Int J Hum Genet, 2(4): 233-242 (2002)

Genetic Affinity Among Eight Ethnic Populations of West Bengal and Manipur, India: A Study Based on Six Polymorphic Functional Loci (HLADQA1, LDLR, GYPA, HBGG, D7S8 and GC)

P. Chattopadhyay[1], T.S. Vasulu[2] and V.K. Kashyap[1]

1. DNA Typing Unit, Central Forensic Science Laboratory, Kolkata, West Bengal, India
2. AHGU, Indian Statistical Institute, Kolkata, West Bengal, India

KEY WORDS HLADQA1; LDLR; GYPA; HBGG; D7S8; GC; reverse dot blot; genetic variation; Tribes of Manipur, castes of Bengal, and population affinities

ABSTRACT Extent of genetic variation at six DNA loci: HLADQA1, LDLR, GYPA, HBGG, D7S8 & GC, which are widely used in forensic testing, were investigated, using allele specific oligonucleotide (ASO) probes and reverse dot blot methodology, in three castes, four tribal and one religious group from West Bengal and Manipur. All studied loci were found polymorphic in the studied populations. No departure from Hardy-Weinberg expectations among the selected loci was observed. The Power of Discrimination for the six loci in the eight population groups ranged between 0.925 and 0.494. Heterozygosity values for the six loci ranged between 0.430 and 0.841. The average heterozygosity for the eight populations range from 0.466 in Naga to 0.644 in Brahmin. The gene differentiation among the studied loci is high (G_{ST} = 0.046). The dendrograms based on UPGMA method by six PCR based markers show good correspondence with the spatial and ethnic affiliations of the tribal and caste populations. The UPGMA based phylogenetic tree constructed on the basis of the generated data shows very low genetic distance between the Brahmin and Kayastha communities in relation to the Garo. Genetic affinities among the populations of Manipur reveal very close association between the Meitei, Naga, Hmar and Kuki. Our study suggests that the six PCR based loci, used so far mostly for forensic investigations, can be used fruitfully for micro evolutionary studies as well

INTRODUCTION

The development of molecular genetic technology has led to the discovery of large number of polymorphic loci in the human genome. This has renewed the interest in investigating genomic diversity (and affinity) between human populations with much more depth and clarity than that was possible with the erstwhile traditional serological and biochemical genetic markers. In this regard DNA polymorphism of the human genome, especially HLA region, particularly HLA DQA1 (Gyllensten et al. 1988; Saiki et al. 1989), Low Density Lipoprotein Receptor (LDLR) (Yamamoto et al. 1984), Glycophorin A (GYPA) (Siebert et al. 1987), Hemoglobin G Gammaglobin (HBGG) (Slightom et al. 1980), D7S8 (Horn et al. 1990), and Group Specific Component (GC) (Yang et al. 1985) loci in the expressed region and VNTR and STR loci in the noncoding regions reveal large genetic variation with high levels of heterozygosity and mutation rates. The detection of HLA class II polymorphism is valuable in the areas of individual identification, tissue typing for transplantation, and genetic susceptibility to specific autoimmune disease due to their high degree of polymorphic nature. The polymorphism at HLA complex region though high among structural part of human genome, but relatively low in comparison to VNTRs. However these functional loci (HLADQA1 in HLA region and LDLR, GYPA, HBGG, D7S8 and GC at PM loci), which exhibit low levels of mutation rates are useful in investigating the genetic diversity and microevolution among human populations.

These DNA markers are widely used in gene mapping, forensic studies and information on population structure and population genetics of regional and global populations (Clark 1987; Chakraborty 1990; Edward et al. 1992; Bowcock et al. 1994; Jin and Chakroborty 1995; Papiha et al. 1996; Robinson et al. 1996; Deka et al. 1999; Bhasin and Walter 2001). Some of the

Address for Correspondence: Dr. V.K. Kashyap, DNA Typing Unit, Central Forensic Science Laboratory, 30 Gorachand Road, Kolkata 700014, West Bengal, India.
Telephone: (O): 91 33 284 1753,
Fax (O): 91 33 384 9442, 91 33 284 1753.
E-mail: vkk2k@hotmail.com

recent studies on DNA polymorphism have addressed the issue of genomic diversity (or phylogentic relationship) and geographical, ethno-linguistic affiliation and the population structure of Indian populations. Most of the studies are based on highly polymorphic STR loci (Papiha et al. 1996; Deka et al. 1999; Mukherjee et al. 1999, 2000; Dutta and Kashyap 2001a,b; Reddy et al. 2001). However studies based on HLA region among Indian populations are very few (Balakrishnan et al. 1996; Pitchappan 1986, 1988), but studies on DNA polymorphism of HLA and other functional loci are yet to be attempted. In this study we report the genetic diversity at six PCR based expressed loci (especially HLADQA1, LDLR, GYPA, HBGG, D7S8 and GC) of DNA polymorphism among eight population groups from West Bengal and Manipur (Table 1), to (i) evaluate the usefulness of HLA DQA1 and other five functional loci in assessing the genetic affinities and diversity among regional populations and (ii) to test the hypotheses: (a) that tribal populations show genetic similarity with respect to their geographical contiguity irrespective of their origin or language affiliation, and (b) that Manipur Muslims are genetically distinct from other regional populations.

