Innate Immunity

High variability of TLR4 gene in different ethnic groups in Iran

Mihai Ioana, Bart Ferwerda, Shirin Farjadian, Luiza Ioana, Abbas Ghaderi, Marije Oosting, Leo AB Joosten, Jos WM van der Meer, Giovanni Romeo, Donata Luiselli, Dan Dediu and Mihai G Netea

The online version of this article can be found at:
http://ini.sagepub.com/content/18/3/492

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

International Endotoxin & Innate Immunity Society

Additional services and information for *Innate Immunity* can be found at:

Email Alerts: http://ini.sagepub.com/cgi/alerts

Subscriptions: http://ini.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://ini.sagepub.com/content/18/3/492.refs.html

>> Version of Record - May 28, 2012

OnlineFirst Version of Record - Oct 18, 2011

OnlineFirst Version of Record - Oct 12, 2011

What is This?

High variability of TLR4 gene in different ethnic groups in Iran

Mihai Ioana[1,2,3]*, Bart Ferwerda[1,4]*, Shirin Farjadian[5,6], Luiza Ioana[1], Abbas Ghaderi[7], Marije Oosting[1], Leo AB Joosten[1], Jos WM van der Meer[1], Giovanni Romeo[8], Donata Luiselli[9], Dan Dediu[10] and Mihai G Netea[1]

Abstract

Infectious diseases exert a constant evolutionary pressure on the innate immunity genes. TLR4, an important member of the TLR family, specifically recognizes conserved structures of various infectious pathogens. Two functional TLR4 polymorphisms, Asp299Gly and Thr399Ile, modulate innate host defense against infections, and their prevalence between various populations has been proposed to be influenced by local infectious pressures. If this assumption is true, strong local infectious pressures would lead to a homogeneous pattern of these ancient TLR4 polymorphisms in geographically-close populations, while a weak selection or genetic drift may result in a diverse pattern. We evaluated TLR4 polymorphisms in 15 ethnic groups in Iran, to assess whether infections exerted selective pressures on different haplotypes containing these variants. The Iranian subpopulations displayed a heterogeneous pattern of TLR4 polymorphisms, comprising various percentages of Asp299Gly and Thr399Ile, alone or in combination. The Iranian sample, as a whole, showed an intermediate mixed pattern when compared with commonly-found patterns in Africa, Europe, Eastern Asia and the Americas. These findings suggest a weak, or absent, selection pressure on TLR4 polymorphisms in the Middle-East that does not support the assumption of an important role of these polymorphisms in the host defense against local pathogens.

Introduction

The innate immune system is constantly exposed to pressures from infectious diseases, suggesting that one of the evolutionary selective forces shaping our genome during human history is represented by infectious pathogens.[1,2] Among immunity genes, positive natural selection, and especially balancing selection, is relatively common compared with other functional gene classes.[1] The initial step in the initiation of an immune response is represented by specific recognition of conserved structures of bacteria, viruses, fungi and protozoa by so-called pattern recognition receptors

[1]Radboud University Nijmegen Medical Centre, and Nijmegen Institute for Infection, Inflammation and Immunity (N4i), Department of Medicine, Nijmegen, The Netherlands
[2]Department of Molecular and Cellular Biology, University of Medicine and Pharmacy of Craiova, Romania
[3]University of Medicine and Pharmacy 'Carol Davila' Bucharest, Romania
[4]Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA
[5]Department of Immunology, Shiraz University of Medical Sciences, Shiraz, Iran
[6]Allergy Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

[7]Shiraz Institute for Cancer Research, Shiraz University of Medical Sciences, Shiraz, Iran
[8]Unità di Genetica Medica, Policlinico Universitario S. Orsola-Malpighi, Bologna, Italy
[9]Dipartimento di Biologia Evoluzionistica Sperimentale, Fisiche e Naturali, Università di Bologna, Italy
[10]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

*These authors contributed equally to this article.

Corresponding author:
Mihai Netea, Department of Internal Medicine (463), Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.
Email: m.netea@aig.umcn.nl

(PRRs).[3] The most studied PRRs are the TLRs.[4] Among them, TLR4, encoded by the gene with the same name located on chromosome 9, is the master receptor for the LPS component of Gram-negative bacteria, but it also recognizes other pathogen-associated molecular patterns (PAMPs) from mycobacteria, fungi, viruses and even parasites, such as malaria.[5–8]

So far, more then 35 TLR4 polymorphisms have been described,[9] among which the most studied are two non-synonymous single-nucleotide polymorphisms (SNPs) located in the leucine-rich repeat domain responsible for ligand recognition: an A/G transition at SNP rs4986790 (896A/G) that causes an Asp/Gly amino acid change at position 299 of the molecule, and a C/T transition at SNP rs4986791 (1196C/T) that causes a Thr/Ile amino acid change at position 399. These mutations affect the ligand-binding region (Asp299Gly) of TLR4 and the co-receptor-binding region (Thr399Ile) of the receptor respectively.[10,11] It has been shown that these TLR4 polymorphisms have important functional consequences related to the production of pro- and anti-inflammatory cytokines. In addition, they modulate the systemic inflammatory response syndrome in septic shock[12] and influence susceptibility to Gram-negative infections.[13]

Based on the prevalence in various populations around the globe, it has been shown that both these polymorphisms are ancient and occurred more than 65,000 years ago in Africa, before the migration of *Homo sapiens* out of Africa. Important differences have been described in the prevalence of TLR4 polymorphisms in various populations, possibly depending on local infectious pressure and population migration. The non-synonymous polymorphism Asp299Gly has a high prevalence in sub-Saharan Africa, and it has been proposed to have protective effects against mortality from malaria.[6] However, because of its effects in increasing susceptibility to severe bacterial infections, the TLR4 haplotype containing solely this polymorphism seems to have disappeared from Asians and Americans. In contrast, Asp299Gly has been found present in co-segregation with Thr399Ile in Europeans; this haplotype showing selective neutrality.[6,14]

