designed using Primer Premier 5.0 according to the sequence deposited in GenBank (No. AB208697). The primer sequences were as follows: 1F-GAA AGC TTA TGG GAG ACT GCC TGG TC (forward), 1R-GCT CTA GAC TAG GAG ATG GTC ACG CTT TG (reverse). The fragment was amplified with high fidelity *Pfu* DNA polymerase (TransGen, Beijing, China) using a 3-min step at 95 °C, followed by 30 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 2 min each, with a 7-min final extension at 72 °C. The PCR products were sequenced at the Beijing Genomics Institute (Beijing, China), using the primer pair 1F/1R and other two primers (2F-CAG ATA CCC CTT GTG TGC as the forward primer and 2R-GGG TGA CAG TGA ACA AGA TG as the reverse one).

Sequence analysis

The full-length gene contig was built using the SEQMAN module of LASERGENE (version 6.0). SNPs were detected by sequence alignment with the DNAMAN package (version 5.2.2), and each of them was confirmed by manual inspection of the sequence diagram using the Chromas software. For SNPs occurring in fewer than three individuals, the fragments were reamplified and sequenced again to reduce the risk of reporting PCR artifacts as polymorphisms.

The PROC HAPLOTYPE procedure of SAS9.1 was used to estimate haplotype frequencies by implementing the Expectation Maximization algorithm which is commonly used in haplotype phasing and frequency estimation. Confidence intervals and standard errors were estimated

under a binomial assumption, which is default in the procedure for each haplotype frequency estimate. The estimates were then utilized to assign the probability that each individual possesses a particular haplotype pair. Individual haplotypes predicted with probabilities < 99% were subject to further verification. The fragments were reamplified, sub-cloned into the pMD-18T vector, and sequenced. In each case, at least five sub-clones were sequenced, including at least one copy of each allele.

Tajima's test was performed to test for deviation from neutrality. This test proposed by Tajima (1989) is based on the fact that, under the neutral model, estimates of S/a_i and k are unbiased estimates of theta (the population mutation rate $4Ne\mu$). DNA sequence polymorphism (DnaSP) software (version 5.10) was used to calculate the Tajima's D value and its confidence limits (two-tailed test) by assuming this statistic to have a beta distribution (Tajima, 1989). At the same time, nucleotide diversity (P_i) was calculated using Equation 10.5 of Nei (1987) with DnaSP.

Relative synonymous codon usage (RSCU) value, which is defined as the ratio of observed frequency of a codon in a gene to that expected under equal codon usage, was calculated as below (Sharp *et al.*, 1986) by using the program CodonW (version 1.4.2),

$$RSCU_i = \frac{X_i}{\frac{1}{n} \sum_{i=1}^n X_i}$$

where X_i is the observed frequency of codon i in the gene under study, n is the number of synonymous codons for the amino acid studied.

Results

Nucleotide polymorphism

A total of 40 SNPs were identified in the coding sequence of the porcine TLR5 gene, with 14 of these SNPs leading to amino acid transformations. Most of them were present at frequencies lower than 3% and/or were specific to a single pig breed. Only five of the SNPs (12.5%) were transversions, while the others (87.5%) were transitions, this showing an obvious transitional bias. The pattern and distribution of base substitution are shown in Figure 1.

The nucleotide diversity value, P_i , was calculated within populations. Wild boars displayed the highest variability among all populations ($P_i = 4.2 \times 10^{-3}$), while Yorkshire pigs showed the least variability ($P_i = 1.29 \times 10^{-3}$). P_i was higher in the Chinese local populations ($P_i = 3.29 \times 10^{-3}$) than in the Western commercial pig breeds ($P_i = 1.89 \times 10^{-3}$).

Nineteen SNPs showed an intermediate level of polymorphism ($0.25 < PIC < 0.5$) (Table 1). Three SNPs, c.137G > A, c.834T > G and c.1246A > T, were non-synonymous, resulting in amino acid replacements of Gly by Asp at position 46, His by Gln at position 278, and Thr by

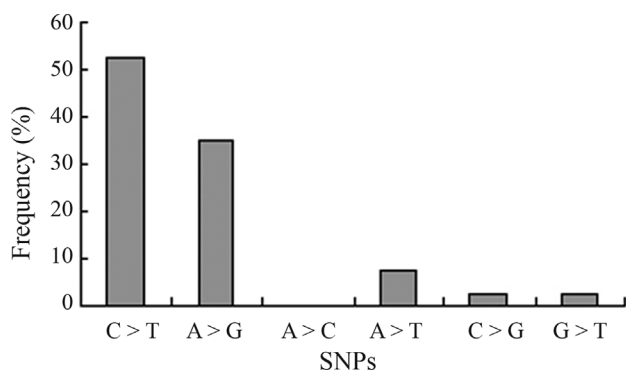


Figure 1 - Base substitution types and the respective frequencies.

