International Journal of Pharmacy and Biological Sciences
ISSN: 2321-3272 (Print), ISSN: 2230-7605 (Online)
IJPBS | Volume 7 | Issue 4 | OCT-DEC| 2017 | 255-262
Research Article | Biological Sciences | Open Access | UGC Approved | MCI Approved Journal

GENETIC DIVERSITY OF TEN TLR GENES IN SOME ETHNIC POPULATIONS OF NORTH BENGAL REGION OF INDIA-A SYSTEMIC STUDY

Avishek Das, Pokhraj Guha and Tapas Kumar Chaudhuri*

Cellular Immunology Laboratory, Department of Zoology, University of North Bengal, Raja Rammohunpur, Siliguri, West Bengal 734013, India–

***Corresponding Author Email: dr_tkc_nbu@rediffmail.com**

ABSTRACT

Objectives: *Toll like Receptors (TLRs) are conserved transmembrane proteins that recognize Pathogen Associated Molecular patterns (PAMPs) and initiate innate immune system functions.* ***Method:*** *Hereby, we have aimed to study the diversity and frequency distribution of 10 TLR genes in the three ethnic populations of North Bengal namely Rabha, Gurkha and Muslim. We took 315 samples of which 125 Gurkhas, 140 Muslims and 50 Rabha samples. PCR-SSP was done for all the ten human TLR genes. We have constructed the phylogenetic tree, Principal component analysis (PCA) and Genetic distance for all three populations.* ***Results:*** *We have documented the highest frequency of TLR4 among the Rabhas (0.90) and Gurkhas (0.96) while TLR5 (0.97) have the highest frequency among the Muslims. Based on TLR frequencies the genetic distances were calculated which revealed that Rabha and Muslim are distantly related (0.089) while Gurkha and Muslim (0.023) are much closer. This observation is interesting as both Rabha and Gurkha are East-Asian origin while Muslim belongs to the Middle East lineage.* ***Conclusion:*** *This may be because of the effect of the environment in combination with the pathogens present in that environment, as TLR is mainly responsible to participate in innate immune response. Convergent evolution also plays a significant role in shaping the three populations in this region inspite of their different ethnicity. This study is a pioneering report on population based TLR frequency distribution in North Bengal region.*

KEY WORDS

Toll Like Receptors, Pathogen Associated Molecular Pattern, Genetic distance, East-Asian, Principal Component analysis.

INTRODUCTION

Infectious pathogens act as the major force for generating selection pressure in human evolutionary history [1]. Migrations of humans to different parts of the world resulted in the exposure of the immune response genes to local pathogenic antigens and thus get modified as per environmental demands. Thus, pressures exerted by the local pathogens cause positive selections of some genetic markers in the population for developing protection against the pathogens [1]. TLRs are genetically conserved pattern recognition receptors (PRR) which are capable of recognizing diverse sets of conserved antigens [2]. Ten members constitute this receptor family in both human with their respective genes located in different chromosomes of the human genome [3]. Majority of the TLR molecules are cell surface receptors which includes TLR1, 2 4 6 8 and 10 respectively, while others are endosomal in nature (TLR 3, 7 and 9). The members of the TLR family are capable of recognizing different conserved antigens like lipopolysaccharide, flagellin, CpG DNA and even double stranded RNA [2] and may also regulate the susceptibility of a population to pathogenic invasions and disease progression [4] [5].

Indian population comprises of various religions, tribes and castes each having their unique socio-cultural and ethnic background, most of which are strictly endogamous [6]. The Indian subcontinent has experienced several human migration events. One such major event includes the migration of Indo-European-speaking people from West Eurasia whose admixture with indigenous Dravidian populations led to the subsequent establishment of the Hindu caste system [7]. Such extensive admixture and enormous genetic diversity among the Indians make them primary target for genetic diversity analyses [7] [8].

The Rabha population is a very small ethnic tribal population inhabitating the Eastern Terai and Dooars regions of northern part of West Bengal [8] [9]. Historical evidences suggest their East-Asian origin [10]. They have their own socio-cultural and linguistic heritage and are considered as an important tribal population of the state as well as the country. On the other hand, Gurkhas constitute the major inhabitants of the hilly region of North Bengal. They are very hard working and courageous people. They have unique cultures and traditions which make them an important subject of population genetics study. Another population having a very interesting historical background is the Muslim population of West Bengal constituting 27% of the total population of the state [9]. Recent studies have documented the admixture of the Indian Muslim populations with the local Hindu residents resulting in differential ancestral patterns in different parts of India [11] [12] [13].

