Int J Hum Genet, 2(4): 255-260 (2002)
PRINT: ISSN 0972-3757 ONLINE: 2456-6360
DOI: 10.31901/24566330.2002/02.04.06

Genetic Variation of Serum Proteins Among the Koch Sub-populations of West Bengal, India

M. K. Das[1], K. Das[2] and S. S. Mastana[3]

1. Kothari Medical Centre, Kolkata, India
2. Anthropology and Human Genetics Unit, Indian Statistical Institute, Kolkata, India
3. Human Genetics Lab., Department of Human Sciences, Loughborough University, Loughborough, UK

KEY WORDS Koch; Boro Deshi; Chhoto Deshi; West Bengal; serum proteins.

ABSTRACT The phenotype and allele frequency distribution of group specific component (GC), transferrin (TF), alpha-1-antitrypsin (PI) and Apolipoprotein E (APOE) was determined by isoelectric focusing of plasma samples from three subpopulations (Boro-Deshi-Dinajpur District, Boro-Deshi-Malda District, and ChhotoDeshi-Malda District) of the Koch Scheduled castes of West Bengal. An appreciable level of allele frequency variation was observed in populations, which highlighted social and geographical isolation among these subpopulations and which is also compatible with the ethnohistory of the groups studied. Average heterozygosity for these IEF subtype systems was high (46-61%) and gene diversity among these population groups was of low to moderate range (1.5%). Overall analysis showed that these polymorphisms are useful anthropological markers for micro-evolutionary and genetic structure studies.

INTRODUCTION

Anthropogenetic and population studies from different parts of the world and India show the considerable importance of genetic markers to the analysis of genetic structure and differentiation. Comprehensive studies on conventional and DNA polymorphisms have highlighted the use and applications of allele frequency variation among populations for applications in disease diagnostics, forensics and evolution studies. A large amount of data on ethnic, regional and social distribution of genetic markers has already been reported on Indian populations, but our knowledge about genetic variation of different genetic markers is still limited (Bhasin et al. 1994; Walter et al. 1993a,b,c; Das et al. 1996; Mastana et al. 2000).

Address for Correspondence: Dr. Sarabjit Mastana, Human Genetics Laboratory, Department of Human Sciences, Loughborough University, Loughborough. LE11 3TU. UK
Tel: +44-1509-22304, *Fax:* +44-1509-223941
E-mail: S.S.Mastana@LBORO.AC.UK

In order to enlarge our knowledge of genetic serum protein markers, we report here on the distribution of Group specific component (GC), Alpha-1-antitrypsin (PI), Transferrin (TF) and Apolipoprotein E (APOE) serum proteins among three subpopulations of the Koch scheduled caste group of West Bengal. The allele frequencies of these three samples are also compared to those observed in other populations of the region.

POPULATIONS AND METHODS

The People and Ethnohistory

The Koch ethnic group of West Bengal represents a tribe-caste continuum in the Hindu social system (Saha et al. 1990). They are a Mongoloid group settled in the northern parts of the state and have been described as one of the most ancient peoples of India. There are many subgroups of the main Koch population including Polyia, Deshi and Tiyor. They have probably been admixed with certain Caucasoid castes from time to time. Boro Deshi and Chhoto Deshi are the two subgroups of the Koch, and erstwhile a tribe, now a schedule caste in West Bengal. There is very limited information about the Kochs population of North Bengal prior to the 16th Century. They claim to be autochthonous of the region of the part of North Bengal –erstwhile Gauda, capital of Bengal. Dalton (1872) describes Koch as one of the most ancient peoples in India. He considers them as belonging to Dravidian stock and probably a Branch of the great Bhuiya family. Majumder (1943) describes Bhuiyas belong to Bodo, Koch and Mech tribes. Risley (1891) describes that Koch, Koch-Mandi, Rajbanshi, Palliya and Desi belong to a large Dravidian tribe of North Eastern and Eastern Bengal amongst whom these are grounds for suspecting some admixture of

Mongolian blood.

Boro Deshi and Chhoto Deshi separated from each other long ago and Boro Deshi looks down upon the Chhoto Deshi because of their different lifestyle. Chhoto Deshis are numerically small group. There had been lot of admixture with other local caste population with Chhoto Deshi reported in the local literature. Both of them are traditionally horticultural and agricultural people.