Table 1 - Distributions of medium polymorphic SNPs among pig populations.

No.	SNP	First allele frequency %					Total
		M	B	W	Y	L	
1	c.137 G > A	85.2	60.9	60.00	100.0	100.0	80.7
2	c.360 T > C	44.4	39.1	0	92.3	93.3	56.7
3	c.834 T > G	42.6	39.1	20.0	53.9	76.7	48.2
4	c.1065 T > C	63.0	97.8	80.0	26.7	83.3	76.5
5	c.1246 A > T	66.7	97.8	100.0	53.9	76.7	75.9
6	c.1269 G > A	66.7	97.8	100.0	53.9	76.7	75.9
7	c.1278 C > T	66.7	97.8	100.0	53.9	76.7	75.9
8	c.1398 C > T	68.5	97.8	90.0	53.9	76.7	77.1
9	c.1812 A > T	72.2	45.7	40.0	100.0	100.0	72.3
10	c.1815 T > C	72.2	45.7	40.0	100.0	100.0	72.3
11	c.1836 G > A	72.2	45.7	30.0	100.0	100.0	71.7
12	c.2100 T > C	72.2	37.0	30.0	96.2	80.0	65.1
13	c.2124 A > G	27.8	63.0	70.0	3.9	20.0	34.9
14	c.2127 A > G	27.8	63.0	70.0	3.9	20.0	34.9
15	c.2170 T > C	72.2	41.3	90.0	96.2	80.0	69.9
16	c.2178 C > T	72.2	41.3	90.0	96.2	73.3	68.7
17	c.2346 C > G	38.9	71.7	80.0	19.2	20.0	44.0
18	c.2391 T > C	70.4	37.0	30.0	96.2	90.0	66.3
19	c.2394 A > G	70.4	37.0	30.0	96.2	90.0	66.3

Breed abbreviations as follows: M, Min pig; B, Beijing black pig; W, Wild boar; Y, Yorkshire pig; L, Landrace pig.

Ser at position 416, respectively. The SNP c.137G > A (p.Gly46Asp) was predicted to cause an acquisition of charge, from neutral to negative, and c.834T > G (p.His278Gln) a change of charge from positive to neutral. SNPs c.137G > A, c.1812A > T, c.1815T > C and c.1836G > A were only found in Oriental pig breeds (Min pig, Beijing black pig, and Wild boar). At the loci of the SNPs c.360T > C and c.834T > G, the major alleles in Oriental populations were different from those in Western breeds. All of the medium polymorphic SNPs were distributed in structural domains, seven in LRRs, eight in TIR, and others in LRRNT and LRRCT (Figure 2).

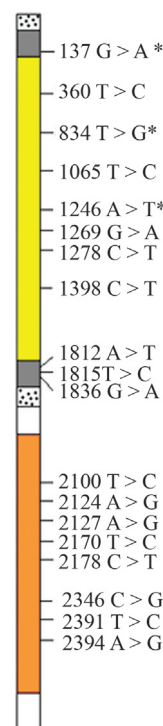


Figure 2 - Distributions of medium polymorphic SNPs in TLR5. The TIR domain was predicted by SMART, other domains were predicted as described by Matsushima *et al.* (2007). Signal peptides and transmembrane regions are indicated with dotted boxes. The LRRNT and the LRRCT are indicated with grey boxes. The LRR and the TIR domains are indicated with yellow and orange boxes, respectively. Synonymous SNPs are indicated with an asterisk.

Haplotype diversity

Thirty haplotypes were predicted and grouped for the presence of missense substitutions using the 19 medium polymorphic SNPs (Figure 3). Haplotype A1 was present at the highest frequency (36.14%) in the whole sample, whereas the other haplotypes showed frequencies ranging between 0.60% and 12.05%. Group A segregated at the highest frequency (46.38%), while variants carrying non-synonymous mutations had frequencies ranging between 0.60% and 23.49%. The SNPs c.1246A > T, c.1269G > A and c.1278C > T always cosegregated in the 83 individuals from the five pig breeds. The SNPs c.2100T > C, c.2124A > G and c.2127A > G also cosegregated in the pigs investigated, and the same cosegregation existed in the SNPs c.1812A > T and c.1815T > C. Compared with the Western commercial pig breeds, Beijing black pigs and Min pigs had more haplotypes, whereas Yorkshire pigs had the fewest among the pig breeds investigated.

