ANNALS OF HUMAN BIOLOGY, 1987, VOL. 14, NO. 2, 155–167

Immunoglobulin (Gm and Km) allotypes in nine endogamous groups of West Bengal, India

R. CHAKRABORTY

Center for Demographic and Population Genetics, University of Texas Health Science Center at Houston

H. WALTER and P. SAUBER

Department of Human Biology/Physical Anthropology, University of Bremen

and B. N. MUKHERJEE, K. C. MALHOTRA, S. BANERJEE and M. ROY

Anthropometry and Human Genetic Research Unit, Indian Statistical Institute, Calcutta

Received 1 May 1986; revised 18 August 1986

Summary. Blood samples from 898 individuals of nine endogamous groups of West Bengal, India were typed for determining the haplotypic structure in the γ-light chain (Gm) and κ-light chain (Km) of immunoglobulin (IgG). The Gm haplotype frequencies detected by Glm (1), Glm (2) and G3m (5) markers suggest that in this eastern state of India there is considerable variation of frequencies of the typical Mongoloid haplotype $Gm^{1,5}$, which shows a high incidence in *Rajbanshi, Rabha, Garo* and *Lodha* groups. On the contrary, this haplotype is probably absent in the high caste groups, *Rarhi Brahmin* and *Vaidya*, and is relatively infrequent in *Jalia Kaibarta*, a scheduled caste of the south-western part of the state. The Km^1 allele is also high in frequency among *Rajbanshi, Rabha, Garo* and *Munda* in comparison with *Rarhi Brahmin* and *Vaidya*, suggesting the former four groups' strong Mongoloid affiliation.

This survey signifies that there is considerable variation in the extent of Mongoloid admixture in Bengali populations. Such admixture is not restricted in specific social class either. It further demonstrates that heterogeneity of the genetic structure of Bengali populations do not correspond to the present social ranking on the basis of caste hierarchy.

1. Introduction

The γ-chain allotypes (Gm) of immunoglobulin G (IgG) and the κ-light chain allotypes (Km, earlier known as Inv) have been proven to be useful markers for detecting genetic variation as well as for finding racial affinity of populations, since these systems exhibit a great extent of genetic polymorphism in terms of heterozygosity and number of haplotypes (Steinberg and Cook 1981). Even though the immunoglobulin heavy-chain allotypes (Gm) are coded by at least three linked genes γ1, γ2 and γ3 (arranged in the order γ3–γ1–γ2, on the q32·3 band of human chromosome 14), and are separated by at least 60 kb distance (Kirsch, Morton, Nakahara and Leder 1982, Migone, de Lange, Piazza and Cavalli-Sfarza 1985), most surveys thus far reported to describe the genetic polymorphism at this gene region involve only Gm (1), Gm (2) and Gm (5) at most.

In India, data on Gm and Km haplotypic distributions are severely limited. Reviews by Steinberg (1974) and Bhasin, Walter, Singh, Hilling and Bhardwaj (1981) and subsequent studies indicate that the Gm and Km surveys in Indian populations are restricted to populations of the Western Himalayan region (state of Himachal Pradesh), Central Himalayas (Kumaon district of Utter Pradesh), South India (Tamil Nadu and Kerala), and some tribal and migrant communities (Parsis and Iranis) of Western India (Chopra, 1970, Steinberg, Undevia and Tepfenhart 1973, Saha, Kirk, Shanbhag, Joshi and Bhatia 1976, Singh, Bhasin, Walter, Hilling, Bhasin and Singh 1982, Bhasin, Singh,

Walter, Sudhakar, Bhardwaj, Dannewitz and Chahal 1985). More recently, Schanfield and Kirk (1981) studied 652 individuals belonging to eight populations from Delhi, Bombay and Madras for G1m, G3m, A2m and Km systems and found evidence of occurrences of IgG haplotypes in Indian high castes of Central Asian origin. The easternmost populations of India so far studied for Gm are the *Oraons* of Bihar (Vos, Kirk and Steinberg 1963), and two communities (*Brahmins* and *Meiteis*) of Manipur (Singh, Mukherjee, Walter, Lindenberg, Gilbert, Dannowitz, Malhotra, Bannerjee, Roy and Dey 1986). Vogel, Krüger, Chakravartti, Flatz and Ritter (1971) typed 961 individuals from the districts of Purulia and Bankura, West Bengal for Km(1) allotypes. However, their data were not subdivided into specific caste and tribe affiliations (only caste Hindu, tribe, and scheduled caste classifications were recorded). Some Indians from Malaya have also been typed for Gm (Vos *et al.* 1963), but they are believed to be immigrants from South India. From these scanty distributions there is a notion that the Indian populations have experienced a significant amount of Mongoloid admixture, as evidenced by the presence of $Gm^{1,5}$ haplotype in considerable frequencies, which may truly represent the haplotype $Gm^{1,3,5,13,14}$ characteristic of Mongoloid populations (Steinberg 1974).

The haplotype $Gm^{1,3}$ is, however, not always present in all Indian populations surveyed so far. Furthermore, if the $Gm^{1,5}$ haplotype in India came through Mongoloid admixture, it is reasonable to expect that its frequency will be considerably higher in populations of eastern states of India which had, in the long history of the country, the highest probability of Mongoloid admixture. With this rationale, 10 endogamous groups of West Bengal had been surveyed through a collaborative research programme of the Indian Statistical Institute, Calcutta and the Bremen University, FR Germany, to test if the various groups of West Bengal suggest the spread of $Gm^{1,5}$ in India through Mongoloid admixture. This paper describes the phenotypic and haplotypic distributions of Gm and Km systems in nine of these groups that are supposed to have different degrees of Mongoloid admixture. The observed distributions are in rough agreement with the known ethnohistory of these groups. However, since the haplotypic determinations are done by fitting Hardy–Weinberg model of phenotype frequencies on population data, like all other Gm surveys, the tentative nature of the above conclusions are discussed in the light of the recent molecular understanding of the IgG gene cluster.