Despite the progress in understanding the biology of TLR4, several important questions remain regarding the factors influencing the prevalence of TLR4 polymorphisms in various populations. One of the most intriguing questions to be answered is that of the degree to which the Asp299Gly and Thr399Ile SNPs (and especially of the haplotype containing both mutations) have influenced susceptibility to infections (especially in Europe and West Asia), with two possible scenarios: one in which these TLR4 polymorphisms strongly influenced susceptibility to infections and, subsequently, their prevalence was under selective pressure, and another one in which they had little influence of infection susceptibility and their prevalence in the

Eurasian landmass was mainly influenced by genetic drift, as previously proposed.[14] One approach to assess whether infections exerted selective pressure on the TLR4 variants is to investigate these polymorphisms in populations of different ethnic origins that have been living in the same geographical location for a long period of time and under the same infectious pressure. One would expect that in case of strong infectious pressure, the prevalence of ancient polymorphisms such as these TLR4 SNPs and haplotypes would become similar in the populations, irrespective of their ethnicity. The Middle East, and especially Iran, is an ideal target for such a study, considering its rich ethnic diversity and its key location on the routes of migration during the out of Africa human migration.

Materials and methods

Description of the Iranian subpopulations

A total of 738 individuals were recruited from 15 different ethnic and cultural groups from Iran: Pars, Armenian, Gilak, Mazendarani, Baloch, Kurd, Zoroastrian, Lur of Luristan, Lur of Kohgiluyeh and Buyer-Ahmad, Turkmen, Qashquaee, Azeri, Jew, People of Qeshm and Arab. Eight individuals were excluded because of missing data. The ethno-linguistic diversity in Iran, reflected by the important number of languages (77, 75 of which are living languages and 2 may only recently have become extinct),[15] is well represented among our sampled groups that belong to three major linguistic families: Indo-European (Indo-Iranian and Armenian subfamilies), Altaic (Turkic subfamily) and Afro-Asiatic (Semitic subfamily). A few of these populations are present in Iran only (e.g. Gilak and Mazendarani), but most of them are also found in other Middle-Eastern locations (Table 1).

Sample collection

Blood samples were collected from unrelated healthy volunteers after obtaining informed consent. All selected participants were self-reported third-generation members of a specific ethnic group, as described above. Those individuals of mixed ancestry or from mixed marriages were excluded. The evaluated samples were collected from different Iran provinces as shown in Figure 1A.

Genotyping

DNA samples were extracted from whole blood using the salting out method.[16] To screen for the TLR4 polymorphisms Asp299Gly (rs4986790) and Thr399Ile (rs4986791), the amplified sequences were digested with the restriction enzymes Nco-I and Hinf-I (New England BioLabs, Beverly, MA, USA) and separated

Table 1. List of the 15 Iranian subpopulations evaluated in this study

Ethnic group	N	Middle East location	Language family
Pars (Pa)	66	Iran, United Arab Emirates, Bahrain, Iraq, Kuwait, Oman, Afghanistan, Pakistan, Uzbekistan, Tajikistan	Indo-European
Armenian (Am)	64	Armenia, Georgia, Iran, Lebanon, Israel, Syria, Jordan, Iraq	
Gilak (Gi)	23	Iran (South-Eastern Caspian Sea-coast)	
Mazendarani (Ma)	51	Iran (South-Eastern Caspian Sea-coast)	
Baloch (Ba)	64	Iran, Pakistan, Afghanistan	
Kurd (Ku)	47	Iran, Iraq, Turkey, Syria, Armenia, Georgia, Lebanon, Azerbaijan, Turkmenistan	
Zoroastrian (Zo)	50	India, Pakistan, Iran	
Lur (LL)	28	Iran (South-West)	
Lur (LKB)	27	Iran (South-West)	
Turkmen (Tu)	53	Turkmenistan, Afghanistan, Iraq, Iran, Pakistan	Altaic/Turkic
Qashquaee (Qa)	44	Iran (South-West)	
Azeri (Az)	51	Azerbaijan, Iran, Turkey, Georgia, Kazakhstan	
Jew (Je)	60	Israel, Iran, Iraq, Georgia, Bahrain, Yemen, Lebanon, Syria, Afghanistan, Ethiopia	Afro-Asiatic/Semitic
People of Qeshm (Qe)	55	Qeshm Island (Southern coast of Iran)	Mixed
Arab (Ar)	47	Egypt, Lebanon, Saudi-Arabia, Iraq, Yemen, Syria, Sudan, Somalia, Jordan, Kuwait, UAE, Oman, Qatar, Bahrain, Israel, Iran, Turkey, Pakistan	Afro-Asiatic/Semitic
Total	*730*		

on agarose gel 2% stained with ethidium bromide. PCR was performed as described by Van Der Graaf et al.[17]

Single-nucleotide polymorphisms and haplotype data analysis

Data analysis used the open-source statistical environment R[18] and several of its packages (available for free download at: www.r-project.org). The linkage disequilibrium between the two loci (D' and r)[19] was computed using the LD function in package genetics,[20] and the deviation from Hardy-Weinberg equilibrium (HWE) for each locus was tested using function HWE.test in the same package. Genotype and haplotype frequencies between populations were compared using Fisher's exact test with 10,000 permutations as implemented by the Fisher test in the statistics package,[18] as well as G-tests[21] using function test.g in package hierfstat,[22] also with 10,000 permutations. We computed Wright's fixation index F_{ST}[23] between populations using calcFst in package polysat[24] and used these as genetic distances between populations. In order to search for structure in the genetic data, we projected the F_{ST} distances in two dimensions using classical multidimensional scaling (function cmdscale in package stats)[18] and we also conducted hierarchical agglomerative clustering (hclust in the same package). The geographic distances were computed as great circle distances using function distance in package argosfilter.[25] The Mantel correlations[26] between genetic and geographic distances were computed with function mantel in package vegan[27] using 99,999 permutations. Geneland[28–30] searches for population structure using genetic and geographic data using a Bayesian approach; we used this package in an attempt to identify patterns in our samples [with and without geographic structure, 5,000,000 Markov Chain Monte Carlo (MCMC) iterations].