Test for deviation from neutrality

Tajima's test, applied separately to individual Oriental breeds and to Western populations, was used to explore whether deviation from neutrality could be detected in the porcine TLR5 gene on the basis of all SNPs detected. For

H	SNP position	N	%	Tot. %	Distribution				
					M	B	W	Y	L
	00011111111222222222								
	1380222388811111333								
	3636467911302277499								
	7045698825604708614								

A1	GTTTAGCCATGTGGTCGTA	60	36.14	46.38	17	12	0	12	19
A2	.C.....C..	5	3.01		3	1	0	1	0
A3	.C.....TCACAACCTCCG	5	3.01		1	2	2	0	0
A4CAACTCCG	3	1.81		0	1	0	0	2
A5CAA..C..	1	0.60		0	1	0	0	0
A6CAACTC.G	1	0.60		0	0	0	1	0
A7T...	1	0.60		0	0	0	0	1
A8C..	1	0.60		0	0	0	0	1
B1	.CG....TCACAACCTCCG	13	7.83	12.05	3	8	2	0	0
B2	.CG....TCACAACCTCC	5	3.01		5	0	0	0	0
B3	.CG....CAACTCCG	1	0.60		0	1	0	0	0
B4	.G....TCACAACCTCCG	1	0.60		0	1	0	0	0
C1	AC.....C..	1	0.60	1.20	1	0	0	0	0
C2	AC.....CAACTCCG	1	0.60		0	1	0	0	0
D1	ACG....TCACAACCTCCG	17	10.24	16.26	5	10	2	0	0
D2	ACG.....CCG	3	1.81		1	2	0	0	0
D3	ACG.....	2	1.20		0	0	2	0	0
D4	ACG.....CAACTCCG	1	0.60		0	1	0	0	0
D5	A.G....CAACTCCG	1	0.60		0	1	0	0	0
D6	A.G....TCACAACCTCCG	2	1.20		0	2	0	0	0
D7	ACG....TCACAA..C..	1	0.60		0	1	0	0	0
E1	.CGCTATT.....	11	6.63	23.49	9	1	0	0	1
E2	.GCTATT.....	20	12.05		7	0	0	11	2
E3	.GCTAT.....	1	0.60		1	0	0	0	0
E4	.CGCTAT.....C..	1	0.60		0	0	1	0	0
E5	.CGCTATT..ACAACCTCCG	1	0.60		0	0	1	0	0
E6	.CG.TATT.....CC..	1	0.60		0	0	0	1	0
E7	.GCTATT..CAACTC..	3	1.81		0	0	0	0	3
E8	.CGCTATT.....T...	1	0.60		0	0	0	0	1
F1	AC.CTATTTCACAACCTCCG	1	0.60	0.60	1	0	0	0	0

Figure 3 - Haplotypes and their distributions among five pig breeds. H, haplotype; *, non-synonymous SNPs; N, numbers of each haplotype; %, frequency of each haplotype; Tot. %, frequency of each haplotype group. Haplotypes named with the same capital letter carry synonymous substitutions.

each breed, and for Oriental breeds, Western breeds, and all breeds combined, respectively, both positive and negative values were obtained for both the full gene and ectodomain, but they were not significantly different from zero ($p > 0.05$) (Table 2).

Discussion

Non-synonymous SNPs

In this study, the genetic diversity of the porcine TLR5 gene was analyzed using PCR and a direct sequencing method. Nineteen SNPs of intermediate polymorphism were identified in the coding sequence, three of which were non-synonymous substitutions and clustered within the extracellular domain (Table 1, Figure 2). The SNP c.137G > A (p.Gly46Asp) is located in the LRRNT domain, which is composed of 27 amino acids (positions 21 to 47 in the polypeptide, as predicted by Matsushima *et al.*, 2007) and is hypervariable among species (Figure 4). Both

