ANNALS OF HUMAN BIOLOGY, 1982, VOL. 9, NO. 3, 235–251

Genetic heterogeneity and population structure in north-west India

S. S. PAPIHA[1], B. N. MUKHERJEE[2], S. M. S. CHAHAL[1], K. C. MALHOTRA[2] and D. F. ROBERTS[1]

[1] Department of Human Genetics, University of Newcastle upon Tyne, and [2] Anthropometric and Human Genetics Unit, Indian Statistical Institute, Calcutta, India

Received 23 October 1981; revised 15 December 1981

Summary. Genetic markers consisting of 11 blood group and red cell enzyme systems were investigated in 14 endogamous groups of north-west India. Genetic differentiation among the samples as indicated by F_{ST} is appreciable, reflecting the ethnic diversity characteristic of this region. Local variation within each state is lower, indicating a geographical component to the total variation. This variation is refined by calculations of genetic distances, which show that the tribals and low-caste groups are closer together but well separated from high-caste Brahmins and other non-tribal middle castes. There is a slight possibility of disruptive selection, but the analyses suggest that the differences in genetic structure in north-west India are more likely to be due to their breeding structure, differential migration and ethnic affiliation.

1. Introduction

North-west India is usually regarded as extending southwards as far as the river Narbada and therefore includes the northern part of the states of Gujarat and Madhya Pradesh. The western limits run along the border of Pakistan, whereas the eastern boundary is less well defined and may continue to the western borders of Bihar state. In the north, in spite of the high mountains, the Himalayan valleys have acted as access routes for many waves of migrants and for ages have promoted intermingling of many races in this region. The plateau of north-west India, therefore, contains populations with diverse ethnic elements. The complexity of ethnic variation was further compounded with the social restrictions in mating enforced by a highly complicated system of inter-regional caste groups which seems to be not very old, dating back to about 500 B.C.

Qualitative anthroposcopic and anthropometric studies show that the southern peripheral region of north-west India has a considerable proportion of an original autochthonous element, Dravidian and proto-Australoid, in constrast to the generally fair peoples of the northern states described broadly as Indo-Aryan. The other foreign influx in north India was the mongoloid, mainly concentrated in the hills and valleys of eastern India, but also penetrating obviously into the north-western states, especially Himachal Pradesh. The mixture of these ethnic elements is further indicated by the large number of languages and dialects prevalent in this region.

The study of the present-day population of north-west India, which is in fact an amalgamation of many streams of ethnic movements and whose social structure consists of many hierarchical sub-populations, poses many interesting problems. In certain states intermixing has occurred to such an extent that it is difficult to differentiate the populations on purely ethnic grounds. But the social structures of the populations provide a basis for genetic study of the various tribes and caste groups.

For genetic studies of the populations of north-west India, there are isolated reports from various states (Blake, Kirk, McDermid, Omoto, and Ahuja 1971, Papiha, Roberts, Wig and Singh 1972, Singh, Sareen and Goedde 1974, Papiha, Roberts and

Gulati 1976, Papiha, Chahal, Roberts and Singh 1980, Das, Mukherjee, Malhotra and Majumder 1978). These studies are either on heterogeneous State samples or from caste groups from a single State, and provide data on either red cell enzymes or serum proteins. From such data there are limitations to comment on the genetic affinities between populations of various states. The present paper reports part of a large Indo-Soviet project in which three north-western states of India, Rajasthan, Punjab and Himachal Pradesh (see figure 1) were extensively investigated for anthropometric, dermatoglyphic, odontology, body constitution and genetic variables. Here we report the comprehensive genetic analysis of the different tribal and caste groups for 11 single gene characters.