The molecular genetic variation within, and among, populations was tested using different implementations of Analysis of Molecular Variance (AMOVA) as given in the R packages: pegas (function amova),[31] ade4 (functions amova and randtest)[32] and vegan (function adonis).[33,34]

When necessary, the P-values were adjusted for multiple comparisons using Holm's correction,[35] as implemented by function p.adjust in package stats.[18] Statistical significance was accepted for P-values < 0.05.

Results

TLR4 haplotype frequencies for the overall Iranian sample and for each of the studied Iranian subpopulations

Iranian TLR4 haplotype frequency pattern and the haplotype frequencies for all the studied groups are shown in Table 2 and Figure 1. The calculated frequencies of the minor alleles for rs4986790 was 0.05, whereas
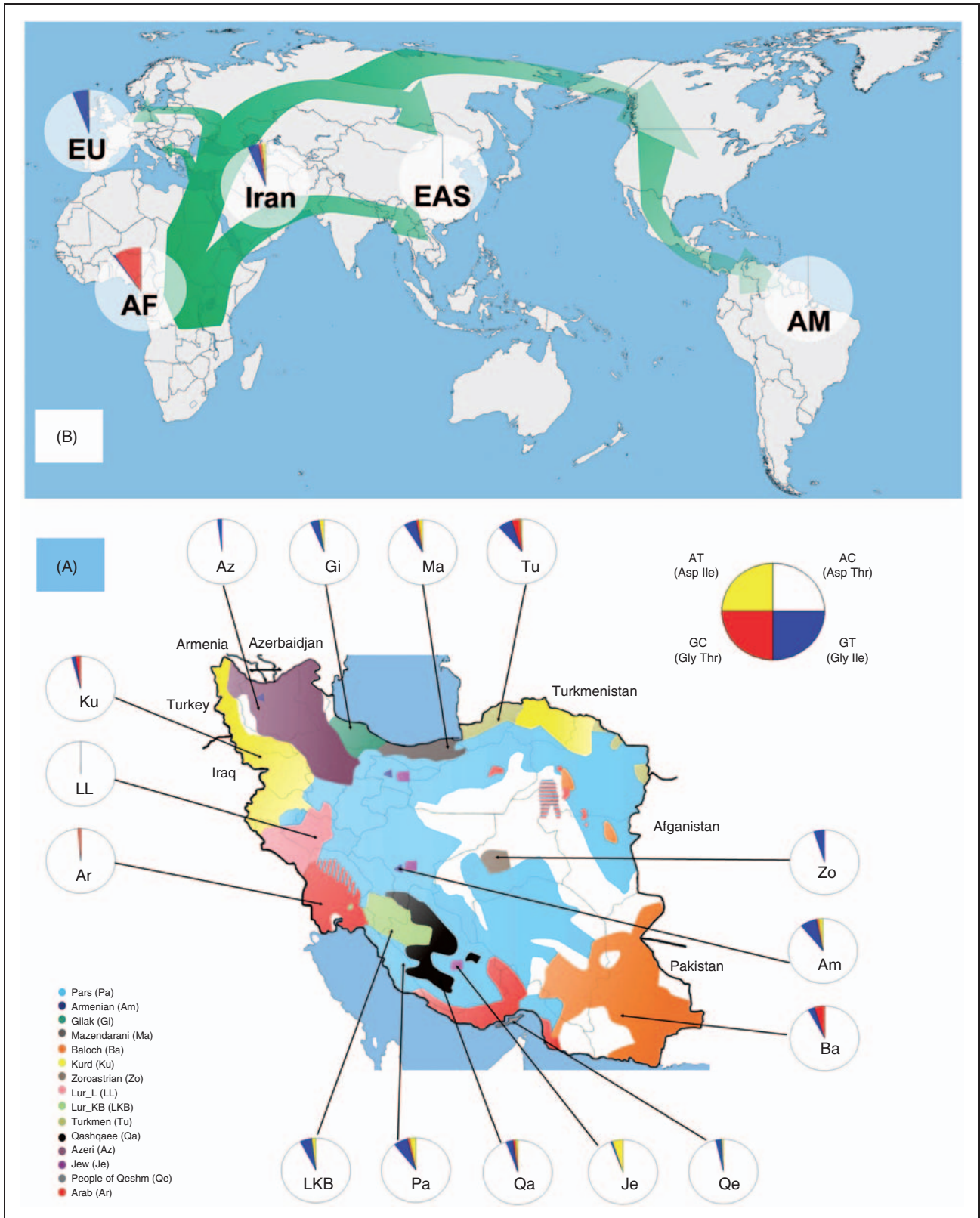
Figure 1. (A) Iranian distribution of TLR4 haplotypes among the 15 groups and the main locations of these groups in Iran. Approximate geographic sampling location of each group is indicated by black arrows. Circles indicate allele frequency (red, Asp299Gly; yellow, Thr399Ile; blue, Asp299Gly/Thr399Ile). (B) The most likely paths of migration during the 'out of Africa' migration of modern humans and the world distribution pattern of TLR4 haplotypes in Europe (EU), Africa (AF), East-Asia (EAS) and the Americas (AM), and the key location of Iran on the route.