2. Materials and methods

Blood samples were collected in the field and kept refrigerated until they were brought to the Human Genetics Laboratory of the Indian Statistical Institute, Calcutta, where the allotypic determinations were made for heavy-chain markers G1m (1), G1m (2) and G3m (5), and for the light-chain markers Km (1), following the methods described by Grubb (1970). The antisera used were obtained from Dr Molter Ltd (Heidelberg). The Gm haplotype frequencies and their standard errors were evaluated by the maximum-likelihood method, following a modified version of Kurczynski and Steinberg's (1968) program, adapted to include various genetic models of Hardy–Weinberg proportions, varying the nature of haplotypes presumed to exist in these populations. The goodness of fit of the models were tested by a Chi-square test criterion (Rao 1965). The Km gene frequencies and their standard errors were evaluated by maximum-likelihood theory of a recessive two-allele model (Li 1955). Even though we designated the Km alleles as Km^1 and Km^{-1}, the latter probably represents Km^3, since the Km^1 allele may in part consist of the haplotype $Km^{1,2}$, as the phenotype

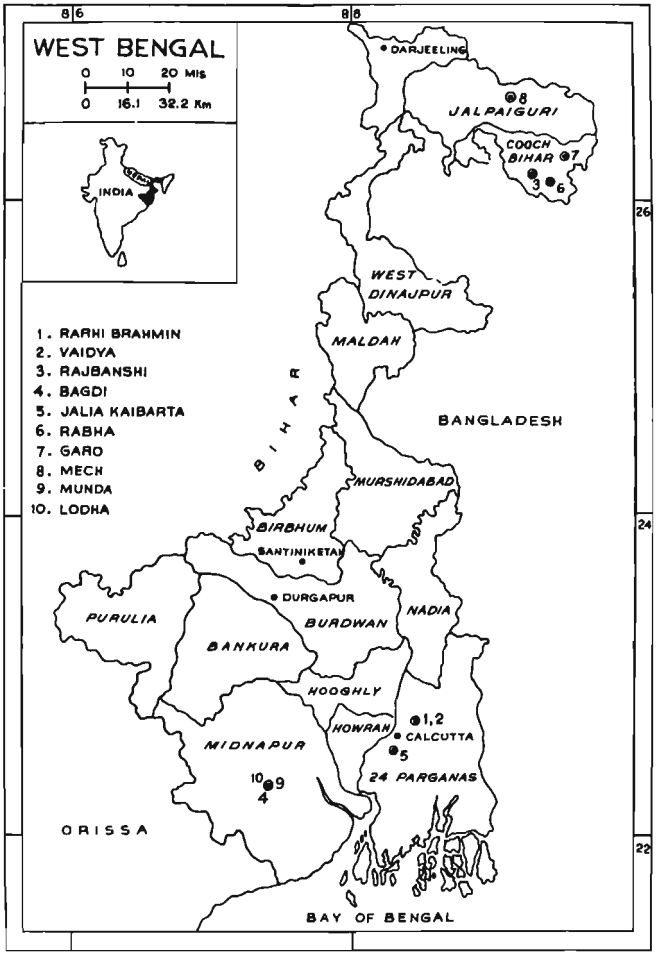

Figure 1. Geographical location of the populations studied in relation to the district boundaries of the State of West Bengal, India.

scoring technique used here involved only antisera Km (1).

The choice of the populations was made on the basis of three criteria: (i) the groups chosen are strictly endogamous, as revealed from their present mating system; (ii) the three major social ranks of the Bengali populations (high caste, scheduled caste and tribe) are all represented by at least one group sampled; (iii) the groups are selected in such a way that geographic regions of West Bengal with low and high chances of Mongoloid admixture are well represented in these samples. Accordingly, four groups (*Rajbanshi, Rabha, Garo* and *Mech*) were chosen from two northern districts of West Bengal where the possibility of Mongoloid admixture was high, and three each from south-west (*Bagdi, Munda* and *Lodha* from Midnapur) and south-east (*Rarhi Brahmin, Vaidya* and *Jalia Kaibarta* from 24-Parganas and Calcutta) zones of the state, where the chance of Mongoloid admixture was supposed to be low. The population names, their distribution in the state, location of the sampling sites, language spoken by them, social rank and sample sizes are given in table 1. Figure 1 presents the geographical location of the sampling sites in relation to the state boundaries. Details of ethnographic accounts of these populations are given in Chakraborty, Walter, Mukherjee, Malhotra, Sauber, Bannerjee and Roy (1986), in connection with other genetic systems surveyed in these populations.

Briefly, the two high-caste groups, *Rarhi Brahmin* and *Vaidya*, may have evolved from immigrants of Caucasoid ancestry from North-west India, although some possibility of mixed blood in them from admixture with lower caste and/or tribes exist (Dutt 1969). Of the three scheduled castes (*Rajbanshi, Bagdi* and *Jalia Kaibarta*), the *Rajbanshis* are generally thought to have strong Mongoloid affiliation (Risley 1908). Little is known about the true ancestry of the *Bagdis* and *Jalia Kaibartas*. Scanty evidence, however, suggests that they may have been transformed into their present social ranking through intermarriages of now extinct tribes with other caste Hindus, including the high-caste ones (Dutt 1969). The three northern tribes (*Rabha, Garo* and *Mech*) are also known to have affiliations with the Assamese tribes of considerable Mongoloid admixture (Dalton 1872, Risley 1908, Das and Raha 1967). The two tribes of Midnapur (*Munda* and *Lodha*) are probable immigrants in Bengal from Bihar and Orissa, respectively. Nevertheless, in these tribes too considerable admixture may have taken place in the past, particularly with low-ranking groups since their arrival in this part of the country.