MATERIAL AND METHODS

Sample Information: A total of 624 blood samples from unrelated healthy individuals were collected from three castes, four tribal and one religious group from West Bengal and Manipur states, India. The West Bengal sample consists of 245 healthy individuals from two higher castes Kayastha (N=114), Brahmin (N=51) from southern and from Garo (N=80), a tribal group from northern region. In case of Manipur in Northeast region, 379 blood samples were collected from different parts of Manipur Valley, viz., higher caste group Meitei (N=102), three tribal groups: Naga (N=76), Kuki (N=75), Hmar (N=60) and from

Table 1: Studied populations, sample sizes, locations and the ethnic backgrounds

Population	*No. of samples*	*Sample locations*	*Traditional Occupation; Socio-cultural affiliation*	*Ethnic background*	*Linguistic affiliation*
Kayastha	114	From the region surrounding Kolkata, districts of 24 Pgs(N) and 24 Pgs (S), Burdwan,	Agriculture, Clerical & Accounts jobs, Business. Upper Caste	Caucasoid*	Indo-European
Brahmin	51	From the region surrounding Kolkata, districts of 24 Pgs (N) and 24 Pgs (S), Burdwan, Darjeeling	Priests, Business, Teaching, Astrology Upper Caste	Caucasoid*	Indo-European
Garo	80	Darjeeling and Jalpaiguri, Coochbehar regions of West Bengal	Shifting Cultivators, Labourer. Tribe	Mongoloid*	Tibeto-Burman
Meitei	102	Imphal, Churachandrapur Moirang, Ukhrul	Agriculture, Carpenters, Fisherman. Upper Caste	Mongoloid*	Tibeto-Burman
Naga	76	Eastern districts of Manipur mainly Imphal and Ukhrul	Shifting Cultivators, Labourer Tribe	Mongoloid*	Tibeto-Burman
Kuki	75	Churachandrapur Moirang, Ukhrul	Shifting Cultivators, Weaving, Blacksmith Tribe	Mongoloid*	Tibeto-Burman
Hmar	60	Imphal and Churachandrapur districts of Manipur	Shifting Cultivators (Jhum) Tribe	Mongoloid*	Tibeto-Burman
Muslim of Manipur	66	Imphal	Trade, Agricultural and wage labour. Religious Group	Mixed Populatio of Indo-European and Tibeto-Burman	Indo-European and Tibeto-Burman

* Singh, K.S. (1998)

Manipuri Muslim group (N=66). The above samples were collected from unrelated volunteers in EDTA containing vacutainers and in few cases as a stain on filter paper. The medical history of each donor was also recorded. Blood was aliquoted (700mL) and stored at –20°C prior to DNA extraction. The details of sample size, location, social status and linguistic affiliation of the studied populations are shown in Table 1 (Singh 1998).

DNA Extraction: DNA was extracted from aliquoted blood samples following a phenol-chloroform method (Comey et al. 1994). Briefly, the samples were incubated in 50mM Tris-HCl, 150mM NaCl and 100mM EDTA. Na_2, with the addition of SDS (1.25%) and 0.03mg/mL proteinase K, and precipitated with absolute ethanol after two extractions with phenol: chlorofrom:isoamyl alcohol (25:24:1), respectively. The quantity of DNA in each sample was estimated using the slot-blot procedure described by Waye et al. (1989) and by Hoefer's DyNA Quant 200 Flurometer. Twenty nanogram of DNA was used for PCR amplification.

PCR Amplification and Typing of Amplified Products: The DNA samples were amplified and typed for the HLA DQA1 and PM loci by using the Amplitype PM+DQA1 PCR Amplification and Typing Kit (Perking and Elmer Corporation) according to the manufacturers protocol (Ampli type: User Guide, Cetus Corporation, CA, 1998). The amplification reaction contained 40mL of Amplitype PM+DQA1 PCR Reaction mix, 40mL of Amplitype Primer Set and 20mL of DNA (20ng). Amplification was carried out in a Perkin Elmer GeneAmp® 2400 Thermal Cycler for 32 cycles: 30s at 95°C for denaturation, 30s at 72°C for primer extension. After 32 cycles, the samples were incubated for an additional 10 min at 72°C. The presence and size of Amplitype PM+DQA1 PCR products was verified by electrophoresis through agarose gels (3% NuSeive and 1% SeaKem) in 0.05X TBE buffer (44.5 mM Tris, Boric Acid and 1mM EDTA) containing 0.05mL/mL ethidium bromide. The Gibco BRL 123bp ladder was used as molecular weight marker. The gel was run at 110V for approximately 30 min. The chromosomal locations, number of alleles of these six functional loci are shown in Table 2.