In this study, we have aimed to study the TLR genetic profile in Gurkha, Rabha and Muslim populations and analyzed the role of TLRs in selection and phylogenetic analyses, if any.

MATERIAL AND METHODS

1 Study populations

The study population consisted of 125 Gurkhas, 140 Muslims and 50 Rabha samples. All blood samples were collected from Darjeeling, Coochbehar and Jalpaiguri districts (26° 20'- 27° 03' N and 88° 18'- 89° 29' E) of Northern West Bengal, India. The samples were collected on the basis of their ethnicity, caste and health conditions. Individuals having three generations of common pedigree were excluded from the analyses. All the donors were informed regarding the purpose of the study before the collection of the blood samples and written consents were provided by the volunteers. The investigation was approved by the Human ethics committee of the Department of Zoology, University of North Bengal (Zoo/4133/2011) and performed in accordance with the Declaration of Helsinki, 1964. 3 mL of blood sample was taken from each volunteer by vein puncture method under the guidance of a medical practitioner and was stored in EDTA containing vials at -20°C until use.

2 DNA extraction and TLR specific PCR- SSP typing

Genomic DNA was extracted from the samples by the standard Phenol- Chloroform extraction method with slight modifications. This was followed by PCR-SSP typing for all the 10 TLR genes [14] **(Table no. 1)**. Primers were designed based on the conserved sites of the ten human TLR genes using NCBI BLAST http://www.ncbi.nlm.nih.gov/tools/primer-blast/.

Primers were supplied by the Integrated DNA Technologies, Inc, Iowa, USA. Each PCR reaction mixture of 25µL volume contained 5X PCR buffer (Promega Corporation, Agora, Fitchburg Center, Fitchburg, Wisconsin), 5 µL of 10mM dNTPs, 1.5µL of 25mM MgCl2, 1.5µL of primers, and 1-1.5 U of Taq DNA polymerase. DNA samples were checked before use for their 260/280 absorbance and 1.5-2ul of 100ng DNA samples were then added to the PCR mixture. The reaction conditions for PCR consisted of an initial denaturation step of 94°C for 3 minutes followed by 30 cycles of 94° C for 30 s, 56.9° C for 50 s and 72°C for 1minutes and final extension of 72^{0}C for10 minutes respectively with slight modifications of annealing temperature for different primer sets. The PCR products were then analyzed on ethidium bromide prestained 2% agarose gel by electrophoresis. Samples were then visualized on UV transilluminator. All the lanes of the product loaded gel showed a control band, except for the negative control lane. The reactions were repeated in case of false reactions where no control bands were found.

3 Statistical analyses

Statistical data were analysed in Kyplot, MS-excel, GenAlEx (ver-6.5), SPSS (Ver-15.0) and MINITAB (Ver-6). Observed frequencies were performed by direct counting the number of the gene present by the total number of the sample. Gene frequencies and Chi- square data were also calculated. Correlation studies were carried out among the three populations. Euclidean

distance based hierarchical cluster analysis was performed from the observed frequency data. Principal component analysis (PCA) score plot and neighbour joining (NJ) tree was constructed using SPSS (ver. 15.0).

Table No. 1: List of forward and reverse primers for the 10 TLRs in human.

Genes	Forward Primers (5'-3')	Reverse Primers (3'-5')	Product size(bp)	GC content (%)
TLR1	TCAACCAGGAATTGGAATAC	AGTTCCAGATTTGCTACAGT	382	40
TLR2	GGATGGTTGTGCTTTTAAGTACTG	AAGATCCCAACTAGACAAAGACTG	2671	41.67
TLR3	ATTGGGTCTGGGAACATTTCTCTTC	GTGAGATTTAAACATTCCTCTTCGC	792	44/40
TLR4	TTCTTCTAACTTCCTCTCCTGTG	TTAGCTGTTCGGCTCTACTATGG	1087	
TLR5	CATTGTATGCACTGTCACTC	CCACCACCATGATGAGAGCA	446	45/55
TLR6	ACAACCCTTTAGGATAGCCACTG	AAACTCACAATAGGATGGCAGG	398	47.83
TLR7	AGTGTCTAAAGAACCTGG	CTTGGCCTTACAGAAATG	545	44
TLR8	CAGAATAGCAGGCGTAACACATCA	AATGTCACAGGTGCATTCAAAGGG	637	45.83
TLR9	TCTAGGGGCTGAATGTGACC	ACAACCCGTCACTGTTGCTT	1106	55
TLR10	GTCGAAGACCCAATATACAG	ATTAAGCAATAGAACCGATG	954	45/35
Growth Hormone (Positive control)	CTTCCCAACCATTCCCTTA	CGGATTTCTGTTGTGTTTC	424	47/42