Even though all of these 10 endogamous groups have been sampled in our survey, because of the limited amount of specimens available the *Mech* group could not be typed for the Gm and Km systems, leaving only 9 populations analysed in the present paper.

3. Results

Gm system

Table 2 presents the observed phenotype frequencies for the Gm system in each of the nine populations surveyed here. These phenotypes can be explained by two possible hypotheses. One that presumes four haplotypes Gm^1, $Gm^{1,2}$, Gm^5 and $Gm^{1,5}$ are present in these populations, which in turn implies that the populations contain in their gene-pool some Mongoloid element (because of the presence of $Gm^{1,5}$ haplotype). A second that assumes the absence of $Gm^{1,5}$, suggesting that Mongoloid admixture did not take place in these populations. Therefore, we fitted both models to each of the populations to evaluate the haplotype frequencies, their standard errors and goodness-

Table 1. Populations of West Bengal sampled, their location, social rank, and sample sizes.

Groups	Population size in district	Location of sampling sites: Village/town	District	Lat. (°N)	Long. (°E)	Social rank	Language	N
Rarhi Brahmin	—	Barasat, Sodepur, Madhyamgram, Baranagore	24 Pgs	22·50	88·25	H. caste	Bengali	100
Vaidya	—	Madhyamgram, Barasat, Baranagore	24 Pgs.	22·40	88·30	H. caste	Bengali	103
Rajbanshi	418893	Putimari, Dinhata, Coochbehar town	Coochbehar	26·18	89·32	S. caste	Bengali	115
Bagdi	119187	Makrampur, Hirapur, Bhara, Middapara, Narayangarh, Chaturi, Bidisa, Pakursani	Midnapur	22·25	87·24	S. caste	Bengali	100
Jalia Kaibarta	5561	Dum Dum, Beliaghata, Bowbazar	Calcutta	22·35	88·21	S. caste	Bengali	101
Rabha	1608	Chhatrampur, Barasalbari, Taliguri, Tufangung town	Coochbehar	26·25	89·50	Tribe	Bengali and Tibeto-Burman	114
Garo	1279	Garopara, Baro Atalbari, Phalkata	Coochbehar	26·22	89·42	Tribe	Bengali and Tibeto-Burman	97
Mech	153	Mahakalguri, Nawabgung	Jalpaiguri	26·30	88·50	Tribe	Bengali and Tibeto-Burman	96
Munda	16960	Belda, Doharpur, Bidisa, Makrampur, Hirapur, Katiaghora	Midnapur	22·50	87·40	Tribe	Bengali and Austro-Asiatic	100
Lodha	11205	Birkanda, Belda, Bidisa, Markunda, Shaldanga, Kuki, Narayangarh	Midnapur	22·25	87·28	Tribe	Bengali	74

Table 2. Gm phenotype frequencies in nine populations of West Bengal, India.

Populations		Gm phenotypes						Test criterion		
		1	1,2	1,2,5	1,5	5	Total	χ^2	d.f.	P
R. Brahmin:	obs	13	11	22	34	18	98			
	exp[1]	12·1	15·6	16·5	32·3	21·5	98	3·88	2	0·144
	exp[2]	10·6	14·8	17·4	36·1	19·1	98	2·96	1	0·085
Vaidya:	obs	11	15	17	33	17	93			
	exp[1]	12·0	16·0	15·9	30·2	18·9	93	0·69	2	0·708
	exp[2]	10·7	15·4	16·6	33·2	17·1	93	0·03	1	0·863
Rajbanshi:	obs	5	5	15	62	7	94			
	exp[1]	15·5	9·5	10·0	37·0	22·0	94	38·98	2	$<10^{-6}$
	exp[2]	4·6	5·8	14·0	62·5	7·1	94	0·23	1	0·632
Bagdi:	obs	12	2	8	69	8	99			
	exp[1]	22·7	5·1	4·7	44·6	21·9	99	31·35	2	$<10^{-6}$
	exp[2]	10·9	3·6	6·3	70·1	8·1	99	1·34	1	0·247
J. Kaibarta:	obs	7	13	21	42	15	98			
	exp[1]	11·2	15·9	17·5	31·4	22·0	98	8·64	2	0·013
	exp[2]	6·8	13·4	20·5	42·2	15·1	98	0·64	1	0·842
Rahba:	obs	0	1	8	86	5	100			
	exp[1]	18·9	4·1	4·7	45·2	27·1	100	78·34	2	$<10^{-8}$
	exp[2]	0·1	0·6	8·6	85·7	5·0	100	0·50	1	0·480
Garo:	obs	0	4	16	69	7	96			
	exp[1]	13·7	8·9	10·6	37·4	25·4	96	59·30	2	$<10^{-6}$
	exp[2]	0·4	2·7	17·9	68·1	6·9	96	1·28	1	0·258
Munda:	obs	2	0	3	79	15	99			
	exp[1]	17·4	1·3	1·7	46·9	31·7	99	46·57	2	$<10^{-6}$
	exp[2]	1·8	0·4	2·6	79·2	15·0	99	0·53	1	0·467
Lodha:	obs	1	0	18	96	6	121			
	exp[1]	19·8	8·0	9·4	51·0	32·8	121	95·34	2	$<10^{-8}$
	exp[2]	0·4	1·7	15·4	97·2	6·1	121	2·96	1	0·085

n.b., exp[1] are the Hardy–Weinberg expectations for the 3-haplotype model, whereas exp[2] are the same for the 4-haplotype model (see text for details).

of-fit statistics. The expected Hardy–Weinberg phenotype frequencies are, therefore, computed under the above two models (exp[1] giving the expected frequencies under the three-haplotype model—Gm^1, $Gm^{1,2}$ and Gm^5, and exp[2] giving the same for the four-haplotype model—$Gm^{1,5}$ in addition to the above).