Statistical Analysis: The allele frequency for each loci was calculated by the gene count method. Possible divergence from the Hardy-Weinberg expectations (HWE) was determined on the basis of the likelihood ratio test (Weir 1992) and the exact test (Guo and Thomson 1992). The level of significance of the test, where applicable, was determined by shuffling 2000 times for each test (Chakraborty et al. 1993). The power of discrimination and the probability of exclusion (Garber and Morris 1983) were calculated from the heterozygosity values of different markers observed for different populations and were used for checking the suitability of six genetic markers in human identification testing and in making Human DNA Database. Average heterozygosity, Genetic diversity and genetic distance were estimated by the genetic distance and Phylogenetic Analysis Software (Nei 1973 and Nei et al. 1983) and the dendrograms were drawn by unweighted

Table 2: The chromosomal location of HLADQA1 and PM Loci and other genotype information

	HLADQA1	*LDLR (Low Density Lipoprotein Receptor)*	*GYPA (Glyco-phorin A)*	*HBGG Hemoglobin G Gammaglobin)*	*D7S8*	*GC (Group Specific Component)*
Chromosomal Location	6p21.3	19p 13.1 – 13.3	4q 28 - 31	11p15.5	7q 22 – 31.1	4q 11- 13
PCR Product (bp)	239/242	214	190	172	151	138
Number of Alleles	7*	2	2**	3	2	3

* The HLA DQA1 4.2 and 4.3 alleles are detected but not distinguished from each other by the AmpliType HLADQA1 DNA Probe Strip.

** The GYPA A and A' alleles are detected but not distinguished from each other by the AmpliType PM DNA Probe Strip.

group-method with arithmetic mean (UPGMA) (Sneath and Sokal 1973). The bootstrap method (Felsenstein 1985) was applied to test the stability of the topology of the UPGMA dendrogram. The same number of loci that were used chosen at random with replacement, and a new distance matrix (D_A) was computed. Using this matrix a dendrogram based on UPGMA method was constructed. This process was repeated 1000 times in order to examine the frequency of each cluster in the original topology. The frequencies (percentages) of all the clusters are shown on the nodal points in the dendrogram. High frequency of a cluster reflects its higher stability.

RESULTS

Estimated allele frequencies at 6 functional PCR based loci: HLADQA1, LDLR, GYPA, HBGG, D7S8 and GC, observed in eight populations from West Bengal and Manipur regions are shown in Table 3 and Table 4. In case of HLADQA1 locus Kayastha and Brahmin populations show almost similar frequencies whereas Garo show differences especially with respect to allele 4.1 and 4.2/4.3. Among the Manipur populations, Muslim group shows distinct differences with the other populations at alleles 1.2, 3 and 4.1. Naga and Hmar show the highest frequency in case of allele 4.2/4.3 (0.42 and 0.40), while the lowest is observed in case of Kuki for allele 2 (0.010) and Hmar for allele 4.1

Table 3: HLA DQA1 allele frerquencies in eight populations of West Bengal and Manipur

		HLA DQA1 Allele						
	N	*1.1*	*1.2*	*1.3*	*2*	*3*	*4.1*	*4.2/4.3*
West Bengal								
Kayastha	114	0.228	0.149	0.096	0.285	0.065	0.096	0.078
Brahmin	51	0.186	0.117	0.147	0.294	0.049	0.127	0.078
Garo	80	0.256	0.125	0.081	0.125	0.062	0.031	0.318
Manipur								
Meitei	102	0.245	0.117	0.083	0.181	0.049	0.039	0.284
Naga	76	0.320	0.100	0.040	0.040	0.040	0.040	0.420
Kuki	75	0.333	0.166	0.083	0.041	0.041	0.000	0.333
Hmar	60	0.283	0.117	0.083	0.066	0.033	0.016	0.400
M. Musim	66	0.250	0.055	0.166	0.111	0.027	0.027	0.361

N: Number of Samples, M. Muslim: Manipur Muslim

Table 4: Allele frequency distribution of LDLR, GYPA, HBGG, D7S8 and GC loci in eight populations of West Bengal and Manipur

		PM Allele											
	N	*LDLR*		*GYPA*		*HBGG*			*D7S8*		*GC*		
		A	*B*	*A*	*B*	*A*	*B*	*C*	*A*	*B*	*A*	*B*	*C*
West Bengal													
Kayastha	114	0.495	0.504	0.552	0.447	0.377	0.622	0.0	0.692	0.307	0.321	0.192	0.486
Brahmin	51	0.460	0.539	0.578	0.421	0.441	0.558	0.0	0.725	0.274	0.343	0.196	0.460
Garo	80	0.375	0.625	0.731	0.268	0.362	0.637	0.0	0.775	0.225	0.212	0.387	0.400
Manipur													
Meitei	102	0.426	0.573	0.696	0.303	0.328	0.671	0.0	0.750	0.250	0.289	0.343	0.367
Naga	76	0.300	0.700	0.700	0.300	0.400	0.600	0.0	0.780	0.220	0.280	0.320	0.400
Kuki	75	0.333	0.666	0.708	0.291	0.312	0.687	0.0	0.604	0.395	0.145	0.375	0.479
Hmar	60	0.300	0.700	0.716	0.283	0.433	0.566	0.0	0.716	0.283	0.183	0.333	0.483
M. Musim	66	0.305	0.694	0.555	0.444	0.361	0.638	0.0	0.722	0.277	0.222	0.166	0.611

Table 5: Test of Hardy-Weinberg equilibrium of six genetic loci (HLADQA1 and PM) in populations of West Bengal and Manipur.