RESULTS AND DISCUSSION

The observed frequencies of all known TLR genes estimated in Gurkha, Muslim and Rabha populations respectively are represented in **(Table no. 2).** In one of our previous studies, we have screened TLR1-5 genes in the Rabha population [15]. It was observed that TLR4 was found in very high frequency among the Rabhas, while TLR5 was found to be the least frequent among the studied genes [14]. It has been observed that among the 10 TLR genes, TLR4 has the highest frequency among the Gurkhas and the Rabhas while TLR5 was found to be the highest among the Muslims **(Fig. 1).** When compared among the three populations, it was observed that TLR5 has the highest calculated frequency value (0.971) in the Muslim population followed by TLR4 in the Gurkha population (0.968). In contrast, it was interestingly observed that TLR4 has the lowest frequency in the Muslims (0.557) while TLR5 gene was the least frequent among the Rabhas. Apart from TLR2, TLR5 gene also showed low frequency in the Rabha population. Another interesting observation reported from our study was the low frequency of the TLR2 gene in all the studied populations. The gene frequencies of all the 10 TLR genes in the three populations were also calculated and presented in **(Table no. 3)**. Chi-square analyses ($\chi 2$) were performed to compare the differences in carrier frequencies (OF) of TLR genes among the three populations **(Table no. 2).** It was found that no significant differences were observed for 7 out of 10 TLR loci among the Gurkhas and the Muslims which outnumbered the non-significant cases in Gurkhas vs. Rabhas and Rabha vs. Muslims comparisons respectively. No significant differences were found for TLR2, 8 and 9 in any of the comparisons. Mean unbiased genetic diversity of TLR genes in the three populations was calculated to be 0.240 ± 0.038 for Gurkhas, 0.258±0.049 for Muslims and 0.410±0.033 for Rabhas respectively. Hierarchical cluster analysis was also performed followed by the construction of a neighbour joining tree based on the Euclidean distances calculated from the observed frequencies of the TLR genes in the above-mentioned populations, as shown in **(Fig 2a)**. It was quite surprising to see from the tree that the Gurkhas clustered with the Muslims while the Rabha population occupied a different branch of the tree.

Table No. 2: Observed frequencies of the 10 TLR genes in the three populations. χ^2 values were also mentioned where each gene was compared between two populations for any statistical differences.

	Muslims (M)	Gurkha (G)	Rabha (R)	MXG	GXR	RXM
TLR1	0.886	0.928	0.760	0.9314	8.020**	3.687
TLR2	0.571	0.608	0.600	0.229	0.005	0.034
TLR3	0.943	0.856	0.740	4.691*	2.549	13.427***
TLR4	0.557	0.968	0.900	57.426***	2.134	17.499***
TLR5	0.971	0.928	0.460	1.820	45.051***	66.880***
TLR6	0.907	0.896	0.640	0.009	14.349 ***	17.349
TLR7	0.943	0.864	0.680	3.927*	6.745**	20.745***
TLR8	0.829	0.824	0.700	0.0041	2.591	2.986
TLR9	0.793	0.784	0.820	0.0006	0.1057	0.0424
TLR10	0.793	0.832	0.540	0.4299	14.6650***	10.612**

*** P< 0.05, ** P< 0.01, *** P< 0.001**

Table No. 3: Gene frequencies of the 10 TLR genes. The gene frequencies were calculated from the observed frequencies of the 10 TLR genes using the formula 1-√ {1-f}, where f is the observed frequency.