The estimated haplotype frequencies under both models are presented in table 3 along with their standard errors. As is intuitively clear, the four-haplotype model always fits better, judging from the goodness-of-fit criterion presented in the last three columns of table 2. However, a close look at the parameter estimates presented in table 3 suggests that there are at least two groups (*Rarhi Brahmin* and *Vaidya*) where the $Gm^{1,5}$ haplotype frequency may not be significantly different from zero. In these two populations, the improvement of fit by adding the extra parameter ($Gm^{1,5}$ haplotype frequency) is not statistically significant. In fact, for the *Rarhi Brahmins* the fit for the three-parameter model is better than that for the four-parameter model. The *Jalia Kaibartas* are likely to be another candidate where the four-parameter model is only marginally better (since the residual χ^2 for the three-parameter model is not significant at 1% level of significance). Note that for these three populations, the estimates of the remaining three haplotype (Gm^1, $Gm^{1,2}$ and Gm^5) are not significantly altered even if the presence of $Gm^{1,5}$ is assumed. On the contrary, for the remaining six population groups, the assumption of the absence of $Gm^{1,5}$ haplotype leads to poor fit, judged by the χ^2 statistic (see table 2). Furthermore, in these six remaining groups (*Rajbanshi, Bagdi, Rabha, Garo, Munda* and *Lodha*) the parameter estimates differ significantly in three- and four-parameter models. Thus, we conclude that the evidence of Mongoloid admixture, as detected by genetic variation at the Gm system, is restricted to populations of lower caste and tribes of Bengal. Even if the Mongoloid genes are present in high castes, they may constitute only a small component of their present gene pool.

Table 3. Gm haplotype frequencies and their standard errors (in percent) in nine populations of West Bengal, India.

Populations	Model†	Gm^1	$Gm^{1,2}$	Gm^5	$Gm^{1,5}$
Rarhi Brahmin	3-hap	35·1±3·5	18·0±2·9	46·9±3·6	—
	4-hap	32·9±4·2	18·0±2·9	44·2±4·5	4·9±5·8
Vaidya	3-hap	36·0±3·6	18·9±3·0	45·1±3·6	—
	4-hap	34·0±4·3	19·0±3·0	42·9±4·6	4·1±5·9
Rajbanshi	3-hap	40·6±3·6	11·0±2·3	48·4±3·6	—
	4-hap	22·1±4·5	11·2±2·4	27·4±5·0	39·4±6·7
Bagdi	3-hap	47·9±3·6	5·1±1·6	47·0±3·5	—
	4-hap	33·2±4·5	5·1±1·6	28·6±4·8	33·0±6·6
Jalia Kaibarta	3-hap	33·8±3·5	18·8±2·9	47·4±3·6	—
	4-hap	26·3±4·2	19·1±2·9	39·2±4·6	15·4±6·0
Rabha	3-hap	43·5±3·5	4·5±1·5	52·0±3·5	—
	4-hap	3·8±4·2	4·7±1·5	22·3±4·9	69·2±6·5
Garo	3-hap	37·8±3·5	10·7±2·3	51·5±3·6	—
	4-hap	6·6±4·2	11·4±2·3	26·8±4·9	55·2±6·5
Munda	3-hap	41·9±3·5	1·5±0·9	56·6±3·5	—
	4-hap	13·4±4·8	1·5±0·9	39·0±4·6	46·1±6·7
Lodha	3-hap	40·5±3·2	7·4±1·7	52·1±3·2	—
	4-hap	5·8±3·8	7·4±1·7	22·4±4·4	64·4±5·9

† 3-haplotype model consists of haplotypes Gm^1, $Gm^{1,2}$, and Gm^5; while the 4-haplotype model consists of haplotype $Gm^{1,5}$ in addition to the above haplotypes.

Km system

Table 4 presents the Km phenotypes and allele frequencies, along with standard errors of the estimates. These results are also of significance since Km data are even more sparse for Indian populations. The fact that the *Rajbanshi, Rabha, Garo* and *Munda* show higher incidence of Km^1 allele strengthen the assertion that these populations have significant Mongoloid affiliations. However, it is to some extent surprising that the *Lodhas* have the lowest incidence of Km^1 among the nine populations surveyed here. In fact, even the high-caste groups (*Rarhi Brahmin* and *Vaidya*) show an incidence of this allele that is smaller than other caste Hindus sampled from the Western and Central Himalayas regions (about 10%, see Chopra 1970, Singh *et al.* 1982, Bhasin *et al.* 1981). It is interesting that even at this system, the two scheduled castes, *Bagdi* and *Jalia Kaibarta*, depict allele frequencies closer to the high-caste groups, which is in accordance with their genetic proximity with high-caste Hindus, as revealed from 10 other genetic systems (Chakraborty *et al.* 1986). It is interesting to note that the Km^1 frequencies in the *Rarhi Brahmins* and *Vaidyas* (4·0% and 8·2% respectively) encompass the allele frequency (6·5%) reported by Vogel *et al.* (1971) for the caste Hindus from the districts of Purulia and Bankura, West Bengal, while the three tribes (*Rabha, Garo* and *Munda*) have Km^1 frequencies somewhat larger than that of the tribes of Purulia and Bankura (11·6%; Vogel *et al.* 1971). However, the incidence of Km^1 in the *Rarhi Brahmins* of Bengal seems to be smaller than those in Brahmins of Delhi and Madras, reported by Schanfield and Kirk (1981).