	West Bengal				*Manipur*			
Locus	*Kayastha (N=114)*	*Brahmin (N=51)*	*Garo (N=80)*	*Meitei (N=102)*	*Naga (N=76)*	*Kuki (N=75)*	*Hmar (N=60)*	*Muslim (N=66)*
HLADQA1								
Heterozygosity (observed)	0.833	0.843	0.713	0.784	0.560	0.750	0.633	0.778
Power of Discrimination	0.920	0.925	0.920	0.925	0.854	0.887	0.883	0.871
Prob. Exclusion	0.661	0.680	0.447	0.570	0.245	0.509	0.332	0.558
P (HW Exact Test)	0.175	0.274	0.068	0.092	0.074	0.864	0.214	0.228
LR (HW Liklihood Ratio)	0.113	0.380	0.147	0.130	0.127	0.802	0.294	0.146
LDLR								
Heterozygosity (Observed)	0.447	0.490	0.425	0.480	0.360	0.416	0.333	0.500
Power of Discrimination	0.648	0.627	0.623	0.625	0.586	0.601	0.587	0.549
Prob. Exclusion	0.145	0.185	0.179	0.183	0.165	0.172	0.165	0.166
P (HW Exact Test)	0.271	1.000	0.469	0.842	0.615	1.000	0.364	1.000
LR (HW Liklihood Ratio)	0.347	1.000	0.469	1.000	0.615	1.000	0.364	0.636
GYPA								
Hete ozygosity (observed)	0.578	0.529	0.362	0.411	0.360	0.416	0.433	0.666
Pow r of Discrimination	0.572	0.598	0.559	0.581	0.585	0.570	0.558	0.495
Pro Exclusion	0.185	0.183	0.157	0.165	0.165	0.163	0.161	0.185
P (HW Exact Test)	0.095	0.774	0.567	0.824	0.632	1.000	1.000	0.337
LR (HW Liklihood Ratio)	0.095	0.593	0.567	0.824	0.632	1.000	1.000	0.200
HBGG								
Heterozygosity (observed)	0.508	0.490	0.500	0.480	0.480	0.458	0.600	0.500
Power of Discrimination	0.591	0.623	0.587	0.576	0.614	0.573	0.551	0.587
Prob. Exclusion	0.178	0.184	0.176	0.171	0.182	0.167	0.184	0.176
P (HW Exact Test)	0.434	1.000	0.627	0523	1.000	1.000	0.310	1.000
LR (HW Liklihood Ratio)	0.434	1.000	0.492	0.403	1.000	1.000	0.310	1.000
D7S8								
Heterozygosity (observed)	0.508	0.509	0.375	0.441	0.360	0.458	0.433	0.555
Power of Discrimination	0.547	0.519	0.513	0.525	0.508	0.622	0.558	0.494
Prob. Exclusion	0.166	0.158	0.143	0.152	0.142	0.142	0.161	0.159
P (HW Exact Test)	0.045	0.077	0.731	0.118	1.000	1.000	1.000	0.253
LR (HW Liklihood Ratio)	0.045	0.055	0.552	0.083	1.000	1.000	1.000	0.181
Gc								
Heterozyosity (observed)	0.719	0.764	0.612	0.705	0.680	0.708	0.633	0.388
Power of Discrimination	0.757	0.721	0.793	0.793	0.840	0.691	0.778	0.704
Prob. Exclusion	0.458	0.534	0.305	0.436	0.398	0.441	0.332	0.106
P (HW Exact Test)	0.198	0.120	0.348	0.183	0.788	0.398	1.000	0.065
LR (HW Liklihood Ratio)	0.159	0.131	0.390	0.180	0.764	0.471	1.000	0.175

N: Number of Samples, M. Muslim: Manipur Muslim

Table 6: Gene Diversity analysis of six genetic loci (HLADQA1 and PM) among West Bengal and Manipur populations

Locus	*Ht*	*Hs*	*Gst*
HLDQA1	0.8417	0.7849	0.0675
LDLR	0.4636	0.4477	0.0344
GYPA	0.4718	0.4595	0.0260
HBGG	0.5095	0.4845	0.0491
D7S8	0.4309	0.4235	0.0172
GC	0.6468	0.6086	0.0590
Average	0.5607	0.5348	0.0463

(0.016). Allele 4.1 was totally absent in studied individuals of Kuki. In case of five PM loci (Table 4) Kuki population shows the lowest frequency for alleles HBGG (A), D7S8 (A) and GC (A) and Garo and Manipur Muslim populations show the lowest allele frequency for the allele GYPA (B) and GC (B). Naga and Muslim groups from Manipur show the highest frequency of D7S8 (A) and GC (C) alleles respectively. Allele C of HBGG is totally absent in any of the studied

Table 7: Pair-wise genetic distances (D_A) based on six genetic loci (HLADQA1 and PM) among the eight populations from West Bengal and Manipur regions

Population	*Kaystha*	*Brahmin*	*Garo*	*Meitei*	*Naga*	*Kuki*	*Hmar*	*M. Muslim*
Kaystha								
Brahmin	0.0020							
Garo	0.0204	0.0219						
Meitei	0.0133	0.0149	0.0020					
Naga	0.0322	0.0336	0.0053	0.0087				
Kuki	0.0415	0.0473	0.0119	0.0174	0.0121			
Hmar	0.0303	0.0314	0.0040	0.0084	0.0035	0.0068		
M.Muslim	0.0222	0.0223	0.0128	0.0128	0.0132	0.0193	0.0093	

population. All the populations from West Bengal and Manipur show no detectable deviation from the Hardy-Weinberg Equilibrium for the alleles at HLA DQA1 and PM loci as can be observed from the Homozygosity Test, Likelihood Ratio Test and Exact Test (Table 5). Observed heterozygosities among studied populations are given in Table 5. The populations show higher heterozygosity values (63% to 83%) for HLADQA-1 and GC loci than for other loci (33% to 58%).