Table 2. TLR 4 896A/G (Asp299Gly) and 1196C/T (Thr399Ile) haplotype frequencies in Iran and in the various ethnic groups

Populations	Haplotype frequencies			
	AC (AspThr)	GT (GlyIle)	GC (GlyThr)	AT (AspIle)
Pars	0.894	0.068	0.008	0.030
Armenian	0.891	0.078	0.008	0.023
Gilak	0.935	0.043	-	0.022
Mazendarani	0.910	0.060	0.010	0.020
Baloch	0.922	0.031	0.047	-
Kurd	0.958	0.021	0.021	-
Zoroastrian	0.950	0.050	-	-
Lur (Luristan)	1.000	-	-	-
Lur (Kohgiluyeh and Buyer Ahmad)	0.925	0.056	-	0.019
Turkmen	0.887	0.066	0.038	0.009
Qashquaee	0.944	0.034	0.011	0.011
Azeri	0.980	0.020	-	-
Jew	0.942	0.008	-	0.050
People of Qeshm	0.964	0.027	-	0.009
Arab	0.989	-	0.011	-
Iran	0.936	0.039	0.012	0.013

Table 3. The overall Iranian distribution of 896A/G and 1196C/T diplotypes

1196 C/T (Thr399Ile)	896A/G (Asp299Gly)		
	AA	AG	GG
CC	641 (87.81)	16 (2.19)	-
CT	18 (2.46)	50 (6.85)	1 (0.14)
TT	1 (0.14)	-	3 (0.41)

Data are presented as number (%) of diplotypes in Iranian individuals.

for rs4986791 it was 0.018. Fisher's exact test showed no deviation from Hardy–Weinberg equilibrium for both evaluated SNPs, both in the overall Iranian sample (*P*-values of 0.1221 for Asp299Gly and 0.1322 for Thr399Ile) and in each individual group (data not shown). We observed a high degree of linkage disequilibrium between the rs4986790 and rs4986791 SNPs, both in the whole sample (D' = 0.75, r = 0.74, $\chi^2_{(730)} = 798.72$, *P* = 0) and in each population separately, confirming the reported co-segregation (data not shown).[36]

The differences between populations for rs4986790 did not reach statistical significance, although a tendency towards significance was observed (Fisher's exact test with 10,000 permutations, *P* = 0.066). In contrast, a statistically significant difference between populations was found for the rs4986791 SNP (*P* = 0.010). Table 3 shows the diplotype distribution among all included individuals. When assessing the distribution of TLR4 haplotypes based on these two SNPs, significant differences between populations (Fisher's exact

test with 10,000 permutations *P* = 0.022) seemed to be present. However, although several pairs of populations did seem to differ at first sight (Table 4 shows uncorrected Fisher's exact test *P*-values), many of these paired tests did not survive Holm's multiple testing correction (data not shown).

To evaluate the significance of genetic differences, we also used G-tests [overall: 10,000 permutations, G = 69.24, *P* = 0.0024 for each pair of populations, see supplementary data (supplementary Table 1)]. As in the case of Fisher's exact test, no paired test survived Holm's multiple testing correction. These results were confirmed by AMOVA. The values of F_{ST} genetic distances between population pairs (supplementary Table 2) are graphically represented in Figure 2 (white = min, black = max). It can be seen that Lurs of Luristan and Arabs are the most different, but the differences are relatively small (min = 0.0, max = 0.039, mean = 0.0092, median = 0.0056) and statistically not significant. Multidimensional scaling (MDS),[37] a method for visually

Table 4. Fisher's exact test for TLR4 diplotype differentiation between populations. (uncorrected *P*-values; no test survives Holm's multiple testing correction)

	Pa	Am	Gi	Ma	Ba	Ku	Zo	LL	LKB	Tu	Qa	Az	Je	Qe	Ar
Pa	NA														
Am	0.984	NA													
Gi	I	0.934	NA												
Ma	0.982	I	I	NA											
Ba	0.020	0.018	0.124	0.044	NA										
Ku	0.192	0.167	0.342	0.297	0.709	NA									
Zo	0.273	0.346	0.448	0.419	0.028	0.217	NA								
LL	0.147	0.131	0.085	0.135	0.299	0.408	0.161	NA							
LKB	I	I	I	I	0.070	0.289	0.512	0.052	NA						
Tu	0.477	0.618	0.842	0.813	0.171	0.493	0.268	0.115	0.801	NA					
Qa	0.819	0.731	I	0.846	0.270	0.745	0.529	0.516	0.904	0.789	NA				
Az	0.108	0.068	0.242	0.078	0.071	0.446	0.436	0.535	0.143	0.060	0.425	NA			
Je	0.134	0.058	0.542	0.087	0.006	0.082	0.039	0.676	0.216	0.017	0.327	0.154	NA		
Qe	0.331	0.296	0.507	0.396	0.049	0.445	0.480	0.709	0.356	0.161	0.840	I	0.280	NA	
Ar	0.012	0.007	0.057	0.015	0.133	0.423	0.055	I	0.023	0.015	0.146	0.350	0.109	0.206	NA
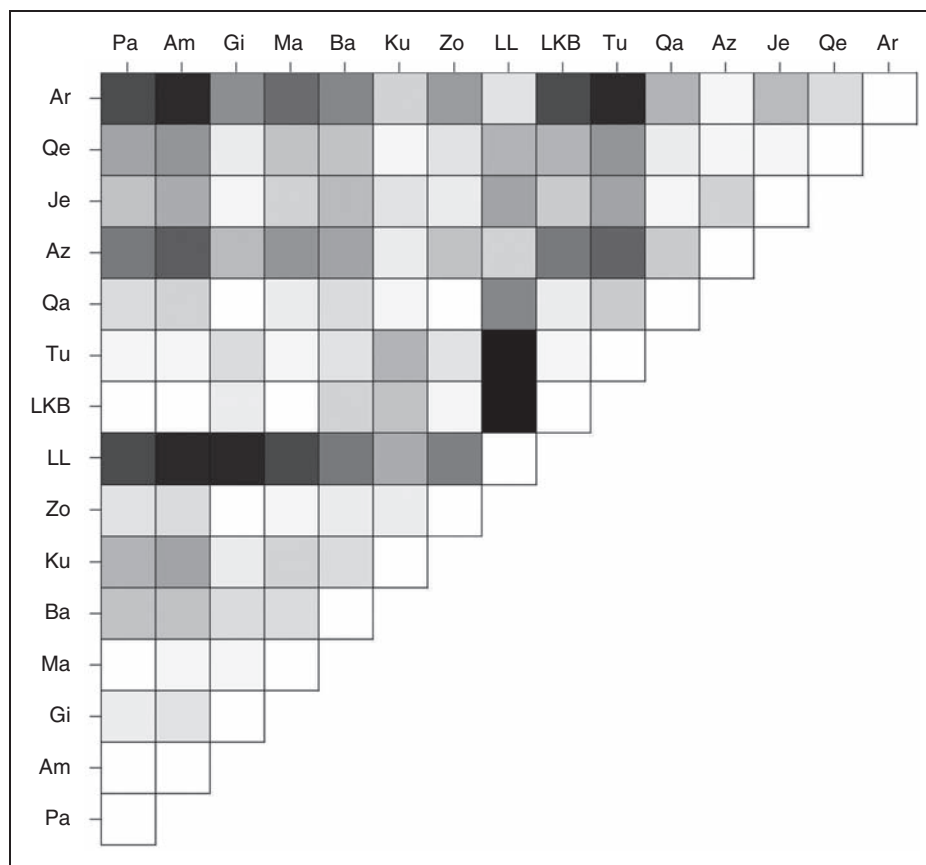
Figure 2. Graphic representation of the F_{ST} distances (white, min; black, max).