Table 4. Km phenotype and allele frequencies in nine populations of West Bengal, India.

Populations	Km phenotypes			Allele frequencies (%)		
	Km (1) +	Km (1) −	Total	Km^1	Km^{-1}	S.E.
Rarhi Brahmin	5	59	64	4·0	96·0	1·7
Vaidya	11	59	70	8·2	91·8	2·4
Rajbanshi	23	71	94	13·1	86·9	2·6
Bagdi	13	86	99	6·8	93·2	1·8
Jalia Kaibarta	19	82	101	9·9	90·1	2·1
Rabha	31	69	100	16·9	83·1	2·8
Garo	34	63	97	19·4	80·6	3·0
Munda	24	75	99	13·0	87·0	2·5
Lodha	8	113	121	3·4	96·6	1·2

Principal component analysis of Gm and Km haplotype frequencies

The haplotype frequencies at the *Gm* locus and allele frequencies for the *Km* locus are significantly different over these nine populations, as can be seen from a heterogeneity χ^2 (pooled χ^2 over four independent haplotypes is 365·7 with 24 degrees of freedom; $P < 10^{-6}$), or from Wright's $F_{ST} = 0{\cdot}051 \pm 0{\cdot}016$, averaging over all four independent haplotypes (Gm^1, $Gm^{1,2}$, Gm^3 and Km^1). To examine if the clustering of these populations agrees with our genetic distance analysis of inter-group affinity, as detected by 10 other genetic systems (Chakraborty *et al.* 1986), we conducted a principal component analysis of the haplotype frequencies studied in this paper. The eigenvalues and eigenvectors of the pooled within-population covariance matrix of haplotype frequencies was computed, when a single within-population covariance matrix was constructed by averaging the within-population covariance matrix of haplotype frequencies over all nine populations. The first two components of this matrix explain 46·7 and 26·1% of allelic variability, respectively. Based on normalized eigenvectors,

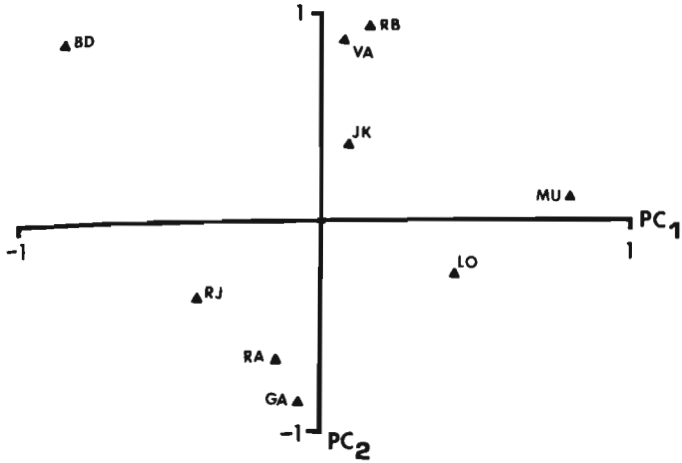

Figure 2. Principal component representation of nine population groups of West Bengal based on Gm haplotype and Km allele frequencies. The population names are abbreviated: RB, *Rarhi Brahmin*; VA, *Vaidya*; RJ, *Rajbanshi*, JK, *Jalia Kaibarta*; BD, *Bagdi*; RA, *Rabha*; GA, *Garo*; MU, *Munda*; LO, *Lodha*.

principal component scores of the haplotype vector for each populations were computed for these two components, and these are represented graphically in figure 2.

It seems clear that there are apparently three clusters in which these nine populations may be grouped. The first cluster comprises two high-caste groups, *Rarhi Brahmin* (RB) and *Vaidya* (VA), which may have some affinity with the scheduled caste group *Jalia Kaibarta* (JK). The second cluster consists of the two tribes, *Rabha* (RA) and *Garo* (GA), and one low-caste group, *Rajbanshi* (RJ). The other two tribes, *Munda* (MU) and *Lodha* (LO), constitute the third cluster. The *Bagdi* (BD) group seems to stand alone on the basis of their Gm and Km haplotype structure. Excepting this last group, the clustering is almost parallel to the clustering that we obtained in the dendrogram analysis from genetic distances computed from the allele frequencies at 10 other genetic systems (Chakraborty *et al.* 1986). With the exception of the *Lodhas* and *Bagdis*, the separation of the remaining seven populations with respect to the second principal component (PC_2) are parallel with the $Gm^{1,5}$ haplotype frequencies; a higher $Gm^{1,5}$ yielding a smaller normalized score for PC_2. From these features, we may conclude that the separation of populations on the basis of the second principal component may indicate the extent of Mongoloid admixture in these nine groups; suggesting that Mongoloid elements may be most predominant in *Rabha, Garo* and *Rajbanshi*, while they are least observed among *Rarhi Brahmins, Vaidyas* and possibly *Jalia Kaibartas*. The Mongoloid admixtures in the remaining three populations are somewhat intermediate between these two extremes.

4. Discussion and conclusion

Even though this survey was conducted with a relatively small number of Gm antibodies, G1m (1), G1m (2) and G3m (5), results of the present analysis are significant on several counts.