Locus-wise and average genetic diversity values (Ht, Hs and Gst) or the degree of differentiation among the studied populations are shown in Table 6. The extent of gene differentiation differs for different loci: the highest (0.0675) is observed for HLDQA1 locus and the lowest is observed at D7S8 locus (0.0172) in the studied groups of populations. The average heterozygosity is observed to be 54% and the average Gst value is 0.0463. The overall G*st* value is however higher than that observed for traditional markers in Indian populations. The genetic distance (D_A) based on the studied loci was computed for the eight populations (Table 7). Brahmin and Kayastha, the two caste populations of West Bengal show the least distance values, whereas Kuki (Manipur) and Brahmin (West Bengal) show the highest distance. To assess the affinity between studied populations dendrogram based on UPGMA (unweighted pair group method with arithmetic mean) is drawn (Fig .1). The figure shows two main clusters. Kayastha and Brahmin of West Bengal show a single close cluster and are distinctively different from the Garo and Manipur

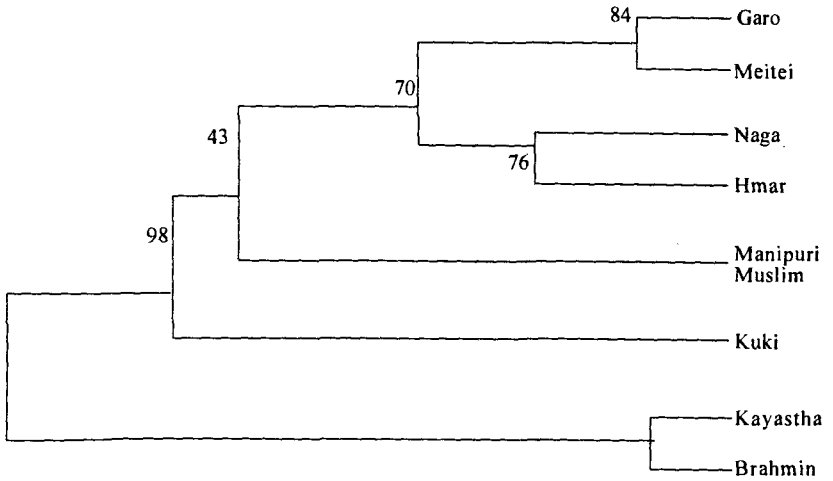

Fig. 1. Genetic affinities between 8 populations from West Bengal and Manipur Regions based on HLADQA1 and PM loci by DA distance and UPGMA Clustering method. The numbers represent the bootstrap values obtained in 1000 replications.

samples. The second cluster consists of two subclusters consisting of the four tribal populations: especially, Naga and Hmar forming a single close subcluster and Garo (West Bengal) and Meitei (Manipur) forming another close subcluster. The clustering of four populations in the second cluster is consistent with the 70% of the bootstraps and the two subclusters are stable at 84% and 76% of the bootstraps respectively. The clustering of Manipur Muslim populations is stable at 98% of the bootstraps. Manipur Muslim groups are distinctly different from other tribal and caste populations. The results of the study are compared with data from USA (white), Pakistan, Arabia, Japan and Korea. The dendrogram shown in Figure 2 shows clustering of Kayastha and Brahmin with US whites, Pakistanis and Arabians, whereas Japanese and Koreans are distinctly different from the Manipur population samples, though both belong to the same racial stock.

DISCUSSION

The genetic profile based on classical genetic markers indicates that caste populations (Brahmin, Kayastha etc.) are separate from northeast tribal population. Further the Mongoloid affiliated populations of West Bengal show genetic affinity with those in Northeast region (Roychoudhury 1992). The results based on studied six genetic loci confirm the above findings thus supporting the results of earlier studies. One of the objectives of the study is to examine the usefulness of HLADQA1 and PM loci in making population database and assessing the genetic affinities among the regional populations. Values of Power of