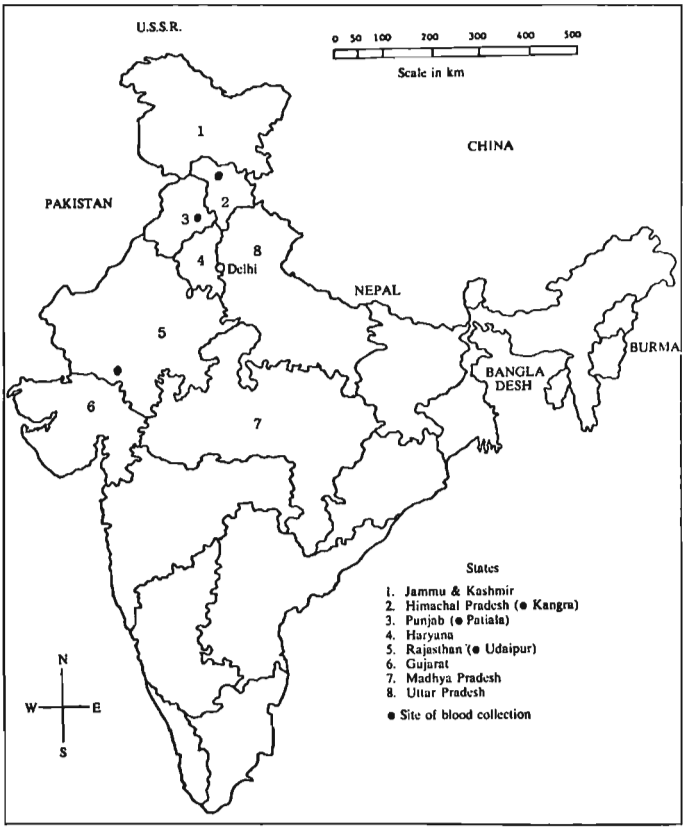

Figure 1. Map of India showing states and sites of blood collection.

2. Subjects and methods

Rajasthan State

The six populations selected for study were a high-caste group of Paliwal Brahmin (PB), a hierarchical caste population of Kshatriya known as Rajput (RR), a trader caste group of Oswal Mahajan (OM), a scheduled low caste, the Meghwal (MR) and two tribal groups Bhil (BT) and Meena or Mina (MT).

Punjab State

Three populations were sampled. The endogamous group Jat Sikh (JS) in Punjab corresponds to the Rajput caste of Rajasthan. Ramdasia Sikh (RS) is a separate endogamous group—a lower caste and followers of *Guru Ramdass*. The third group of Punjabi studied were the traders of Mahajan Agarwal (MA).

Himachal Pradesh

The remaining five groups came from the state of Himachal Pradesh and represent different castes. There were a high-caste group of Brahmin (BH), a group of Rajput (GR) but belonging to a tribe of Gaddi, an endogamous group of agriculturist Chowdhury (CW), a low-caste group of Chamar (CH) whose dominant trade is shoe-making. The fifth group in Himachal Pradesh is a heterogeneous immigrant mongoloid group of Nepali (NH), originally inhabitants of different parts of Nepal state.

Methods

The field work was conducted jointly by Indian and Soviet scientists in January–March 1979. For the study of various biological characters, surveys were planned in different districts of the three states of Rajasthan, Punjab and Himachal Pradesh. Subjects were classified according to tribe or caste group of the region. On each subject, after initial biological studies, about 8–10 drops of blood were collected by finger prick in a 1 ml vial containing EDTA as anticoagulant. Specimens were obtained from 1188

Table 1. Populations, their localities and sample sizes.

Population	Code of identification	Sample size
Rajasthan state, Udaipur district		
1. Paliwal Brahmin	PB	94
2. Rajput	RR	115
3. Oswal Mahajan	OM	104
4. Bhil Tribe	BT	73
5. Meghwal	MR	81
6. Meena Tribe	MT	80
Punjab state, Patiala district		
7. Jat Sikh	JS	102
8. Ramdasia Sikh	RS	76
9. Mahajan Agarwal	MA	104
Himachal Pradesh state, Kangra district		
10. Brahmin	BH	105
11. Chowdhury	CW	113
12. Gaddi Rajput	GR	76
13. Chamar	CH	42
14. Nepali	NH	23
Total		1188

individuals and transported on the day of collection by airfreight to the Department of Human Genetics, Newcastle upon Tyne in lots of 100–150 specimens, where the analysis was carried out using local samples as controls. The sample size, population, and location of samples are given in table 1.

The specimens were tested with the following antisera A, A_1, B, M, N, C, c, D, E, e, P and Le^a. Due to limitations of antisera the later batches could not be tested for P and Le^a. After blood grouping the remaining portion of each samples was lysed and stored for the analysis of red cell enzymes. The polymorphic systems, adenosine deaminase (ADA), esterase-D (EsD), adenylate kinase (AK) and phosphoglucomutase (PGM) were examined according to methods listed in Harris and Hopkinson (1976). For ADA and EsD the samples were cleared with 2-mercaptoethanol before electrophoresis. The variation of PGM-2 locus and superoxide dismutase were examined along with the above systems. The gene frequencies for ABO and Rh systems were calculated by the maximum likelihood method; for others the gene counting method was used.