All Deshi people belong to single gotra and therefore marriages are allowed within the gotra. Till recent past they were practising polygeny and both the male and female members have had multiple marriages which have been observed in many pedigrees. Usually they do not marry within relatives but marriages between mother's brother's daughter (MBD) and mother's sister's daughter (MSD) are permissible in the society. Inbreeding level observed to be very low to nil in these populations. They live in contiguous areas of both the districts (Das, MK: Unpublished data).

Boro Deshi are numerically large group and though exact enumeration could not be done because in census some of them declare as Rajbanshi to avail some facilities.

As Chhoto Deshi got isolated long time back from Boro Deshi and it would be interesting to analyse the genetic data to assess whether isolation and fission has created any main differences in the allele frequency spectrum among these populations. We have analysed two samples of Boro Deshi (North Dinajpur and Malda District) and one of Chhoto Deshi for serum protein variation.

Methods: Samples were collected from three non-overlapping generations, nearly of equal sample size, following random sampling method using random numbers. Samples were collected from North Dinajpur and Malda Districts of the West Bengal. Geographical distance between the two districts headquarters is around 75 km. Demographic information reveals no admixture between the two groups during last 5-6 generations. We will refer these groups as Boro Deshi-North Dinajpur (BD-ND), Boro Deshi – Malda District ((BD-MD) and Chhoto Deshi-Malda District (CD-MD) in this paper.

Samples were transported to Human Genetics Laboratory, Department of Human Sciences, Loughborough University, Loughborough, UK, and stored at -20 C until use. The subtyping of GC, TF and PI and pheontyping of APOE was carried out by IEF on polyacrylamide gels essentially following standard techniques described by Constans et al. (1980), Mastana et al. (1996), Mastana and Papiha (1998). For the transferrin system, the plasma samples were treated with ferric ammonium citrate and rivanol before electrophoresis. GC polymorphism was detected using immunofixation of the protein with Anti GC antiserum. APOE polymorphism was detected by western blotting following Eichner et al. (1991)'s method. All typings were read and assessed by two independent scientists.

RESULTS AND DISCUSSION

The phenotype numbers, allele frequencies (with standard errors) and heterozygosities are presented in table 1. All subpopulations and pooled sample were found to be in Hardy-Weinberg equilibrium. As sample sizes are small, some caution is required in interpretations. Comprehensive comparisons with other populations were not possible due to paucity of data analysed for the genetic systems studied in this investigation.

Group specific Component (GC)

GC system is one of the most extensively studied serum systems in Indian populations (Mastana et al. 1996; Walter et al. 1993b). In the present populations, GC*1S was the most common allele and GC 1S-1S the most common phenotype in all populations. GC*1S allele frequency was highest in BD-MD (0.556) and lowest in BD-ND (0.474). CD-MD showed highest frequency of GC*1F (0.380). Although the allele frequencies showed some heterogeneity, the differences were statistically non-significant (all $p > 0.05$). No rare variants were observed in any of the populations. In comparison to other populations from West Bengal, pooled GC*1S allele frequency (0.500) is higher than observed in some other (0.355-378) Koch populations of the region (Saha et al. 1990; Mastana et al.1996). Overall pattern fits well with the observed GC*1F longitudinal cline from West to East (Mastana et al. 1996).

Alpha 1-Antitrypsin (PI)

PI*M1 is the most common allele but it shows

Table 1: Phenotype and allele frequencies of 4 serum proteins in Koch subpopulations