representing distance matrices, was used it to plot the pair-wise F_{ST} distances between the 15 studied populations in order to identify any patterns and clusters in the genetic data (Figure 3).

We also executed 10 independent runs of Geneland using only the genetic information, with 1,000,000 MCMC generations,[38] a thinning of 100 and burn-in of 2000, and all converged to classifying all individuals
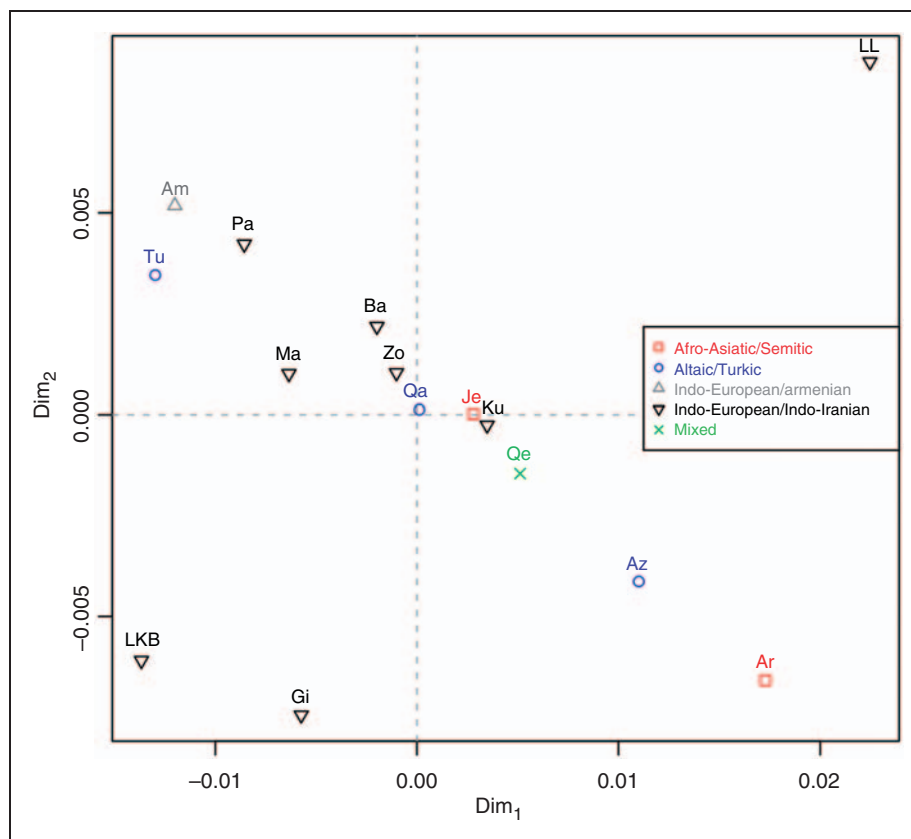
Figure 3. MDS plot of pairwise FST distances in two dimensions. Populations are classified in terms of the language family/subfamily (for Queshm, no such classification was possible).

into a single population. Using F_{ST} to create a hierarchical clustering of the populations resulted in the phenogram represented in Figure 4. Supporting the finding that the genetic structure of these loci does not reflect geography was the very low and non-significant Mantel correlation between geographic distances (great circle distances) and F_{ST} distances between all pairs of populations: $r = -0.014$, $P = 0.49$ (9999 permutations).

TLR4 polymorphisms in different Iranian ethnic groups with language as a grouping factor

Moreover, to investigate the possible differences between populations speaking languages belonging to different language families, we amalgamated all sampled individuals into three 'language groups' (L-groups): Afro-Asiatic (107 individuals), Altaic (148) and Indo-European (420) (the individuals from Qeshm have been excluded from the analysis because of their uncertain linguistic status). Comparing groups based on linguistic affiliation can, potentially, provide insights into processes occurring over several thousand of years, including genetic drift and selection pressures. Both SNPs were in HWE and strong LD in all three L-groups, except Afro-Asiatic where the LD was less pronounced (D' = 0.47, r = 0.25) but still highly significant ($P = 0.0003$).

Genotype frequencies for rs4986790 seemed to differ between L-groups (Fisher's exact test with 10,000 permutations, $P = 0.016$) but not for rs4986791 ($P = 0.21$) differentiation confirmed by the overall g-test ($P = 0.0126$) and as a result of Afro-Asiatic being more different than the other two L-groups, even after Holm's multiple testing correction (Table 5). Haplotype frequencies showed a similar pattern of difference between L-groups (Fisher's exact test with 10,000 permutations $P = 0.039$) with Afro-Asiatic being the outlier, even after Holm's correction (supplementary Table 3). Difference was also reflected by the greater F_{ST} distances involving this L-group (supplementary Table 4). Conducting an AMOVA with language as a grouping factor suggested that there was variation within samples ($P = 0.01699830$) and marginally between samples ($P = 0.06109389$), but not among L-groups ($P = 0.28547145$).