First, this constitutes Gm haplotype data on the farthest eastern population groups of India so far examined with the exception of the two populations from Manipur (Singh *et al.* 1986). On the basis of known history of incorporation of Mongoloid genes in India, we expected that some of these groups are likely to provide the highest frequencies of the typical Mongoloid Gm and Km haplotypes found so far in India. To a certain extent this expectation is realized in the present survey. For example, only among the *Tharus* of the Kumaon region of the Central Himalayas, $Gm^{1,5}$ has a higher incidence (67·9%) (Chopra 1970) than the ones found here. Km^1 allele frequencies in these populations are also considerably smaller than those found among the *Rajputs* and *Tharus* of Kumaon region of the Central Himalays (Chopra 1970). Interestingly, however, the *Garo, Rabha, Rajbanshi* and *Munda* show elevated incidences of Km^1 in comparison with the *Rarhi Brahmin, Vaidya, Jalia Kaibarta, Bagdi* and *Lodha* communities, suggesting again that the extent of Mongoloid admixture in these nine populations may vary considerably. In short, therefore, we might infer that while Risley (1908) classified the high-caste Hindus in Bengal as having significant Mongoloid affiliation, this assertion is not borne out with Gm and Km haplotype data.

Secondly, within eastern India the Gm haplotype data exhibit considerable heterogeneity, and the same is true about the Km^1 allele frequencies. It may be noted that Schanfield and Kirk (1981) did not find any such heterogeneity in the Km^1 frequencies in eight populations chosen from Delhi, Bombay and Madras. Not all endogamous groups of this region carry the characteristic Mongoloid haplotype ($Gm^{1,5}$, which is probably the reduced form of $Gm^{1,5,13,14}$ or $Gm^{1,3,5,13,14}$). One caution, however, has to be exercised in this interpretation: we reached this conclusion on the basis of the traditional analysis of Gm phenotype data (e.g. see, Steinberg and Cook 1981), where goodness-of-fit criterion of Hardy-Weinberg proportions is exploited to infer the existence of certain haplotypes from data on unrelated individuals. The test criterion used is well known for its insensitiveness (low statistical power for detecting certain deviations from the null hypothesis; e.g. see, Li and Horvitz 1953, Smith 1970, Ward and Sing 1970, Chakraborty 1975, Emigh 1980, Haber 1981, Robertson and Hill 1984). Additional family data, including kindreds of individuals with phenotypes G1m (1) and G3m (5) positives, might provide a clearer picture as to the existence of the $Gm^{1,5}$ haplotype in these populations.

Thirdly, the results of this survey also indicate that the clustering of the 9 Bengali endogamous groups on the basis of Gm and Km data is in accordance with the findings of 10 other biochemical markers (Chakraborty *et al.* 1986). The clustering does not always put populations of the same social ranks together, since some of the scheduled castes (e.g., *Jalia Kaibartas*) are in fact closer to some high-caste groups, instead of being closer to their compatriots of the same social rank. Subdivision of populations based on social hierarchy may not uncover the genetic structure of populations, as seen in this analysis.

It is true that ideally one should study the distribution of Gm haplotypes with a larger number of antigen specificities. There is a common notion that at least five Gm antisera are required to derive the maximum information regarding the racial affinity of populations. However, unless these antisera are raised from the same source against various Gm specificities, there could be a number of problems in Gm typing. In

addition, since some of the populations examined here are known to reflect a high environmental antigenetic load due to their poor socio-economic conditions, use of such extensive battery of Gm specificities may show environmental differences, rather than their anthropological affinity.

Futhermore, unless the population surveys are supplemented by family studies, and more ideally combined with molecular examination of DNA cloned from specific individuals, the assertion of certain haplotypes cannot be made definitely. For example, G1m (1) detects specific amino acid residues Asp-Glu-Leu-Thr-Lys in position 356–360 of the coded γ1 gene of the IgG gene cluster, while G3m (5) detects the amino acid residues Phe–Phe in positions 296–436 of the coded γ3 gene of the IgG gene cluster (Zaleski *et al.* 1983). These two genes are known to be at least 28 kb apart from one another (Migone *et al.* 1985). Even though no exact estimate of recombination rates at these gene regions are known, if we assume that this region is under uniform recombination rate of 0·001% per kb (estimate obtained by Kurnit and Hoehn 1979), it can be estimated that nearly 0·028% recombination or one recombinant per 1786 meiosis events are expected over the γ1–γ3 gene region of the IgG gene cluster. If we assume that the history of Mongoloid admixture in this region of India goes back 1000 years (approximately 50 generations), and the average effective size of these endogamous groups is 1000 over this period, we have an estimate of roughly 28 possible recombinants that could occur during this period since first Mongoloid admixture had taken place in these populations. Surely, not all of these recombinants will be detectable. Nevertheless, this calculation suggests that it is quite likely that some of these populations may now possess Gm haplotypes that are not generally found in the racial stocks from which they evolved. This argument becomes even stronger when we include haplotypes detected by an extensive battery of Gm markers which cover all three genes, γ1, γ2 and γ3, covering DNA of at least 60 kb long.

Therefore, the analysis of Gm allotype data from unrelated individuals belonging to populations of mixed origin, with the assumption of the existence of a given array of haplotypes (on the basis of their ethnic ancestry), may be regarded as tenuous. This is particularly so when the history of admixture is long, and the markers cover a DNA region as broad as the IgG gene cluster. However, molecular variation of the human IgG gene cluster is still not sufficiently understood to allow examination of the real mechanism of haplotype diversity at population level (see Migone *et al.* 1985, Chaabani *et al.* 1985 for some recent discussions), and hence there does not seem to be any better alternative at present.