System and Phenotype/Allele Frequency	*Boro Deshi-North Dinajpur*		*Boro Deshi-Malda district*		*Chhota-Deshi Malda District*		*Pooled sample*	
Phenotype	*Obs. No.*	*Exp. No.*	*Obs.No.*	*Exp. No.*	*Obs. No.*	*Exp. No.*	*Obs. No.*	*Exp. No.*
Group Specfic Component								
ISIS	5	8.53	12	13.89	8	10.05	25	32.253
ISIF	13	11.84	11	10.56	18	16.36	42	39.50
IFIF	5	4.11	1	2.01	7	6.66	13	12.09
2IS	13	7.11	15	11.67	9	6.54	27	25.00
2IF	2	4.93	6	4.43	3	5.33	11	15.31
2--2	0	1.48	0	2.45	1	1.07	1	4.84
Total Number		38	45		46		129	
Chi-Square		0.08	0.45		0.77		0.04	
Allele Frequencies								
*IS	0.474	*0.057*	0.0556	*0.052*	0.467	*0.052*	0.500	*0.031*
*IF	0.329	*0.054*	0.211	*0.043*	0.380	*0.051*	0.306	*0.029*
*2	0.197	*0.046*	0.233	*0.045*	0.152	*0.037*	0.194	*0.025*
Heterozygosity	0.63		0.59		0.61		0.62	
Alpha 1-Antitrypsin								
MIM1	12	13.44	17	16.90	30	40.26	59	68.07
MIM2	17	14.67	16	16.25	21	23.14	54	55.05
MIM3	3	1.83	0	0.65	5	4.63	8	6.70
MIS	0	0.61	2	1.30	1	0.93	3	2.98
M2M2	3	4.00	4	3.91	2	3.32	9	11.13
M2M3	0	1.00	1	0.31	0	1.33	1	2.71
M2S	1	0.33	0	0.63	0	0.27	1	1.20
M3M3	0	0.06	0	0.01	0	0.13	0	0.16
M3S	0	0.04	0	0.03	0	0.05	0	0.15
S-S	0	0.01	0	0.03	0	0.01	0	0.03
Total Number	36		40		47		123	
Chi-Sq	0.87		1		0.81		1	
Allele frequencies								
*M1	0.611	*0.057*	0.650	*0.053*	0.737	*0.041*	0.678	*0.028*
*M2	0.333	*0.056*	0.313	*0.052*	0.212	*0.038*	0.274	*0.027*
*M3	0.042	*0.024*	0.013	*0.012*	0.042	*0.019*	0.033	*0.011*
*S	0.014	*0.014*	0.025	*0.017*	0.008	*0.008*	0.015	*0.007*
Heterozygosity	0.51		0.48		0.41		0.35	
Transferrin								
C1C1	17	19.39	15	14.78	26	26.81	58	60.25
C1C2	18	13.40	17	16.81	19	17.37	54	48.33
C1C3	1	0.71	0	0.00	0	0.00	1	0.68
C1D	2	2.12	4	4.64	0	0.00	6	7.49
C2C2	0	2.31	4	4.78	2	2.81	6	9.69
C2C3	0	0.24	0	0.00	0	0.00	0	0.27
C2D	1	0.73	4	2.64	0	0.00	5	3.00
C3D3	0	0.01	0	0.00	0	0.00	0	0.00
C3D	0	0.04	0	0.00	0	0.00	0	0.04
D-D	0	0.06	0	0.36	0	0.00	0	0.23
Total Number	39		44		47	130		
Chi-Sq	0.85		1		1	0.85		
Allel frequencies								
*C1	0.705	*0.052*	0.580	*0.053*	0.755	*0.044*	0.681	*0.029*
*C2	0.244	*0.049*	0.330	*0.050*	0.245	*0.044*	0.273	*0.028*
*C3	0.013	*0.013*	0.000	*0.000*	0.000	*0.000*	0.004	*0.004*
*D	0.038	*0.022*	0.091	*0.031*	0.000	*0.000*	0.042	*0.012*
Heterozygosity	0.44		0.55		0.37		0.46	

System and Phenotype/Allele Frequency	*Boro Deshi-North Dinajpur*		*Boro Deshi-Malda district*		*Chhota-Deshi Malda District*		*Pooled sample*	
Phenotype	*Obs. No.*	*Exp. No.*	*Obs.No.*	*Exp. No.*	*Obs. No.*	*Exp. No.*	*Obs. No.*	*Exp. No.*
Apolipoprotein E								
E2-E2		0.00	0	0.00	1	0.13	1	0.04
E2-E3		0.00	0	0.00	2	3.75	2	3.92
E3-E3	36	36.00	28	28.00	29	28.13	93	92.04
E2-E4		0.00	0	0.00	0	0.00	0	0.00
E3-E4		0.00	0	0.00	0	0.00	0	0.00
E4-E4		0.00	0	0.00	0	0.00	0	0.00
Total Number	36		28		32		96	
Chi-Sq	0.00		0		1		0.85	
Allele frequencies								
APOE*E2	0.000	*0.000*	0.000	*0.000*	0.063	*0.030*	0.021	*0.010*
APOE*E3	1.000	*0.000*	1.000	*0.000*	0.938	*0.030*	0.979	*0.010*
APOE*E4	0.000	*0.000*	0.000	*0.000*	0.000	*0.000*	0.000	*0.000*
Heterozygosity	0.00		0.00		0.12		0.04	

Stardard Errors of allele frequencies are given in italics

wide variation, with lowest value in BD-ND (0.611) and highest in CD-MD (0.737). BD-ND had the highest frequency of PI*M2 (0.333) allele while BD-MD documented higher PI*S allele frequency (0.025). Again statistical differences were non-significant.