Gene	Muslim	Gurkha	Rabha
TLR1	**0.662**	**0.732**	**0.510**
TLR2	**0.345**	**0.374**	**0.368**
TLR3	**0.761**	**0.621**	**0.490**
TLR4	**0.335**	**0.821**	**0.684**
TLR5	**0.831**	**0.732**	**0.265**
TLR6	**0.695**	**0.678**	**0.400**
TLR7	**0.761**	**0.631**	**0.434**
TLR8	**0.586**	**0.580**	**0.452**
TLR9	**0.545**	**0.535**	**0.576**
TLR10	**0.545**	**0.590**	**0.322**

Table No. 4: Nei's Genetic distances among the three populations using GenAlEx (ver- 6.5) software.

	Gurkha	Muslim	Rabha
Gurkha	**0.000**		
Muslim	**0.023**	**0.000**	
Rabha	**0.056**	**0.089**	**0.000**

From the Principal Component Analyses **(Fig. 2b)** it was evident that the three populations occupied three different quadrants of the score plot whereby the Muslims occupied the upper left quadrant, Gurkhas occupied the upper right quadrant and the Rabhas occupied the lower right quadrant of the plot. Nei's genetic distance was also calculated between the populations **(Table no. 4)** whereby, it was observed that the Gurkha-Muslim genetic distance (0.023) was considerably lesser than the Gurkha-Rabha genetic distance (0.056). The genetic distance was found to be the highest among the Muslims and the Rabhas (0.089).

Figure1. Radar chart constructed using MS EXCEL demonstrating the visual differences in the observed frequency distribution of the 10 TLR genes in the three ethnic populations of North Bengal.

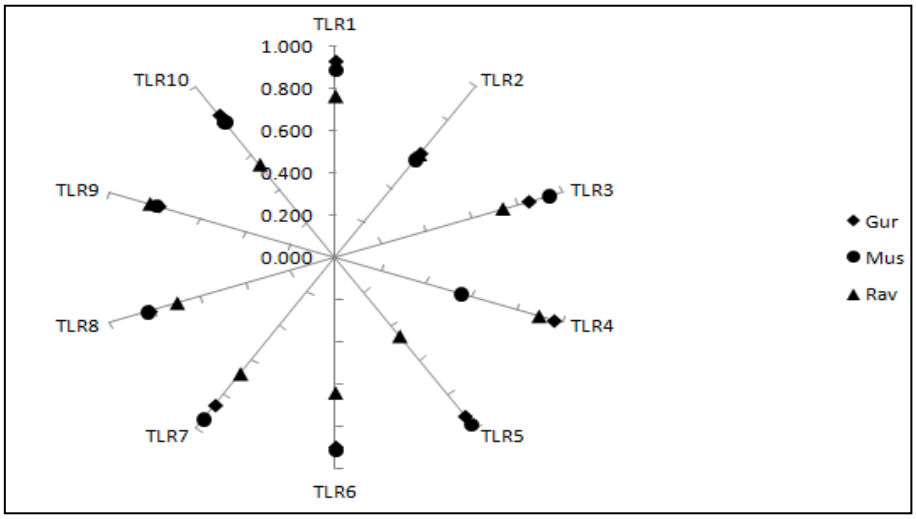
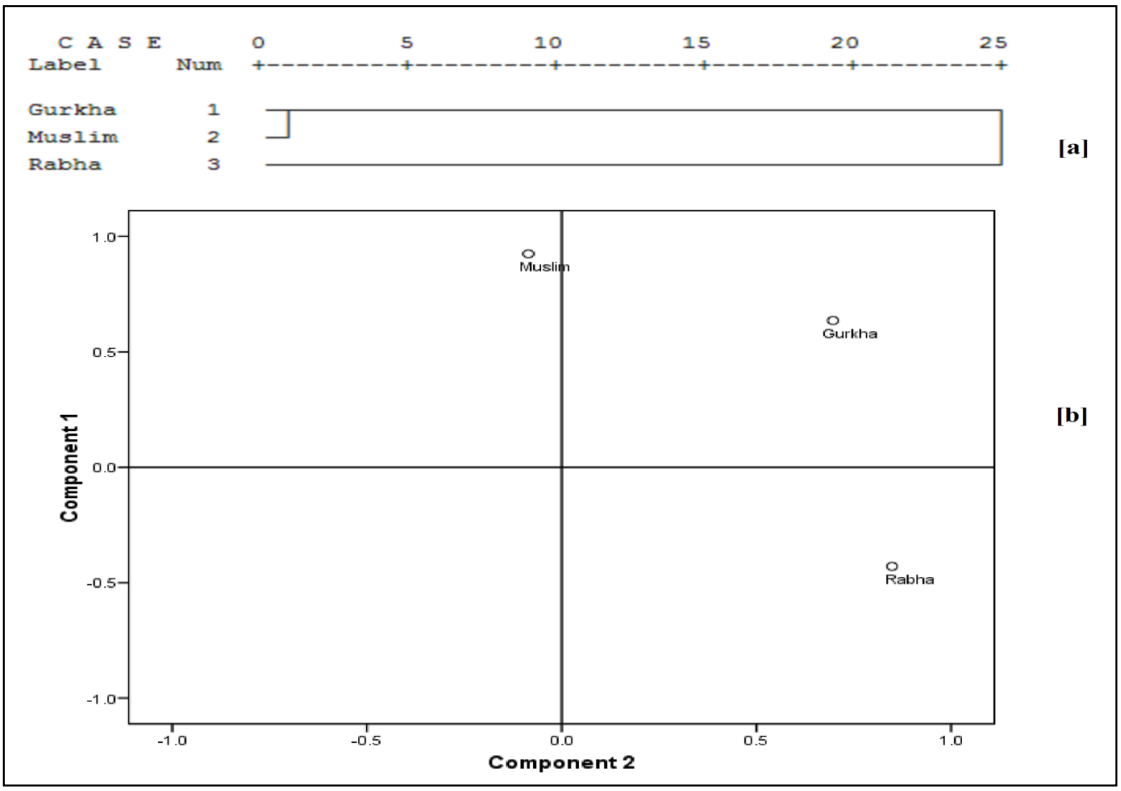