Discussion

Owing to its Middle East location, Iran, the 18th largest country in the world, has a special geographic significance for the various human migrations on the Eurasian landmass, from that of *Homo sapiens* out of Africa,[39] to the Silk Road,[40] and to more recent periods in history. Iran is a diverse country with various
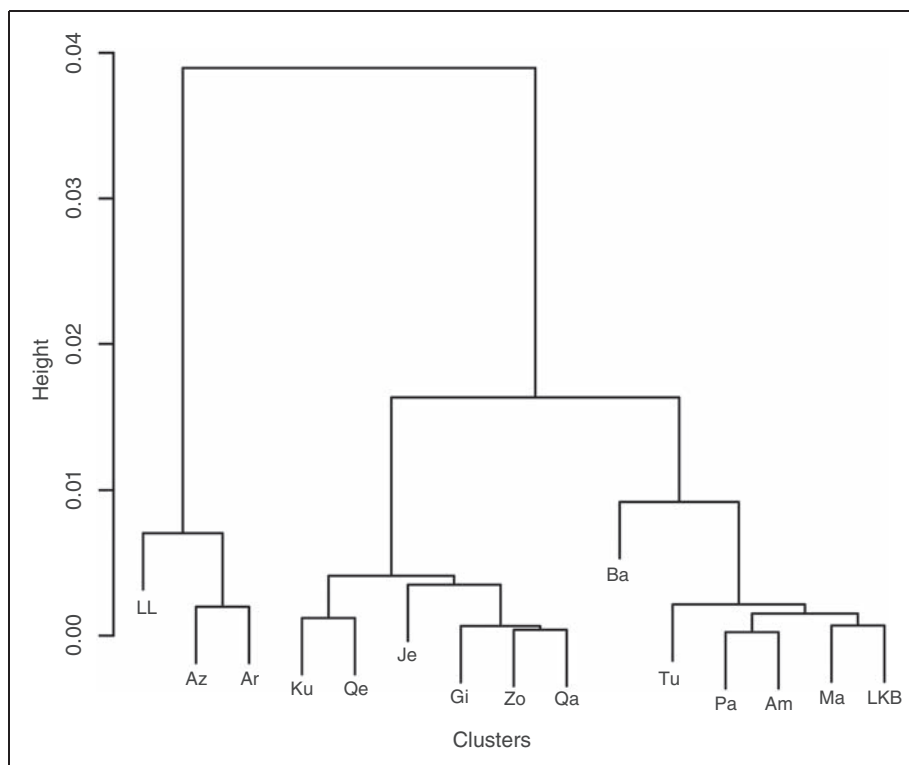
Figure 4. Hierarchical clustering of the populations based on genetic distances.

Table 5. Overall g-test values after multiple testing correction

	Afro-Asiatic	Altaic	Indo-European
Afro-Asiatic	NA		
Altaic	0.034	NA	
Indo-European	0.015	0.560	NA

geographical landforms, a wide range of climatic variation, as well as populations of different religions and ethno-linguistic backgrounds. Iran is home to a large number of different ethnic groups including Pars, Turk, Kurd, Arab and Baloch,[41] which have interacted during history with many other groups, such as Macedonians, Arabs, Turks and Mongols.[42] These characteristics are reflected in the 15 different ethno-religious groups collected from different provinces across the country. In the present study, we took advantage of the ethnic diversity of populations that live in Iran and assessed the prevalence and distribution of TLR4 polymorphisms among these diverse ethnic groups. We hypothesized that if TLR4 polymorphisms would strongly influence susceptibility to infections, the resulting infectious pressure would result in a homogeneous pattern of the ancient TLR4 polymorphisms (Asp299Gly and Thr399Ile) in these ethnically-different populations. In contrast, a situation in which TLR4 polymorphisms

would not have a strong effect on susceptibility to infections, thus resulting in weak selection pressures or even genetic drift, will have, as a consequence, a diverse pattern of TLR4 polymorphisms in the various populations.

The prevalence of the Asp299Gly and Thr399Ile TLR4 polymorphisms between Iranian subpopulations differed, as seen in Table 2 and Figure 1. Interestingly, the heterogeneity of TLR4 polymorphisms and haplotypes in the Iranian subpopulations studied was greater than on the African continent. In populations from both East and West Africa, a homogenous pattern of TLR4 polymorphisms was seen, characterized by the presence of 5–15% individuals bearing the Asp299Gly SNP, and a much smaller group of individuals bearing Asp299Gly/Thr399Ile in linkage, while no Thr399Ile polymorphism is present alone. It has been suggested that this homogeneous distribution of TLR4 polymorphisms may be caused by the protective effect against severe malaria.[6,14] In contrast, it is less clear whether TLR4 polymorphisms also influence susceptibility to infections in the colder climates of Europe and Asia. The TLR4 299Gly/399Ile haplotype comprising both polymorphisms does not seem to modify the response of monocytes to endotoxin and susceptibility to infections, and it has been proposed to be the result of genetic drift in populations from Europe and Asia.[14] If TLR4 polymorphisms would, indeed, not modify susceptibility to infections in the colder climates of

Eurasia, one would expect that the various ethnic populations studied here would maintain a high heterogeneity of these TLR4 polymorphisms. In line with this, the Iranian subpopulations displayed a heterogeneous pattern of TLR4 polymorphisms, comprising various percentages of Asp299Gly and Thr399Ile alone, or in combination. These differences are most probably a result of the specific geographic origin, natural borders and/or ethnic and religion barriers between the populations studied. In this respect, particular patterns can be observed for Baloch, Zoroastrian, Azeri, Jew, Arab, Kurd and Lurs of Luristan (see Figure 1A). The diversity of TLR4 polymorphisms in Iranian populations is reminiscent of the mixed pattern of TLR4 haplotypes described among Israeli subpopulations.[14] The variation of the prevalence of TLR4 polymorphisms in Iranian subpopulations living in close geographic proximity suggest a weak, or absent, infectious pressure on TLR4 SNPs in the Middle East. Evidence for the absence of a strong infectious pressure among most of the Iranian subpopulations included in this study was also found in human leukocyte antigen (HLA) class II genes.[43] A few recently published studies investigated HLA class II allele and haplotype frequencies in order to find the genetic relationship between the major Iranian subpopulations,[43] based on the fact that HLA data can be used to elucidate the genetic history of human populations.[44,45] Parsee, Zoroastrian and Baloch subpopulations were found to be mainly confined to intrapopulation variations, with little subdivision among studied Iranian populations.[41] A closer genetic relationship was present between Iranian Arabs and Iranian Jews when compared with either Iranian Arabs and Middle-Eastern Arabs or Iranian Jews and other Jews.[46] When comparing Kurds and Azeris, another two major ethnic groups in Iran, considerable similarities in HLA class II allele and haplotype distributions were seen, except for DQB1*0503, which was observed with a higher frequency in Kurds in comparison to Azeris.[47]