The present Koch sample has on average higher *M1 and lower *M2 allele frequencies (0.678 and 0.274) when compared to other populations of the region. . It is interesting to note that the PI*S allele is practically missing in most caste and tribal populations studied from North and North- West India (Mastana and Papiha 1994; Walter et al. 1993a,b). Possible selective advantage of the presence of PI*S alelle warrants further epidemiological / ecogenetic studies.

Transferrin (TF)

In the transferrin system, TF*C1 allele is the most common allele in allele populations and its highest frequency was observed in CD-MD (0.755), while BD-MD possessed highest TF*C2 allele frequency (0.333). Interpopulation differences were not significant though allele frequency variation was considerable. Overall frequency spectrum was similar to other West Bengal populations especially Poliya and Deshi (Saha et al. 1990; Mastana and Papiha 1998). TF*C3 allele, which was once described as Caucasoid marker (Kamboh and Kirk 1983; Walter et al. 1983) was observed at lower frequencies around 1.3% in BD-ND only. Average TF*D allele frequency (0.042) is slightly higher than observed in West Bengal populations (average 0.012) but is comparable to many Mongoloid populations of the region (Mastana and Papiha 1998). Overall observed TF*C2 allele frequencies fit well in the increasing West to East geographic cline in India (Mastana and Papiha 1998).

Apolipoprotein E (APOE)

APOE polymorphism, despite its functional and pathological relevance, has not been studied in many populations of India (Mastana et al. 1998). APOE was monomorphic in this study in two sub-populations (BD-ND and BD-MD) but it showed low level of polymorphism in CD-MD. A caution is required due to small sample size, but further sample analysis will confirm whether this system is polymorphic in these populations or founder effect and genetic drift has played a significant role in its allele frequency variation. No APOE*E4 allele was observed in Chotto Deshi population which may have clinical and evolutionary significance. It may be interesting to assess the relationship of

APOE with reference to disease spectrum in tribal populations to assess the usefulness of APOE as heart disease marker. The lack of suitable disease data precludes us to make any concrete suggestions at this stage.

Genetic Diversity Analysis

While there were no statistically significant differences, as assessed by the chi-square method, between the sub-populations studied, allele frequencies showed interesting and considerable variation. To assess the dynamics of this variation, we computed average heterozygosity and genetic distance between three populations.

The highest heterozygosity was observed for GC (0.62 range 0.59-0.63) and lowest for APOE (0.04, range 0.00-0.012) (Table 1). Average population heterozygosity varies from 0.38 ± 0.10 (CD-MD) to 0.41 ±0.14 (BD-MD), while BD-ND have intermediate value (0.40 ±0.14). These heterozygosity values are higher than blood group and serum protein estimates, but are similar to subtyping estimates from many populations (Walter et al. 1993a; Mastana and Papiha; 1994, Das et al. 1996). Overall CD-MD shows isolation but differences are not statistically significant.

G_{ST} value computed from 4 polymorphisms shows that the differentiation is of low to moderate range (0.015 or 1.5%) with H_T value of 0.398 and H_S at 0.393. These values are typical of the region and perhaps suggest barriers (geographical and social) to mating. In the DA genetic distance analysis, BD-MD and CD-MD had the highest distance (0.0293) The UPGMA dendrogram constructed from these data show distinctive geographical position of these populations (Fig. 1). Similar results were obtained by Banerjee et al. (1992) in which Deshi population was clearly differentiated from Poliya and Tiyor-Koch subgroups.

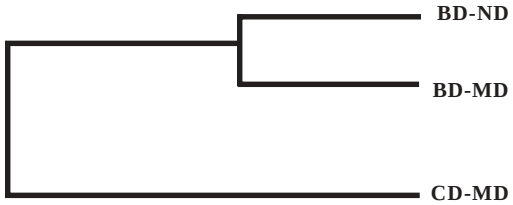

Fig. 1. UPGMA dendrogram of three subpopulations of the Koch Caste.
BD-ND : Boro Deshi-North Dinajpur
BD-MD: Boro Deshi-Malda
CD-MD: Chhoto Deshi -Malda

The results of serum protein systems studied here clearly demonstrate significant heterogeneity amongst the subpopulations of Koch scheduled caste population of West Bengal and highlight that serum protein markers are useful for genetic differentiation and anthropological studies.