Figure 2. (a) Euclidean distance based Hierarchial cluster analysis was performed using from the observed frequency data by using SPSS (Ver- 15) software. (b) Principle Component Analyses (PCA) based on observed frequencies of the 10 TLR genes in the three ethnic populations of North Bengal constructed with Minitab (ver-6).



The Indian population exhibits enormous diversity in its genetic structure which is not only reflected in its diverse cultural and linguistic backgrounds but also renders difficulty in explaining the overall health and disease conditions in different population subgroups [6]. A combined inter-disciplinary approach or method is much needed to explain and understand the disease-associated genetic variants in the populations and their susceptibility [6]. The TLR profile of a population in alliance with the surrounding environment plays a complex role in disease pathogenesis [5] [16] [17]. In one of our previously published reports, we found that

chronic gastritis and associated stomach problems were very much common in Rabha population. This basic observation led us to speculate that Killer Cell immunoglobulin like receptors (KIR) and Toll like receptors (TLR) may join hand in hand to regulate disease progressions in Rabha population [15]. Based on the frequencies of the TLR4 and TLR5 genes, we assumed that such profile of these genes may indicate an up regulation of IFN-γ production by NK cells which in turn may regulate the prevalence of the Helicobacter pylori negative gastritis [18] [19]. Our assumption was also supported by previously published reports which suggested that the TLR genes play crucial roles in disease pathogenesis [16] [5]. Furthermore, the presence of TLR genes can be documented from the genetic pools of each and every population of the world. Thus, this marker may be considered to clarify the genetic diversity and the relatedness among different populations [11]. However, such preliminary observation requires experimental support. Therefore, in this pioneering study, we have analyzed the frequency of TLR genes in Rabha and two other common population groups of West Bengal, India. Considering the above-mentioned fact, it can be said that this study may help to explore the KIR-TLR connections in disease pathogenesis; because it would always be a judicious decision to explore the distribution pattern of these two gene families in populations before finding their relations and roles in immune responses.