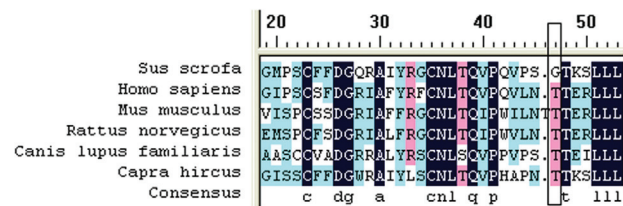


Figure 4 - Comparison of the LRRNT domain of TLR5 among species. SNP p.46Gly in porcine is boxed. Homologous sequences were downloaded from GenBank (AB208697 for *Sus scrofa*, NM003268 for *Homo sapiens*, NM016928 for *Mus musculus*, FJ750588 for *Rattus norvegicus*, EU551146 for *Canis lupus familiaris*, and FJ659852 for *Capra hircus*).

the SNPs c.834T > G (p.His278Gln) and c.1246A > T (p.Thr 416Ser) fell into irregular LRR domains, LRR9 and LRR15, respectively (Matsushima *et al.*, 2007).

The LRRNT can shield the hydrophobic core of the first LRR from exposure to solvent, and therefore is expected to play a role in stabilizing the protein structure (Matsushima *et al.*, 2005). Irregular LRRs are common in mammalian TLRs, and involvement in ligand recognition and binding has been suggested (Bell *et al.*, 2003; Wei *et al.*, 2011). In addition, p.His278Gln occurs within a 228-amino acid region, which has been revealed to confer species-specific flagellin recognition in the mouse and human (Andersen-Nissen *et al.*, 2007). It will be important to further analyze the effect of the SNPs on receptor function by means of eukaryotic expression vectors of the porcine TLR5 gene with site-directed mutagenesis in cell cultures, and/or to analyze their association with immune-related traits in various populations.

Synonymous SNPs

Several synonymous SNPs were found in the LRRs and the TIR domains (Figure 2). Synonymous SNPs do not affect amino acid sequence, and therefore are not expected to alter the function of the protein. However, accumulating evidence indicates that synonymous SNPs can have an effect on protein function through alteration of mRNA stability, mRNA splicing, mRNA expression, mRNA maturation, and mRNA folding (Hoffmeyer *et al.*, 2000; Duan *et al.*, 2003; Johnson *et al.*, 2005). For example, a silent mutation in the MDR1 gene results in decreased levels of mRNA by a change in mRNA stability, thus altering the protein level (Wang *et al.*, 2005). Another example is a synonymous SNP in exon 4 of the human leukocyte antigen (HLA)-A2 gene, which results in the alternative expression of HLA-A*01:01:38L through aberrant splicing (Dunn *et al.*, 2011). In addition, the presence of a rare codon, marked

Table 2 - Tajima'D value.

Tajima'D	M	B	W	Oriental	Y	L	Western	Total
Full gene	0.4229	1.1750	-0.1614	0.9137	-1.5061	0.0030	-0.4938	0.6496
Ectodomain	-0.0834	1.0691	-0.5121	8.6559	-0.9454	-0.3512	-0.6072	-0.0803

by a silent mutation, can affect the translation rate as a result of the scarcity of cognate tRNA, which in turn influences the cotranslational folding and/or insertion of a protein into the membrane, thereby altering the conformation and function of the protein (Kimchi-Sarfaty *et al.*, 2007; Komar, 2007).

RSCU is an effective measurement of codon usage bias. If the synonymous codons are used equally, the respective RSCU values would be 1.00. A codon which is used more frequently than expected will have a value higher than 1.00, and vice versa for a codon which is used less frequently than expected. In this study, six synonymous SNPs, c.2100T > C, c.2124A > G, c.2127A > G, c.2170T > C, c.2178C > T, and c.2394A > G, led to codon replacements with significantly different values of RSCU. In particular, SNPs c.2100T > C, c.2124A > G, and c.2127A > G always cosegregated, in the pattern of TGG or CAA, in the populations under study (Figure 3). In the case of SNP c.2100T > C, the RSCU values for codons GGT and GGC were 0.61 and 1.13 respectively; whereas in c.2124A > G, the RSCU values for GAA and GAG were 1.38 and 0.63; and in c.2127A > G, the RSCU values for GCA and GCG were 1.05 and 0.32, respectively. Furthermore, the substitutions all occurred at the third base in the codon, which suggests a large effect on protein folding (Gupta *et al.*, 2000; Cortazzo *et al.*, 2002). The simultaneous replacement of more frequently used codons GGC-GAA-GCA by the less frequent ones GGT-GAG-GCG at the three adjacent positions probably slows down the rate of translation at the corresponding mRNA region, and as a result, may influence the cotranslational folding and the function of the protein.