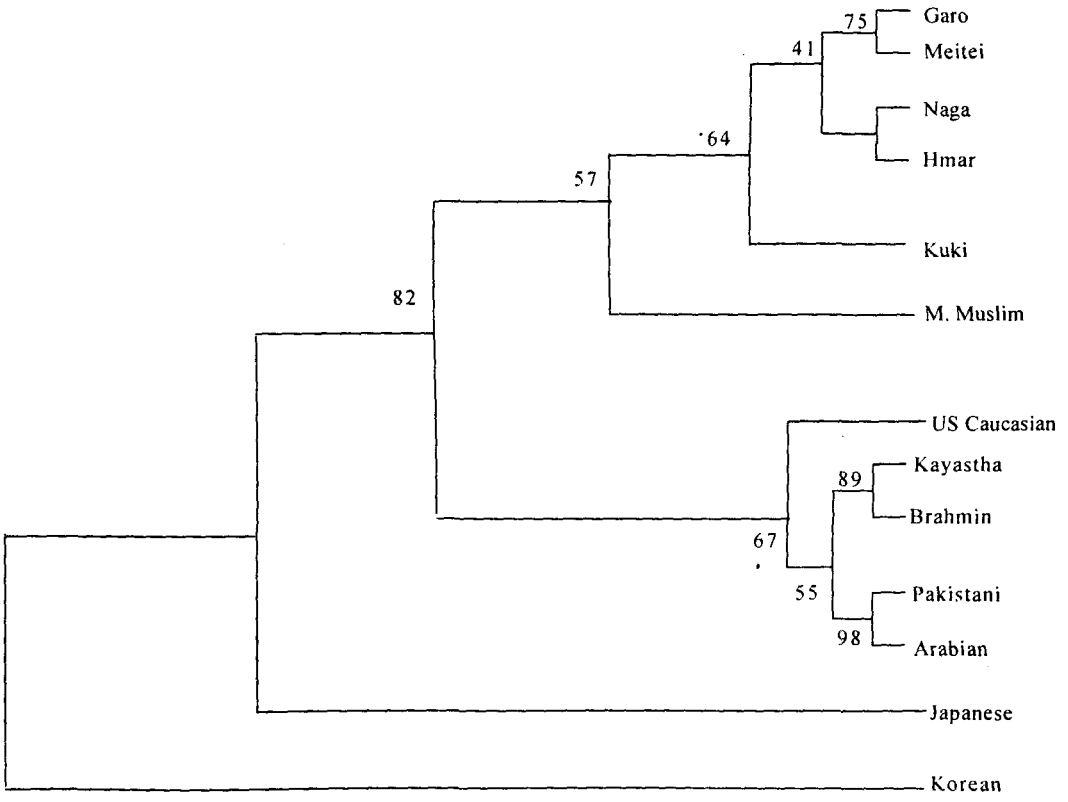

Fig. 2. Comparison of genetic affinities between studied Indian populations with other world populations from published allele frequency data on HLADQA1 and PM loci by DA distance and UPGMA Clustering method

discrimination and Power of Exclusion reveals that these markers are highly suitable for human identification and thus the study generates much needed database essential for identification. Though HLADQA1 and PM System for Human Identification is less informative in comparison to STR marker systems (Chattopadhyay et al. 2001; Dutta 2001a; Dutta 2001b, Dutta et al. 2002), but it is still easier to practice due to the simple detection methodology (reverse dot blot technology) among PCR based systems. In terms of individual allele frequencies HLADQA1 shows distinct differences possibly suggesting the influence of population structure variables.

One of the interesting aspects of the results is about Manipur Muslim population. Unfortunately there were no available records about the history, or origin of the population They are supposed to be a migrant group and maintain their ethnic identity in terms of endogamy and strong cultural values. Their physical appearance of Caucasoid features also suggests that they are morphologically different from the local Mongoloid populations. This possibly accounts for some of the allelic differences observed in them in the study. For example they show the lowest allele frequency (0.055) for allele 1.2 at HLDQA1 locus, highest frequency for allele C of locus GC at PM loci. Thses loci also show some distinct features in Manipur Muslim. And the clustering pattern obtained in the dendrogram also suggest that they are genetically different from other local populations and are not the local converts. Similar results are seen between Meitei and Kuki suggesting least genetic affinity between them. Though Meitei is a caste population, they are originally the Mongoloid origin and in due course of time in the past have adopted Hindu religious tradition. In the dendrogram they cluster with Garo, a Mongoloid population of West Bengal, than with the two regional populations: Naga and Hmar.

The clustering pattern based upon diversity in HLADQA1 and PM loci for the above eight populations is in agreement with the spatial, linguistic and ethnic affiliations. The clustering pattern of the dendrogram depicts that population groups and the tribal group of Manipur formed a major distinct cluster, leaving the other two higher caste groups of West Bengal, Brahmin and Kayastha in another major cluster, and this formation is highly significant (98% of the boot strap). This may be a pointer to the fact that these two major cluster populations were originated from different roots. The high percentages of bootstrap values given at the nodes of each cluster might suggest genetic reality in the observed pattern. These findings also confirm the results obtained from an earlier study based on STR loci for the above eight populations (Chattopadhyay et al. 2001). The UPGMA tree (Fig. 2) shows the genetic relationship of studied Indian populations with other world populations (Budowle et al. 1995; Hayes et al. 1995; Tahir et al. 1997; Tamaki et al. 1991; Woo et al. 1995), which further substantiate the conception of phylogenetic affinities among populations by other DNA markers.

Though there are few reports on HLA region but they are not based on DNA markers, but on biomarkers (for example, Balakrishnan et al. 1996) and using less sensitive and specific lymphocyte toxicity assay. The results based on 6 PCR based loci reflects the influence of population structure variables, especially marriage pattern, migration that need to be investigated for understanding the genetic diversity of the regional populations. In comparison to caste population the tribal population and the Muslim population were relatively isolated since several generations after their migration to northeastern hilly terrain. Thus Manipur Muslims though they live along with Meitei and other tribal populations but show wide genetic diversity. Garo and Kuki possibly had migrated from the same region in Tibet and settled at different places. Though the *Gst* values depend on number and type of loci, the values are similar to other populations studied for Indian and World populations and it is similar to or equivalent to the values found in case of classical markers (Chakraborty et al. 1986, 1987; Singh et al. 1986; Robinson et al. 1996; Das et al. 1996; Papiha 1996). As far as we know, it is the first report on HLDQA1 and LDLR, GYPA, HBGG, D7S8 and GC loci in the populations of West Bengal and Manipur; and also first report on these populations at DNA level.