REFERENCES

Banerjee S, Das MK, Das K, Dey B, Roy M, Mukherjee BN 1992. Genetic differentiation among the three Hindu low caste groups of Malda district, West Bengal, India. *International Journal of Anthropology,* **7:** 7-15

Bhasin MK, Walter H, Danker-Hopfe H 1994. *People of India: An Investigation of Biological Variability in Ecological, Ethno-Economic and Linguistic Groups.* Delhi: Kamla-Raj Enterprises.

Chatterjee SK 1974. *Kirata-Jana Kriti: The Indo-Mongoloids: Their Contribution to the History and Culture of India.* Calcutta: Asiatic Society.

Constans J, Kuhnl P, Viau M, Spielmann W 1980. A new procedure for the determination of transferrin C Tf C subtypes by isoelectric focusing Existence of two additional alleles, Tf C4 and Tf C5. *Hum Genet,* **55:** 111-114.

Dalton ET 1872. *Descriptive Ethnology of Bengal.* Office of Superintendent of Govt Printers Calcutta Pp 89-92.

Das K, 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.

Eichner JE, Kuller LH, Ferrell RE, Kamboh MI 1991. A simplified method for screening the apolipoprotein E polymorphism. *Hum Hered,* **41**: 61-64.

Grigson WV 1938. *The Maria Gonds of Bastar.* Oxford: Oxford University Press.

Kamboh MI, Kirk RL 1983. Distribution of transferrin Tf subtypes in Asian, Pacific and Australian Aboriginal populations: evidence for the existence of a new type TfC6. *Hum Hered,* **33**: 237-43.

Majumder RC 1943. *History of Bengal,* Vol II p 53.

Mastana SS, Calderon R, Pena J, Reddy PH, Papiha SS 1998. Anthropology of the apoplipoprotein E apo E gene: low frequency of apo E4 allele in Basques and in tribal Baiga populations of India. *Ann Hum Biol,* **25:** 137-43.

Mastana SS, Constans J, Papiha SS 1996. New rare variants and interpretation of the genetic diversity at the group-specific component GC locus in populations of India. *Hum Biol,* **68:** 795-818.

Mastana SS, Papiha SS 1994. Genetic structure and microdifferentiation among four endogamous groups of Maharashtra, western India. *Ann Hum Biol,* **21**: 241-62.

Mastana SS, Papiha SS 1998. Genetic variability of transferrin subtypes in the populations of India. *Hum Biol*, **70**: 729-44.

Mastana SS, Reddy PH, Das MK, Reddy P, Das K 2000. Molecular genetic diversity in 5 populations of Madhya Pradesh, India. *Hum Biol*, **72**: 499-510.

Reddy PH, Mastana SS 1995. Genetic variation of serum proteins GC, TF and PI subtypes in the Baigas of Madhya Pradesh, India. *Anthropol Anz*, **53**: 53-6

Risley H 1891. *Tribes and Castes of Bengal*. Firma Mukhopadhyay, Calcutta

Saha N, Tay JS, Das MK, Das K, Roy M, Dey B, Banerjee S, Mukherjee BN 1990. The distribution of some serum protein and red cell enzyme polymorphisms in the Koch ethnic group of West Bengal, India. *Jpn J Human Genet*, **35:** 253-256.

Walter H, Bhasin MK, Danker-Hopfe H, Khanna A, Hahm U, Witt I 1993a. Genetic studies on eleven populations of Jammu and Kashmir, India. *J Hum Ecol*, **4**: 1-14.

Walter H, Eberhardt D, Tegeler M, Wiechmann I, Danker-Hopfe H, Bhattachachrya SK, Das MK, Mukherjee BN, Saha PN, Das K 1993b. HP, GC and PI polymorphisms in sixteen central Indian tribal populations. *Gene Geogr*, **7:** 193-202.

Walter H, Naidu JM, Danker-Hopfe H, de Beek M, Harms M, Babu BV, Yasmin Devi SS 1993c. Genetic serum protein markers in eight south Indian caste and tribal populations. *Z Morphol Anthropol*, **79**: 355-65.

Walter H, Stach M, Singh IP, Bhasin MK 1983. Transferrin subtypes in four Northwest Indian tribal populations and some remarks on the anthropological value of this new polymorphism. *Am J Phys Anthropol*, **61**: 423-8.