Variability in populations

Among the populations investigated, wild boar showed the highest variability, which is consistent with the common assumption that genetic variability decreases as a result of the selective breeding. Compared with the Western commercial populations, Min and Beijing black pigs had higher variability, as revealed by the *Pi* value and the number of haplotypes (Figure 3), which reflects the known breeding history of each breed. The Min pig, a Chinese indigenous breed, and the Beijing black pig, which was produced by crossing Chinese indigenous breeds with Western commercial populations, have an extensive genetic basis, whereas many alleles have been fixed in the Western commercial breeds during a long period of intensive breeding.

Tajima's test

The Tajima's test is a commonly used test of neutrality in population genetic studies. It tests whether sequences in question fit the neutral theory model at equilibrium between genetic drift and mutation. If the D value is close to zero, the null neutral hypothesis is accepted, which implies that the population has not experienced any contraction or

recent growth, and that no selection is acting at the locus (Tajima, 1989). On the contrary, if the D value is significantly different from zero, the null hypothesis is rejected (Tajima, 1989). Compared with other intraspecies neutrality tests that use polymorphic data within a single species such as Fay and Wu's H, Tajima's D test does not require an outgroup sequence (Tajima, 1989, Fay and Wu, 2000). In the present study, the Tajima's D statistic for each breed, and for Oriental breeds, Western breeds, and all breeds combined, showed no significant difference from zero ($p > 0.05$), hence, the null hypothesis was not rejected. This indicates that the evolution of the porcine TLR5 gene is caused by random drift of selectively neutral mutants. The result is consistent with the results obtained for the human and chimpanzee TLR5 genes using the same method (Wlasiuk *et al.*, 2009). Studies have shown that selection direction in TLRs may differ among species, and many factors can influence their prediction by statistical methods, such as relaxation of purifying selection and the analytical methods used (Kryazhimskiy and Plotkin, 2008; Barreiro *et al.*, 2009). In addition, in this study the sample size per population was small, which could also affect the results of the test. Therefore, the results of this study should be interpreted with caution.

Acknowledgments

This work was supported by the Natural Science Foundation of Heilongjiang Province (Item No. 200927) and the National Natural Science Foundation of China (Item No. 31072007).

References

- Akira S, Takeda K and Kaisho T (2001) Toll-like receptors: Critical proteins linking innate and acquired immunity. *Nat Immunol* 2:675-680.
- Andersen-Nissen E, Smith KD, Bonneau R, Strong RK and Aderem A (2007) A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *J Exp Med* 204:393-403.
- Barreiro LB, Ben-Ali M, Quach H, Laval G, Patin E, Pickrell JK, Bouchier C, Tichit M, Neyrolles O, Gicquel B, *et al.* (2009) Evolutionary dynamics of human toll-like receptors and their different contributions to host defense. *PLoS Genetics* 5:e1000562.
- Bell JK, Mullen GE, Leifer CA, Mazzoni A, Davies DR and Segal DM (2003) Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol* 24:528-533.
- Chung W, Nam DH, Jo DJ and Lee JH (2011) Association of toll-like receptor 5 gene polymorphism with susceptibility to ossification of the posterior longitudinal ligament of the spine in Korean population. *J Korean Neurosurg Soc* 49:8-12.
- Cortazzo P, Cerveňanský C, Marín M, Reiss C, Ehrlich R and Deana A (2002) Silent mutations affect *in vivo* protein folding in *Escherichia coli*. *Biochem Biophys Res Commun* 293:537-541.
- Duan J, Wainwright MS, Comeron JM, Saitou N, Sanders AR, Gelernter J and Gejman PV (2003) Synonymous mutations