In summary, we conclude that the present survey suggests that the Mongoloid admixture that had taken place in eastern India does not seem to be uniform over endogamous groups of all social ranks; neither is it clustered in populations of the same social rank. The exact extent of admixture is somewhat uncertain, since there are methodological difficulties in estimating the extent of admixture because of imprecise knowledge of ancestral stocks through which such admixture had occurred, and the inability of asserting the correct haplotypic status of the populations on the basis of the limited nature of data examined here. As the molecular structure of the IgG gene cluster is understood better, the scope of answering such questions will widen through a combination of molecular and population genetic principles.

Acknowledgements

The data collected during this project was funded by collaborative research funds from Deutsche Forschungsgemeinschaft (Wa 122–29/1) and the Indian Statistical

Institute, Calcutta. Data analysis was supported by US Public Health Service Research Grants from the National Institutes of Health. We are grateful to Professor B. P. Adhikari for his encouragement and guidance during the conduct of the survey. Discussions with Drs D. Hewett-Emmett and T. C. Douglas regarding the molecular basis of the IgG gene cluster are greatly appreciated.

References

BHASIN, M. K., SINGH, I. P., WALTER, H., SUDHAKAR, K., BHARDWAJ, V., DANNEWITZ, A., and CHAHAL, S. M. S., 1985, Genetic studies in four tribal populations of the Surat district, Gujarat (India). *Annals of Human Biology*, **12**, 27–39.

BHASIN, M. K., WALTER, H., SINGH, I. P., HILLING, M., and BHARDWAJ, V., 1981, Geographic and ethnic distribution of genetic markers in India. 2. Inv, Gm, Gc, ADA, AK, AP, PGM_1, 6-pgd and EsD polymorphisms. *Anthropologischer Anzeiger (Stuttgart)*, **39**, 200–243.

CHAABANI, H., BECH-HANSEN, N. T., and COX, D. W., 1985, A multigene deletion within the immunoglobulin heavy-chain region. *American Journal of Human Genetics*, **37**, 1164–1171.

CHAKRABORTY, R., 1975, Power of assaying inbreeding through sampling of phenotypes and mating types. *Human Genetics*, **29**, 217–222.

CHAKRABORTY, R., WALTER, H., MUKHERJEE, B. N., MALHOTRA, K. C., SAUBER, P., BANERJEE, S., and ROY, M., 1986, Gene differentiation among ten endogamous groups of West Bengal, India. *American Journal of Physical Anthropology*, **71**, 295–309.

CHOPRA, V. P., 1970, Studies on serum groups in the Kumaon region, India. *Humangenetik*, **10**, 35–43.

DALTON, E. T., 1872, *Descriptive Ethnology of Bengal* (Calcutta: Government of Bengal Printings).

DAS, A. K., and RAHA, M. K., 1963, *The Rabhas of Bengal* (Calcutta: Director of Cultural Research Institute, Government of West Bengal).

DUTT, N. K., 1969, *Origin and Growth of Caste in India*, Vol. II (Calcutta Firma K. L. Mukhopadhyay).

EMIGH, T. H., 1980, A comparison of tests for Hardy–Weinberg equilibrium. *Biometrics*, **36**, 627–642.

GRUBB, R., 1970, *The Genetic Markers of Human Immunoglobulins* (Berlin: Springer-Verlag).

HABER, M., 1981, Exact significance levels of goodness-of-fit tests for the Hardy–Weinberg equilibrium. *Human Heredity*, **31**, 161–166.

KIRSCH, I. R., MORTON, C. C., NAKAHARA, K., and LEDER, P., 1982, Human immunoglobulin heavy chain gene map to a region of translocation in malignant B lymphocytes. *Science*, **216**, 301–303.

KURCZYNSKI, T. W., and STEINBERG, A. G., 1968, A general program for maximum likelihood estimation of gene frequencies. *American Journal of Human Genetics*, **19**, 178–179.

KURNIT, D. M., and HOEHN, H. H., 1979, Prenatal diagnosis of human genome variation. *Annual Review of Genetics*, **13**, 235–258.

LI, C. C., 1955, *Population Genetics* (Chicago: University of Chicago Press).

LI, C. C., and HORVITZ, D. G., 1953, Some methods of estimating the inbreeding coefficient. *American Journal of Human Genetics*, **5**, 107–117.

MIGONE, N., DE LANGE, G., PIAZZA, A., and CAVALLI-SFORZA, L. L., 1985, Genetic analysis of eight linked polymorphisms within the human immunoglobulin heavy-chain region. *American Journal of Human Genetics*, **37**, 1146–1163.

RAO, C. R., 1965, *Linear Statistical Inference and Its Applications* (New York: John Wiley).

RISLEY, H. H., 1908, *The People of India* (London: Thacker and Spink).

ROBERTSON, A., and HILL, W. G., 1984, Deviations from Hardy–Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. *Genetics*, **107**, 703–718.

SAHA, N., KIRK, R. L., SHANBHAG, S., JOSHI, S. H., and BHATIA, H. M., 1976, Population genetic studies in Kerala and the Nilgiris (South West India). *Human Heredity*, **26**, 175–196.

SCHANFIELD, M. S., and KIRK, R. L., 1981, Further studies on the immunoglobulin allotypes (Gm, Am and Km) in India. *Acta Anthropogenetica*, **5**, 1–21.

SINGH, I. P., BHASIN, M. K., WALTER, H., HILLING, M., BHASIN, V., and SINGH, R., 1982, Genetic studies of Pangwalas, Transhumant and settled Gaddis. 2. Serum protein and red cell enzyme polymorphisms. *Zeitschrift fur Morphologie und Anthropologie*, **73**, 175–195.