Polymorphism at HLADQA1 and PM system is low in comparison to set of STR loci, but the present system is highly useful tool in population studies, because, (i) the simple detection

technology (reverse dot blot technology) (Saiki, et al. 1989) and (ii) known highest polymorphism in the structural part of human genome, which reflects true genome diversity (Balakrishnan et al. 1996; Robinson et al. 1996). UPGMA trees constructed on the basis of information from HLADQA1 and PM loci for understanding the clustering pattern, clearly demonstrates information at par with STR markers (Deka et al. 1999; Chattopadhyay et al. 2001; Dutta et al. 2001a; Dutta et al. 2002). Our study clearly demonstrate the diversity at six structural genes, which are in agreement with socio-culture, ethno-history and with previously studied microsatellite data of the studied populations of India.

ACKNOWLEDGEMENTS

Author acknowledges laboratory facilities of C.F.S.L. Kolkata for pursuing experiments. The first author wishes to acknowledge the financial support as SRF from Bureau of Police Research & Development, India.

REFERENCES

Balakrishnan K, Pitchappan RM, Suzuki K, Sankar Kumar U, Santha Kumari R, Tokunaga, K 1996. HLA affinities of Iyers, a Brahmin population of Tamil Nadu, South India. *Hum Biol*, **68**: 523-537.

Bhasin MK, Walter H 2001. *Genetics of Castes and Tribes of India*. Delhi: Kamla-Raj Enterprises.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR and Cavalli-Sforza LL 1994. High resolution of human evolutionary trees with polymorphism evolutionary trees with polymorphic microsatellites. *Nature*, **368**: 455-457.

Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM and Comey CT 1995. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8 and GC (PM Loci) and HLA-DQa using a multiplex amplification and typing procedure. *J Forensic Sci*, **40**: 45-54.

Cetus Corporation 1998. *Amplitype PM+DQA1. User Guide*. Emeryville. C.A.

Chattopadhyay P, Kashyap VK, Vasulu TS 2001. Genomic diversity at three STR tetrameric loci among eight ethnic populations of West Bengal and Manipur, India. *Int J Hum Genet*, **1**: 151-158.

Chakraborty R 1990. Genetic profile of cosmpolitan populations: Effects of hidden subdivision. *Anthropol Anz*, **48**: 313-331.

Chakraborty R, Walter H, Mukherjee BN, Malhotra KC, Sauber P, Banerjee S, and Roy M 1986. Gene differentiation among ten endogamous groups of West Bengal, India. *Am J Phys Anthrop*, **7**: 295-309.

Chakraborty R, Walter H, Sauber P, Mukherjee BN, Malhotra KC, Banerjee S, Roy M 1987. Immunoglobin (Gm and Km) allotypes in nine endogamous Groups of West Bengal, India. *Ann Hum Biol*, **14**: 155-157.

Chakraborty R, Jin L, Zhong Y 1993. On allele frequency compilation from DNA typing data. *Int J Leg Med*, **106**: 103-106.

Clark AG 1987. Neutrality tests of highly polymorphic restriction fragment length polymorphisms. *Am J Hum Genet*, **41**: 948-956.

Comey CT, Kooms BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM, Baechtel FS 1994. DNA extraction strategies for amplified fragment length polymorphisms analysis. *J Forensic Sci*, **39**: 1254-1269.

Das KC, Malhotra KC, Mukherjee BN, Walter H, Majumder PP, Papiha SS 1996. Population structure and genetic differentiation among 16 tribal Populations of Central India. *Hum Biol*, **68**: 679-705.

Deka R, Shriver MD, Yu LM, Heidrecich EM, Jin L, Zhong Y, McGarvey ST, Agarwal SS 1999. Genetic variation at twenty three microsatellite loci in sixteen human populations. *J Genet*, **78**: 99-121.

Dutta R, Reddy BM, Chattopadhyay P, Kashyap VK, Sun G, Deka R 2002. Patterns of genetic diversity at the 9 forensically approved STR loci in the Indian populations. *Hum Biol*, **74**: 33-49.

Dutta R, Kashyap V K 2001a. Genetic variation observed at three tetrameric short tandem repeat loci HumO1, TP0X, and CSF1PO – in five ethnic population groups of northeastern India. *Am J Hum Biol*, **13**: 23-29.

Dutta R, Kashyap K 2001b. Genetic variation at ministaellite loci D1S7, D4S139, D5S110 and D17S79 among three population groups of eastern India. *J Genet*, **80**: 23-30.

Edwards A, Hammond H, Jin L, Caskey CT, Chakraborty R 1992. Genetic variation at five trimeric and tetrameric repeat loci in four human population groups. *Geneomics*, **12**: 241-253.

Felsenstein J 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**: 783-791.

Garber RA, Morris JW 1983. General equations for the average power of exclusion for genetic systems of n codominant alleles in one-parent cases of disputed parentage. In: Walker RH (Ed.): *Inclusion Probabilities in Parentage Testing*. American Association of Blood Banks Arlington VA pp 277-88.

Guo SW, Thompson EA 1992. Performing the exact test of Hardy-Weinberg proportion for mutiple alleles. *Biometrics*, **48**: 361-372.