3. Results

For each population, the observed and expected phenotype numbers for blood groups (table 2), red cell enzymes (table 3) and gene frequencies calculated from the observed numbers are given (table 4). In no system did the expected phenotype numbers calculated assuming Hardy–Weinberg equilibrium show any statistically significant variation from the observed figures.

The range of frequencies of various blood group systems and enzyme systems in 14 population groups investigated are compatible with the ranges described previously for north India, (Mourant, Kopeć and Domaniewska Sobczak 1976, Papiha 1982). The analysis showed that there are significant differences in gene frequency distribution in these various populations. PGM_1 was the only system which showed no significant variation ($\chi^2 = 23{\cdot}61$, d.f. = 26, N.S.); for all the others, genetic heterogeneity was suggested at the following significance levels: ABO, RhC, RhE and P (at 0·1%); Le^{a+} ADA, and AK (at 1%) and MN and EsD (at 5%). All individuals were of a single phenotype (1-1) at the PGM_2 and SOD loci.

Intra-state comparison of the various population groups demonstrated no significant variation in any of the four enzyme systems, in contrast to the well documented heterogeneity in certain blood group systems. Nevertheless, in the enzyme systems, there was an appreciable range of frequencies in all the four isozyme systems in Rajasthan populations, for ADA, AK and PGM_1 systems in the samples in Himachal Pradesh and for the PGM_1 system in Punjab.

For blood group systems Rajasthan populations are not well represented on the gene frequency map compared to the other north-western states, even for ABO distribution. ABO gene frequencies in Rajput by Negi and Das (1962) are not significantly different from the present results ($\chi^2 = 6{\cdot}27$, d.f. = 3, N.S.) and indeed the A_2 and B gene frequencies are very similar. The tribal Bhils showed the highest frequencies of the A_1 and A_2 genes (20% and 2·7% respectively), more similar to the Bhils of Madhya Pradesh, (Papiha, Roberts, Mukerjee, Singh and Malhotra 1978) than to those of Gujarat (Vyas, Bhatia, Sukumaran, Balakrishnan and Sanghvi 1962). The lowest A_1 gene frequency was found in low-caste Meghwal. These differences were at the expense of the O gene, since the B gene showed fairly consistent levels among the populations of Rajasthan (24–28%). In Punjab, the Jat Sikh and Mahajan Agarwal were similar in A_1 gene frequency (11–12%), which is slightly lower than the other reported samples from Punjab (Bhalla 1966), and from these the Ramdasia Sikh with

Table 2. Blood groups, phenotype number observed and expected in different population groups of north-west India.