SINGH, K. S., MUKHERJEE, B. N., WALTER, H., LINDENBERG, P., GILBERT, K., DANNOWITZ, A., MALHOTRA, K. C., BANERJEE, S., ROY, M., and DEY, B., 1986, Genetic markers among Meitels and Brahmins of Manipur, India. *Human Heredity*, **36**, 177–187.

SMITH, C. A. B., 1970, A note on testing the Hardy–Weinberg law. *Annals of Human Genetics*, **33**, 377–383.

STEINBERG, A. G., 1974, The Gm and Inv immunoglobulin allotypes in Indian populations: a review. In *Human Population Genetics in India*, edited by L. D. Sanghvi, V. Balakrishnan, H. M. Bhatia, P. K. Sukumaran and J. V. Undevia (New Delhi: Orient Longman Ltd.), pp. 112–126.

STEINBERG, A. G., and COOK, C. E., 1981, *The Distribution of the Human Immunoglobulin Allotypes* (Oxford: Oxford University Press).

STEINBERG, A. G., UNDEVIA, J. V., and TEPFENHART, M. A., 1973, Gm and Inv studies of Parsi and Irani in India: report of a new polymorphic haplotype, $Gm^{1,3,21}$. *American Journal of Human Genetics*, **25**, 302–309.

VOGEL, F., KRÜGER, J., CHAKRAVARTTI, M. R., FLATZ, G., and RITTER, H., 1971, Inv phenotypes and quantitative gamma globulin determinations in leprosy patients and control populations from India and Thailand. *Humangenetik*, **12**, 35–41.

VOS, G. H., KIRK, R. L., and STEINBERG, A. G., 1963, The distribution of the gamma globulin types (Gm) (a), Gm (b), Gm (x) and Gm-like in South and South East Asia and Australia. *American Journal of Human Genetics*, **15**, 44–52.

WARD, R. H., and SINO, C. F., 1970, A consideration of the power of the χ^2 test to detect inbreeding effects in natural populations. *American Naturalist*, **104**, 355–365.

ZALESKI, M. B., DUBISKI, S., NILES, E. G., and CUNNINGHAM, R. K., 1983, *Immunogenetics* (Boston: Pitman).

Address correspondence to: Dr R. Chakraborty, Center for Demographic and Population Genetics, University of Texas Graduate School of Biomedical Sciences, PO Box 20334, Houston, TX 77225, USA.

Zusammenfassung. Blutproben von 898 Individuen aus neun endogamen Gruppen von West-Bengalen in Indien wurden auf die Haplotyp-Struktur der leichten Gamma-Kette (Gm), der schweren Gamma-Kette (Km) und des Immunoglobins (IgG) typisiert. Die Häufigkeiten der Gm-Haplotypen, untersucht durch die Merkmale G1m (1), G1m (2) und G3m (5), legen nahe, daß es in diesem östlichen Staat von Indien eine beträchtliche Variation der Frequenzen des typisch mongoliden Haplotyps $Gm^{1,5}$ gibt, der ein hohes Vorkommen bei den Gruppen Rajbanshi, Rabha, Garo und Lodha zeigt. Im Gegensatz dazu ist dieser Haplotyp wahrscheinlich bei der hohen Kastengruppe der Rarhi Brahmin und Vaidya abwesend und relativ selten bei den Jalia Kaibarta, einer eingetragenen Kaste des südwestlichen Teils des Staates. Das Allel Km^1 hat ebenfalls eine hohe Frequenz bei den Rajbanshi, Rabha, Garo und Munda im Vergleich mit Rarhi Brahmin und Vaidya, was die starke mongolide Herkunft der erstgenannten vier Gruppen unterstützt.

Diese Untersuchung weist darauf hin, daß es eine beträchtliche Variation im Ausmaß der mongoliden Beimischung bei bengalischen Bevölkerungen gibt. Diese Mischung beschränkt sich auch nicht auf spezifische sociale Klassen. Sie zeigt weiterhin, daß die Heterogenität der genetischen Struktur der bengalischen Bevölkerungen nicht der augenblicklichen sozialen Schichtung in Form der Kasten-Hierarchie entspricht.

Resumé. Des prélèvements sanguins sur 898 personnes provenant de neuf groupes endogames de l'ouest du Bengale (Inde), ont été typés afin de déterminer la structure haplotypique des chaînes legeres (Gm) et lourdes (Km) de l'immunoglobuline (IgG). Les fréquences haplotypiques Gm détectées par les marqueurs G1m(1), G1m(2) et G1m(5), suggèrent une variation considérable des fréquences de l'haplotype mongoloïde typique $Gm^{1,5}$ qui présente une incidence élevée dans les groupes *Rjabanshi, Rabha, Garo* et *Lodha*. Cet haplotype est probablement absent par contre, dans les castes élevées *Brâhmanes Rarhi* et *Vaidya* et peu fréquent chez les *Jalia Kaibarta*, un groupe d'Intouchables du sud-ouest de l'état. L'allèle Km^1 présente aussi des fréquences plus élevées chez les groupes *Rjabanshi, Rabha, Garo* et *Munda*, que chez les *Brâhmanes Rarhi* et les *Vaidya*, suggerant une forte affinité mongoloïde chez les premiers. Cette etude montre qu'il y a une variation importante de l'apport mongoloïde dans les populations du Bengale, mais que cet apport n'est pas limité à certains groupes sociaux. Elle démontre également que l'hétérogénéité génétique des populations du Bengale ne correspond pas à l'ordre social actuel, fondé sur la hiérarchie des castes.