Gyllensten UB, Erlich HA 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQ alpha locus. *Proc Nat Acad Sci USA*,

85: 7652-7656.

Hayes JM, Budowle B, Freund M 1995. Arab Population Data on the PCR-Based Loci: HLA-DQA1, LDLR, GYPA, D7S8, Gc, and D1S80. *J Forensic Sci*, **40**: 888-892.

Horn GT, Richards B, Merrill JJ, Klinger KW 1990. Characterization and rapid diagnostic analysis of DNA polymorphisms closely linked to the cystic fibrosis locus. *Clin Chem*, **36**: 1614-1619.

Jin L, Chakraborty R 1995. Population structure, stepwise mutations, heterozygosity deficiency, and their implications in DNA forensics. *Heredity*, **74**: 274-285.

Mukherjee N, Majumder PP, Roy B, Roy M, Dey B, Chakraborty M, Banerjee S 1999. Variation at 4 short tandem repeat loci in 8 population groups of India. *Hum Biol*, **71**: 439-446.

Mukherjee N, Mitra M, Chakraborty M, Majumder PP 2000. Congruence of genomic and ethnolinguistic affiliations among five tribal populations of Madhya Pradesh (India). *J Genet*, **79**: 41-46.

Nei M, Tajima F, Tateno Y 1983. Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol*, **19**: 153-170.

Nei M 1973. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sc USA*, **70**: 3321-3323.

Papiha SS 1996. Genetic Variation in India. *Hum Biol*, **68**: 607-628.

Papiha SS, Mastana SS, Purandare CA, Jayasekara R, Chakraborty R 1996. Population Genetic study of three VNTR loci (D2S44, D7S22, and D12S11) in five ethnically defined populations of the Indian subcontinent. *Hum Biol*, **68**: 819-835.

Pitchappan RM 1986. HLA and India. In: HLA in Asia-Oceania. M Aizawa (Ed.). Sapporo, Japan: Hokkaido University Press, pp 535-538.

Pitchappan RM 1988. Founder effect explains the distribution of the HLA-A1-B17 but not the absence of the A1-B 8 haplotype in India. *J Genet*, **67**: 101-111.

Reddy BM, Dutta R, Langstieh BT, Kashyap VK 2001. Diversity at three tetrameric STR loci in a substructured Golla caste population of southern Andhra Pradesh, in comparison to other Indian populations. *Int J Hum Genet*, **1**: 101-108.

Robinson S, Gutowski S, Van Oorschot RAH, Fripp Y, Mitchell J 1996. Genetic Diversity among selected ethnic subpopulations of Australia: Evidence from three highly polymorphic DNA loci. *Hum Biol*, **68**: 489-508.

Roychoudhury AK 1992. Genetic relationship of the populations in eastern India. *Ann Hum Biol*, **1**: 489-501.

Saiki RK, Walsh S, Levenson CH, Erlich HA 1989. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc Natl Acad Sci USA*, **86**: 6230-6234.

Siebert PD, Fukuda M 1987. Molecular cloning of human glycophorin B cDNA: nucletide sequence and genomic relationship to glycophorin. *Proc Natl Acad Sci USA*, **84**: 6735-6739.

Singh KS, Mukherjee BM, Walter H, Lindenberg P, GilbertK, Dannewitz A Malhotra K C, Banerjee S, Roy M, Dey B 1986. Genetic markers among Meiteis and Brahmins of Manipur India. *Hum Hered*, **36**:177-187.

Singh KS, 1998. In: *People of India*, National Series Volume: India's Communities (IV–VI) Anthropological Survey of India Oxford University Press.

Slightom JL, Blechl AE, and Smithies O 1980. Human fetal ^{G}g – and ^{A}g -globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell*, **21**: 627-638.

Sneath PHA, Sokal RR 1973. *Numerical Taxonomy*. San Francisco: Freeman.

Tahir M A, Al Khyat A Q, Shamali FA, Budowle B, Novick GE 1997 Distribution of HLA-DQA1 alleles in Arab and Pakistani individuals from Dubai, United Arab Emirates. *Forensic Sc Int*, **85**: 219-223.

Tamaki K, Yamamoto T, Uchihi R, Katsumata Y, Kondo K, Mizuno S, Kimura A, Sasazuki T 19[illegible]. Frequency of HLA-DQA1 alleles in the Japanese population. *Hum Hered*, **41**: 209-214.

Waye JS, Presley L, Budowle B, Shulter GG, Fourney RM 1989. A simple method of quantifying human genomic DNA in forensic specimen extracts. *Biotechniques*, **7**: 852-855.

Weir BS 1992. Independence of VNTR alleles defined by fixed bins. *Genetics*, **130**: 873-887.

Woo K M, Budowle B 1995. Korean Population Data on the PCR-Based Loci LDLR, GYPA, HBGG, D7S8, Gc, HLA-DQA1, and D1S80. *J Forensic Sci*, **40**: 645-648.

Yamamoto T, Davis CG, Brown MS, Schneider WJ, Casey ML, Goldstein JL, Russell DW 1984. The human LDL receptor: a Cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell*, **39**: 27-38.

Yang F, Brune JL, Naylor SL, Apples RL, Naberhaus KH 1985. Human group-specific component (Gc) is a member of the albumin family. *Proc Natl Acad Sci USA*, **82**: 7994-7998.