System and phenotype	PB		RR		OM		BT		MR		MT		JS		RS		MA		BH		CW		GR		CH		NH	
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp
ABO																												
O	31	31·7	42	39·6	28	29·8	17	19·4	30	30·0	31	30·5	34	31·3	14	13·7	44	41·8	25	29·5	29	27·6	13	11·8	14	12·2	9	8·2
A_1	21	20·0	13	16·7	25	22·4	22	18·9	11	11·3	13	13·9	13	15·7	21	21·3	14	16·2	30	25·1	24	27·7	24	25·1	11	13·2	6	7·0
A_2	2	2·1	4	3·3	1	1·3	2	2·1	1	0·7	1	0·7	—	0·6	1	1·2	1	1·4	2	1·4	7	5·2	—	0·6	—	0·0	—	0·0
B	33	32·1	44	46·9	42	39·8	27	24·0	34	34·0	29	29·6	43	46·0	27	27·4	36	38·4	43	37·7	36	38·1	23	24·6	10	12·2	5	6·0
A_1B	6	7·2	11	7·0	7	10·0	4	7·5	5	4·7	6	5·0	11	8·0	12	11·7	8	5·7	5	10·7	16	11·7	15	13·6	7	4·5	3	1·9
A_2B	1	0·9	1	1·6	1	0·7	1	1·0	—	0·3	—	0·3	1	0·4	1	0·8	1	0·5	0	0·7	1	2·7	1	0·4	—	0·0	—	0·0
Total	94		115		104		73		81		80		102		76		104		105		113		76		42		23	
Rhesus																												
CCDEE	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	1	0·1	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0
CCDEe	—	0·6	—	0·0	—	0·0	—	0·9	1	1·3	—	0·9	2	3·2	—	0·0	1	1·6	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0
CCDee	30	31·0	37	38·5	44	46·4	36	32·7	23	26·6	52	51·1	22	22·7	20	22·1	42	41·3	44	44·0	69	68·5	33	32·2	23	25·0	11	10·4
CCddee	—	0·0	—	0·0	1	0·1	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·1	—	0·0	—	0·0	—	0·0	—	0·2	—	0·0
CcDEE	—	0·1	—	0·0	—	0·0	1	0·1	—	0·3	—	0·0	—	0·8	—	0·0	1	0·4	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0
CcDEe	18	13·2	16	17·4	7	6·0	8	8·8	17	13·2	9	10·4	15	14·0	13	13·5	17	19·2	13	13·0	11	13·2	6	6·5	8	6·2	6	6·7
CcDee	29	31·2	43	38·7	42	38·4	18	22·9	29	25·3	15	14·5	38	36·2	29	24·3	28	26·7	34	34·0	27	25·7	27	28·0	10	7·9	3	3·4
Ccddee	1	1·0	—	0·0	—	1·7	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	1	0·8	1	0·9	—	0·0	—	0·0	1	0·6	—	0·0
CcddEE	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	1	0·1	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0
CcddEe	1	0·3	—	0·0	—	0·0	—	0·0	1	0·4	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0
ccDEE	—	1·4	3	2·0	—	0·2	—	0·6	—	1·6	1	0·4	2	1·5	3	2·1	4	2·1	—	1·0	2	0·6	—	0·3	—	0·4	1	1·1
ccDEe	5	6·8	8	8·8	2	2·6	4	3·0	6	6·2	1	1·4	8	9·1	6	7·4	4	6·2	7	5·1	1	1·8	4	2·8	—	1·1	2	1·1
ccDee	2	1·6	3	3·7	—	0·0	4	2·7	2	2·9	1	1·0	5	5·2	2	2·7	2	1·8	1	1·2	1	1·1	3	3·0	—	0·2	—	0·0
ccddEe	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	1	0·7	—	0·0	—	0·0	—	0·0
ccddee	8	6·7	5	6·1	8	8·6	2	1·3	2	3·2	—	0·0	9	9·3	3	4·0	3	2·8	5	5·8	1	1·4	3	3·0	—	0·5	—	0·0
Total	94		115		104		73		81		80		102		76		103		105		113		76		42		23	
MN																												
M	38	37·0	54	54·8	46	49·2	32	29·6	26	26·1	21	23·1	51	52·3	28	25·5	42	43·2	39	36·0	42	40·9	29	27·8	15	12·1	12	12·6
MN	42	43·9	50	48·5	51	44·7	29	33·8	40	39·8	44	39·8	44	41·5	32	37·1	50	47·7	45	51·0	52	54·2	34	36·3	15	20·9	10	8·9
N	14	13·0	10	10·8	7	10·2	12	9·6	15	15·1	15	17·1	7	8·3	16	13·5	12	13·2	21	18·0	19	17·9	13	11·8	12	9·1	1	1·6
Total	94		114		104		73		81		80		102		76		104		105		113		76		42		23	
P																												
P^+	55	—	65	—	64	—	51	—	55	—	40	—	75	—	63	—	75	—										
P_1^-	39	—	50	—	40	—	22	—	26	—	40	—	27	—	13	—	28	—										
Total	94		115		104		73		81		80		102		76		103											
Lewis																												
Le^{a+}	19	—	27	—	34	—	5	—	23	—	11	—	19	—	16	—	17	—										
Le^{a-}	75	—	88	—	70	—	68	—	58	—	69	—	83	—	60	—	87	—										
Total	94		115		104		73		81		80		102		76		104											

Table 3. Red cell enzymes, phenotype numbers observed and expected in different population groups of north-west India.