The Iranian subpopulations assessed in the present study displayed an interesting intermediate mixed pattern of TLR4 haplotypes when compared with commonly found patterns in Africa (higher GC, no AT, less GT),[6,14] Europe (very rare GC, lower AT, higher GT)[14,48–50] and Central/Eastern Asia[14,51–55] and Americas (no GC, no AT, no GT)[14] (see Figure 1B). The driving forces that may influence TLR4 haplotypes frequencies in humans may be represented by local infections for which TLR4 is important as a recognition receptor (e.g. Gram-negative bacteria, mycobacteria). As mentioned above, strong effects of the TLR4 SNPs on infection susceptibility would have resulted in a homogeneous haplotype pattern for the overall Iranian sample. However, this proves not to be the case: all haplotypes were present in various percentages in the Iranian populations, suggesting the absence of strong selective forces. This is in line with Barreiro et al.,[56] who showed a weak negative and/or balancing selection on extracellular TLRs, implicitly of TLR4. As suggested before, a different situation may be present in sub-Saharan Africa and additional studies involving ethnically-different populations living in close proximity should be performed also in warm climates.

To investigate the possible differences between Iranian subpopulations belonging to different language families, the prevalence of TLR4 polymorphisms was analyzed in different ethnic, groups with language as a grouping factor. We conclude that higher-level language classification explains a small part of the genetic variance in TLR4 and highlights Afro-Asiatic populations as more different than the other two language-groups; however, this may be owing to its internal inconsistency.

TLR 4 SNPs have been associated with susceptibility to certain infectious diseases[6,7,12,13] because of its role as a receptor for the LPS of Gram-negative bacteria,[8] mannans from fungi,[7] and cell wall structures of the *Plasmodium* parasite[6] and mycobacteria[5]. Innate immune system responses to specific infectious pressure may be reflected by important differences in the prevalence of TLR polymorphisms in populations, but these differences depend on the functional consequences of certain polymorphisms. The TLR4 299Gly allele has been associated with an increased cytokine response which seems to be associated with protection from severe malaria.[14] In contrast, the TLR4 haplotype containing both 299Gly and 399Ile does not seem to modify the function of the molecule.[14]

In conclusion, in the present study we describe the prevalence of Asp299Gly and Thr399Ile TLR4 polymorphisms in 15 ethnically-different populations from Iran. In contrast to the homogeneity of these polymorphisms in other populations, such as those from the African continent, the Iranian subpopulations display a broad heterogeneity of TLR4 Asp299Gly and Thr399Ile. These findings suggest a weak, or absent, effect of TLR4 polymorphisms on infection susceptibility in the Middle East. However, further studies employing an increased number of markers that are not subject to selective pressure are needed to more thoroughly assess these conclusions.

References

1. Ferrer-Admetlla A, Bosch E, Sikora M, et al. Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol* 2008; 181: 1315–1322.

2. Sabeti PC, Schaffner SF, Fry B, et al. Positive natural selection in the human lineage. *Science* 2006; 312: 1614–1620.

3. Janeway Jr, CA and Medzhitov R. Innate immune recognition. *Annu Rev Immunol* 2002; 20: 197–216.

4. Akira S and Takeda K. Toll-like receptor signalling. *Nat Rev Immunol* 2004; 4: 499–511.

5. Means TK, Wang S, Lien E, et al. Human toll-like receptors mediate cellular activation by *Mycobacterium tuberculosis*. *J Immunol* 1999; 163: 3920–3927.

6. Mockenhaupt FP, Cramer JP, Hamann L, et al. Toll-like receptor (TLR) polymorphisms in African children: Common TLR-4 variants predispose to severe malaria. *Proc Natl Acad Sci USA* 2006; 103: 177–182.

7. Netea MG, Van Der Graaf CA, Vonk AG, et al. The role of toll-like receptor (TLR) 2 and TLR4 in the host defense against disseminated candidiasis. *J Infect Dis* 2002; 185: 1483–1489.

8. Poltorak A, He X, Smirnova I, et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 1998; 282: 2085–2088.

9. Georgel P, Macquin C and Bahram S. The heterogeneous allelic repertoire of human toll-like receptor (TLR) genes. *PLoS One* 2009; 4: e7803.

10. Arbour NC, Lorenz E, Schutte BC, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet* 2000; 25: 187–191.

11. Rallabhandi P, Bell J, Boukhvalova MS, et al. Analysis of TLR4 polymorphic variants: new insights into TLR4/MD-2/CD14 stoichiometry, structure, and signaling. *J Immunol* 2006; 177: 322–332.

12. Lorenz E, Mira JP, Frees KL and Schwartz DA. Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock. *Arch Intern Med* 2002; 162: 1028–1032.

13. Agnese DM, Calvano JE, Hahm SJ, et al. Human toll-like receptor 4 mutations but not CD14 polymorphisms are associated with an increased risk of gram-negative infections. *J Infect Dis* 2002; 186: 1522–1525.

14. Ferwerda B, McCall MB, Alonso S, et al. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. *Proc Natl Acad Sci USA* 2007; 104: 16645–16650.

15. Lewis MP. *Ethnologue: languages of the world*, sixteenth edition. Dallas: SIL International, 2009.

16. Miller SA, Dykes DD and Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; 16: 1215.

17. Van der Graaf CA, Netea MG, Morre SA, et al. Toll-like receptor 4 Asp299Gly/Thr399Ile polymorphisms are a risk factor for *Candida* bloodstream infection. *Eur Cytokine Netw* 2006; 17: 29–34.