- in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum Mol Genet* 12:205-216.
- Dunn PP, Hammond L, Coates E, Street J, Griner L and Darke C (2011) A “silent” nucleotide substitution in exon 4 is responsible for the “alternative expression” of HLA-A*01:01:38L through aberrant splicing. *Hum Immunol* 72:717-722.
- Fay JC and Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics* 155:1405-1413.
- Gewirtz AT, Vijay-Kumar M, Brant SR, Duerr RH, Nicolae DL and Cho JH (2006) Dominant-negative TLR5 polymorphism reduces adaptive immune response to flagellin and negatively associates with Crohn’s disease. *Am J Physiol Gastrointest Liver Physiol* 290:G1157-G1163.
- Gupta SK, Majumdar S, Bhattacharya TK and Ghosh TC (2000) Studies on the relationships between the synonymous codon usage and protein secondary structural units. *Biochem Biophys Res Commun* 269:692-696.
- Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM and Aderem A (2001) The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 410:1099-1103.
- Hawn TR, Verbon A, Lettinga KD, Zhao LP, Li SS and Laws RJ (2003) A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires’ disease. *J Exp Med* 198:1563-1572.
- Hawn TR, Wu H, Grossman JM, Hahn BH, Tsao BP and Aderem A (2005) A stop codon polymorphism of Toll-like receptor 5 is associated with resistance to systemic lupus erythematosus. *Proc Natl Acad Sci USA* 102:10593-10597.
- Hawn TR, Scholes D, Li SS, Wang H, Yang Y, Roberts PL, Stapleton AE, Janer M, Aderem A, Stamm WE, *et al.* (2009) Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLoS One* 4:e5990.
- Hoffmeyer S, Burk O, von Richter O, Arnold HP, Brockmüller J, John A, Cascorbi I, Gerloff T, Roots I, Eichelbaum M, *et al.* (2000) Functional polymorphisms of the human multidrug-resistance gene: Multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. *Proc Natl Acad Sci USA* 97:3473-3478.
- Janeway Jr CA and Medzhitov R (2002) Innate immune recognition. *Annu Rev Immunol* 20:197-216.
- Johnson AD, Wang D and Sadee W (2005) Polymorphisms affecting gene regulation and mRNA processing: Broad implications for pharmacogenetics. *Pharmacol Ther* 106:19-38.
- Kathrani A, House A, Catchpole B, Murphy A, German A and Werling D (2010) Polymorphisms in the TLR4 and TLR5 gene are significantly associated with inflammatory bowel disease in German shepherd dogs. *PLoS One* 12:15740-15747.
- Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE and Calcagno AM (2007) A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* 315:525-528.
- Komar AA (2007) Silent SNPs: Impact on gene function and phenotype. *Pharmacogenomics* 8:1075-1080.
- Kryazhimskiy S and Plotkin JB (2008) The population genetics of dN/dS. *PLoS Genetics* 4:e1000304.
- Matsushima N, Enkhbayar P, Kamiya M, Osaki M and Kretsinger RH (2005) Leucine-rich repeats (LRRs): Structure, function, evolution and interaction with ligands. *Drug Des Rev* 2:305-322.
- Matsushima N, Tanaka T, Enkhbayar P, Mikami T, Taga M, Yamada K and Kuroki Y (2007) Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* 8:e124.
- Medzhitov R and Janeway Jr C (2000) Innate immunity. *N Engl J Med* 343:338-344.
- Medzhitov R, Preston-Hurlburt P and Janeway Jr CA (1997) A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* 388:394-397.
- Merx S, Zimmer W, Neumaier M and Ahmad-Nejad P (2006) Characterization and functional investigation of single nucleotide polymorphisms (SNPs) in the human TLR5 gene. *Hum Mutat* 27:293-299.
- Muzio M, Polentarutti N, Bosisio D, Prahladan MK and Mantovani A (2000) Toll-like receptors: A growing family of immune receptors that are differentially expressed and regulated by different leukocytes. *J Leukoc Biol* 67:450-456.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 pp.
- Poltorak A, He X, Smirnov I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, *et al.* (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene. *Science* 282:2085-2088.
- Sharp PM, Tuohy TM and Mosurski KR (1986) Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res* 14:5125-5143.
- Tajima F (1989) Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Takeda K, Kaisho T and Akira S (2002) Toll-like receptors. *Annu Rev Immunol* 21:335-376.
- Wang D, Johnson AD, Papp AC, Kroetz DL and Sadee W (2005) Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C > T affects mRNA stability. *Pharmacogenet Genomics* 15:693-704.
- Wei T, Gong J, Rössle SC, Jamitzky F, Heckl WM and Stark RW (2011) A leucine-rich repeat assembly approach for homology modeling of the human TLR5-10 and mouse TLR11-13 ectodomains. *J Mol Model* 17:27-36.
- Wlasiuk G, Khan S, Switzer WM and Nachman MW (2009) A history of recurrent positive selection at the toll-like receptor 5 in primates. *Mol Biol Evol* 26:937-949.
- Zhang D, Zhang G, Hayden MS, Greenblatt MB, Bussey C, Flavell RA and Ghosh S (2004) A toll-like receptor that prevents infection by uropathogenic bacteria. *Science* 303:1522-1526.

Internet Resources

SMART, <http://smart.embl-heidelberg.de/> (October 12, 2011).

Associate Editor: Alexandre Rodrigues Caetano

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.