Based on $\chi 2$ analyses it was found that the Gurkha population showed non-significant difference for 7 TLR genes with Muslims in comparison to only 5 TLR genes with the Rabhas, suggesting proximity of the Gurkhas with the Muslim population rather than the Rabhas. A similar observation was also made from the Nei's genetic distance measures where it was seen that the Gurkhas have considerably lesser genetic distance with the Muslim than that of the Rabhas. This observation was also supported by the Neighbour joining dendogram constructed on the basis of Euclidean distances. There is no doubt in the fact that this is a very unlike observation since both Gurkhas and Rabhas are considered to be of East-Asian origin [20] while the Muslims belong to Arabian or Iranian lineages [13]. Thus, the question may be raised regarding the role of TLR in exploring the population phylogenetics and migration pattern. This may be one of the prime reasons why frequency distribution studies based on TLR genes were not conducted earlier in different populations of the world.

Another interesting observation made from our study was that of the TLR4 which has established the predominance over other members of the family, was found in Gurkhas and Rabhas with higher frequencies with no significant differences. In contrary, this gene was found at a very low frequency in the Muslims. On the other hand, TLR5 was present higher frequency in the Muslims and Gurkhas while at a very low frequency in the Rabhas. Such TLR frequency distribution in the three populations suggests the genetic remoteness of the Rabhas with the Muslims while Gurkha occupying an intermediate position. It also suggested the influence of similar environmental exposure on the selection of TLR markers in these populations. This observation was further supported by the Nei's genetic distance measures, Euclidean distance-based NJ tree and PCA score plot. In our study, the three populations have occupied three different quadrants of the score plot and thereby signifying the considerable genetic variability with each other. This was further supported by the mean unbiased diversity measures, whereby it was seen that the Gurkhas showed the lowest value followed by the Muslims and the Rabhas respectively.

According to anthropological evidences, both the Gurkhas and the Rabhas belong to the East Asian origin [10] [15]. However, the TLR profiles in these two populations are quite different. This may have resulted due to the strong influence of the environmental selection on their TLR gene pool. Apart from the selected region of Northern part of West Bengal, the Rabhas are scarcely distributed in the North-Eastern states of India. In contrast, the Gurkhas are distributed over a wide range in the Eastern and North-Eastern part of the country while also being the major population of the neighbouring country of Nepal [20]. Thus, the selection pressure of the surrounding environment was larger on the TLR gene pool of the Gurkha population compared to that of the Rabhas, resulting in the drifting apart of their gene pool from that of the Rabhas.

On the other hand, the Muslim population selected for our study shared the same ancestry with the Bangladeshi Muslims and therefore are distributed over a wide range encompassing the whole of Bengal and Bangladesh. Their TLR gene pool was not only

influenced by inter-regional marriages with Muslims from all over India but also experienced East-Asian influence due to human migration events from neighbouring geographical locations. Furthermore, due to their robust spatial distribution, the selection pressure of the surrounding environment was huge on the TLR gene pool of the Muslim population. Interestingly, it was found that the Muslims population showed genetic proximity to the Gurkhas. This may have occurred due to convergent evolution of TLR genes in these two populations probably due to the selection pressure exerted by the environment which they share [21] [22]. Our speculation has been supported by previously published reports whereby it was suggested that infectious disease like plague may exert influence on the convergent evolution of TLRs in some recent human populations with different genetic ancestry but having exposure to similar environmental condition. Such convergent evolution of TLR was observed among the Romanians and Rroma, which are populations with different origins but sharing the same environment [23].

CONCLUSIONS

In conclusion, it would not be sensible enough to say that TLR gene profile of a population generates sufficient data to establish the genetic connection of the population with other world populations. However, further investigations are required in order to analyze the role of TLR genes in studying population origin and migration events. Furthermore, this study has showed that the TLR gene profile of a population is highly influenced by its ambient environment. Therefore, studies on frequency distribution of TLR genes in different population around the World are very essential since these studies may help us to understand the susceptibility of a disease in a population having a particular genetic makeup and geographical distribution and may also pave the way to further advanced genetic researches for disease eradications.

Acknowledgement

We gratefully acknowledge the support from ICMR for providing the project grants (Project sanction no. 54/2/2012-HUM-BMS). I also express my gratitude to the Blood bank of Siliguri district hospital and pathology section of Malda medical college and hospital that helped us for the collection of the blood samples from that area.

Conflict of interest: The authors declare that they have no conflicts of interest.

REFERENCES

[1] Admetlla AF, Bosch E, Sikora M., Bonet TM, Soriano AR, Muntasell A. Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity Genes. J Immunol. 2008; 181: 1315-1322.