18. R Development Core Team R: A Language and Environment for Statistical Computing. Vienna, 2010.

19. Devlin B and Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 1995; 29: 311–322.

20. Warnes G, Gorjanc G, Leisch F and Man M. Genetics: Population Genetics. *R package. version 1.3.4* 2008.

21. Goudet J, Raymond M, de Meeus T and Rousset F. Testing differentiation in diploid populations. *Genetics* 1996; 144: 1933–1940.

22. Goudet J. hierfstat: Estimation and tests of hierarchical F-statistics. *R package version 0.04-4* 2006.

23. Wright S. The genetical structure of populations. *Ann Eugenics* 1951; 15: 323–354.

24. Clark LV and Jasieniuk M. POLYSAT: an R package for polyploid microsatellite analysis. *Mol Ecol Resour* 11: 562–566.

25. Freitas F. Argosfilter: Argos locations filter. *R package version 0.62* 2010.

26. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res* 1967; 27: 209–220.

27. Oksanen J, Blanchet FG and Kindt R. vegan: Community Ecology Package. *R package version* 2010; 1: 17–23.

28. Guillot G. Inference of structure in subdivided populations at low levels of genetic differentiation-the correlated allele frequencies model revisited. *Bioinformatics* 2008; 24: 2222–2228.

29. Guillot G, Estoup A, Mortier F and Cosson JF. A spatial statistical model for landscape genetics. *Genetics* 2005; 170: 1261–1280.

30. Guillot G, Santos F and Estoup A. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 2008; 24: 1406–1407.

31. Paradis E. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419–420.

32. Excoffier L, Smouse PE and Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 1992; 131: 479–491.

33. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 2001; 26: 32–46.

34. Legendre P and Anderson MJ. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol Monogr* 1999; 69: 1–24.

35. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979; 6: 65–70.

36. Smirnova I, Hamblin MT, McBride C, et al. Excess of rare amino acid polymorphisms in the Toll-like receptor 4 in humans. *Genetics* 2001; 158: 1657–1664.

37. Cox TF and Cox MAA. *Multidimensional scaling*. London: Chapman and Hall, 1994, p.213.

38. Gamerman D. Markov Chain Monte Carlo: Stochastic Simulation of Bayesian Inference. In: Chatfield C and Zidek JV (eds) *Texts in statistical science*. London: Chapman & Hall, 1997.

39. Wells S and Read M. *The journey of man: a genetic Odyssey*. New Jersey: Princeton University Press, 2002, p.75.

40. Boulonois I. *The silk road*. London: Allen & Unwin, 1966.

41. Farjadian S, Moqadam FA and Ghaderi A. HLA class II gene polymorphism in Parsees and Zoroastrians of Iran. *Int J Immunogenet* 2006; 33: 185–191.

42. Lockwood W. *A panorama of Indo-European languages*. London: Hutchinson Ltd, 1972.

43. Farjadian S, Ota M, Inoko H and Ghaderi A. The genetic relationship among Iranian ethnic groups: an anthropological view based on HLA class II gene polymorphism. *Mol Biol Rep* 2009; 36: 1943–1950.

44. Uinuk-Ool TS, Takezaki N, Derbeneva OA, et al. Variation of HLA class II genes in the Nganasan and Ket, two aboriginal Siberian populations. *Eur J Immunogenet* 2004; 31: 43–51.

45. Uinuk-Ool TS, Takezaki N, Sukernik RI, et al. Origin and affinities of indigenous Siberian populations as revealed by HLA class II gene frequencies. *Hum Genet* 2002; 110: 209–226.

46. Farjadian S and Ghaderi A. HLA class II genetic diversity in Arabs and Jews of Iran. *Iran J Immunol* 2007; 4: 85–93.

47. Farjadian S and Ghaderi A. HLA class II similarities in Iranian Kurds and Azeris. *Int J Immunogenet* 2007; 34: 457–463.

48. Buraczynska M, Baranowicz-Gaszczyk I, Tarach J and Ksiazek A. Toll-like receptor 4 gene polymorphism and early onset of diabetic retinopathy in patients with type 2 diabetes. *Hum Immunol* 2009; 70: 121–124.

49. Koch W, Hoppmann P, Pfeufer A, et al. Toll-like receptor 4 gene polymorphisms and myocardial infarction: no association in a Caucasian population. *Eur Heart J* 2006; 27: 2524–2529.

50. Schurmann M, Kwiatkowski R, Albrecht M, et al. Study of Toll-like receptor gene loci in sarcoidosis. *Clin Exp Immunol* 2008; 152: 423–431.
51. Hsiao CH, Wei SC, Wong JM, et al. Pediatric Crohn disease: clinical and genetic characteristics in Taiwan. *J Pediatr Gastroenterol Nutr* 2007; 44: 342–346.
52. Kim YS, Hwang YJ, Kim SY, et al. Rarity of TLR4 Asp299Gly and Thr399Ile polymorphisms in the Korean population. *Yonsei Med J* 2008; 49: 58–62.
53. Liang XH, Cheung W, Heng CK and Wang DY. Absence of the toll-like receptor 4 gene polymorphisms Asp299Gly and Thr399Ile in Singaporean Chinese. *Ther Clin Risk Manag* 2005; 1: 243–246.
54. Tahara T, Arisawa T, Shibata T, et al. Absence of common polymorphisms of Toll like receptor 4 (TLR4): Asp299Gly, Thr399Ile in patients with gastroduodenal diseases in Japan. *J Clin Biochem Nutr* 2007; 40: 62–65.
55. Zheng B, Li Q, Wei C, et al. Lack of association of TLR4 gene Asp299Gly and Thr399Ile polymorphisms with rheumatoid arthritis in Chinese Han population of Yunnan Province. *Rheumatol Int* 30: 1249–1252.
56. Barreiro LB, Ben-Ali M, Quach H, et al. Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS Genet* 2009; 5: e1000562.