[2] Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. Nat Immunol. 2001; 2: 675–680.

[3] Akira S, Takeda K. Toll like receptor signaling. Nat Immunol. 2004; 4: 499-511.

[4] Eder W, Klimecki W, Yu L, Von Mutius E, Riedler J, Braun-Fahrländer C, Nowak D, Martinez FD, and the ALEX study team. Toll-like receptor 2 as a major gene for asthma in children of European farmers. J Allergy Clin Immunol. 2004; 113:482-48.

[5] Netea MG, Wijmenga C, O'Neill LA. Genetic variation in Toll-like receptors and disease susceptibility. Nat Immunol. 2012; 13: 535-542.

[6] Tamang R, Singh L, Thangaraj K. Complex genetic origin of Indian populations and its implications. J. Biosci. 2012; 37: 911-919.

[7] Bamshad M, Kivisild T, Watkins WS, Dixon ME, Ricker CE, Rao BB, Naidu JM, Prasad BR, Reddy PG, Rasanayagam , Papiha SS. Genetic Evidence on the Origins of Indian Caste Populations. Genome Res.2001; 11: 994-1004.

[8] Balakrishnan V. A preliminary study of genetic diseases among some populations of the Indian subcontinents. Hum. Evol. 1978; **7**: 67-68.

[9] Data highlights. The scheduled tribes Census of India Source: Office of the Registrar General, India. 2001.1-5.

[10] Mitra A. West Bengal: District Handbooks: Jalpaiguri, Govt. of West Bengal. 1953.

[11] Papiha SS. Genetic variation in India. Hum. Biol. 1996; 68: 607–628.

[12] Robb P. A History of India. Palgrave Macmillan Publishers, Hampshire, England.2002.

[13] Eaaswarkhanth M, Dubey B, Meganathan PR, Ravesh Z, Khan FA, Singh L, Thangaraj K. Diverse genetic origin of Indian Muslims: evidence from autosomal STR loci. J. Hum. Genet. 2009; 54: 340–348.

[14] Das A, Guha p, Chaudhuri TK. Environmental selection influences the diversity of TLR genes in ethnic Rajbanshi population of North Bengal region of India. JGEB. 2016; 14: 241-245.

[15] Guha P, Das A, Dutta S, Bhattacharjee S, Chaudhuri TK. Study of genetic diversity of KIR and TLR in the Rabhas, an endogamous primitive tribe of India. Hum Immunol. 2015; 76: 789–794.

[16] Cook DN, Pisetsky DS, Schwartz DA. Toll-like receptors in the pathogenesis of human disease. Nat. Immunol. 2004; 5: 975–979.

[17] Liu F, Lu W, Qian Q, Qi W, Hu J, Feng B. Frequency of TLR 2, 4, and 9 gene polymorphisms in Chinese population and their susceptibility to type 2 diabetes and coronary artery disease. J. Biomed. Biotechnol. 2012; 3:739-745.

[18] Karttunen R, Karttunen T, Ekre HP, MacDonald TT. Interferon gamma and interleukin 4 secreting cells in the gastric antrum in Helicobacter pylori positive and negative gastritis. Gut. 1995; 36**:** 341–345.

[19] Kanevskiy LM, Telford WG, Sapozhnikov AM., Kovalenko EI. Lipopolysaccharide induces IFN-gamma production in human NK cells. Front. Immunol. 2013; 4: 11.

[20] Roy B. Gorkhas and Gorkhaland. Darjeeling, India: Parbati Roy Foundation. 2012.

[21] Grueber CE, Wallis GP, Jamieson IG. Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. Mol. Ecol. 2013; 22: 4470–4482.

[22] Quach H, Wilson D, Laval G, Patin E, Manry J, Guibert J, et al. Different selective pressures shape the evolution of toll-like receptors in human and African great ape populations. Hum. Mol. Genet. 2013; 22: 4829–4840.

[23] Laayouni H, Oosting M, Luisi P, Ioana M, Alonso S, Ponce IR. Convergent evolution in European and Rroma populations reveals pressure exerted by plague on Toll-like receptors. PNAS. 2014; 111: 2668-2673.

***Corresponding Author:**
Tapas Kumar Chaudhuri*
Email: **dr_tkc_nbu@rediffmail.com**