System and phenotype	PB		RR		OM		BT		MR		MT		JS		RS		MA		BH		CW		GR		CH		NH	
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp
ADA																												
1–1	41	42·9	64	64·1	80	80·0	40	39·6	59	58·8	54	53·6	79	79·4	63	59·1	77	74·5	89	88·7	100	98·5	66	65·4	34	33·5	19	19·2
2–1	22	18·2	17	16·8	18	18·0	9	9·8	13	13·5	23	23·7	22	21·2	8	15·9	22	27·1	15	15·6	11	14·0	9	10·2	7	8·0	4	3·7
2–2	—	1·9	1	1·1	1	1·0	1	0·6	1	0·8	3	2·6	1	1·4	5	1·1	5	2·5	1	0·7	2	0·5	1	0·4	1	0·5	—	0·2
Total	63		82		99		50		73		80		102		76		104		105		113		76		42		23	
AK																												
1–1	54	54·3	67	67·7	83	83·7	40	40·5	59	59·7	67	67·5	86	86·6	62	62·6	88	88·6	89	89·6	91	92·0	60	60·8	35	35·3	21	21·0
2–1	8	7·5	15	13·6	16	14·7	10	9·0	14	12·7	13	11·9	16	14·8	14	12·7	16	14·8	16	14·8	21	19·0	16	14·3	7	6·4	2	1·9
2–2	—	0·3	—	0·7	—	0·7	—	0·5	—	0·7	—	0·5	—	0·6	—	0·6	—	0·6	—	0·6	—	1·0	—	0·8	—	0·3	—	0·0
Total	62		82		99		50		73		80		102		76		104		105		112		76		42		23	
PGM																												
1–1	30	29·1	37	38·2	48	45·3	26	25·6	34	32·9	34	30·8	53	53·7	30	31·6	61	60·8	48	49·4	52	51·8	37	38·4	14	14·1	12	12·4
2–1	25	26·7	36	34·2	38	43·3	19	20·6	30	32·2	30	36·4	42	40·6	38	34·8	37	37·5	48	45·3	47	48·1	34	31·3	20	19·9	9	8·3
2–2	7	6·1	7	7·6	13	10·3	5	4·2	9	7·9	14	10·8	7	7·7	8	9·6	6	5·8	9	10·4	12	11·2	5	6·4	7	7·1	1	1·4
7–1	—	0·0	2	1·4	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	—	0·0	2	1·4	—	0·0	—	0·0	—	0·0
Total	62		82		99		50		73		78		102		76		104		105		113		76		41		22	
EsD																												
1–1	43	42·8	42	42·1	51	51·6	23	20·5	38	36·3	36	37·4	64	65·1	51	49·8	69	66·2	53	57·2	57	56·9	48	47·4	25	25·9	12	11·6
2–1	17	17·4	32	31·9	41	39·7	16	21·1	27	30·3	36	33·2	35	32·8	21	23·5	28	33·5	49	40·6	45	45·1	24	25·3	16	14·1	8	8·7
2–2	2	1·8	6	6·1	7	7·6	8	5·5	8	6·3	6	7·4	3	4·1	4	2·8	7	4·2	3	7·2	9	8·9	4	3·4	1	1·9	2	1·6
Total	62		80		99		47		73		78		102		76		104		105		111		76		42		22	

System gene/haplotype	PB	RR	OM	BT	MR	MT	JS	RS	MA	BH	CW	GR	CH	NH
ABO														
P_1	0·1565	0·1090	0·1700	0·2017	0·1041	0·1262	0·1240	0·2473	0·1111	0·1878	0·1931	0·2994	0·2380	0·2156
P_2	0·0193	0·0241	0·0119	0·0273	0·0069	0·0071	0·0055	0·0176	0·0107	0·0124	0·0441	0·0092	0·0000	0·0000
q	0·2431	0·2804	0·2825	0·2560	0·2800	0·2492	0·3164	0·3108	0·2443	0·2702	0·2685	0·2982	0·2228	0·1888
r	0·5811	0·5865	0·5355	0·5150	0·6091	0·6175	0·5541	0·4243	0·6339	0·5296	0·4942	0·3932	0·5393	0·5956
Rhesus														
CDe	0·5544	0·5783	0·6402	0·6691	0·5727	0·7943	0·4720	0·5395	0·6091	0·6288	0·7786	0·6513	0·7122	0·6612
cDE	0·1221	0·1304	0·0433	0·0872	0·1406	0·0349	0·1191	0·1645	0·1434	0·0952	0·0476	0·0657	0·0951	0·2133
cde	0·2677	0·2303	0·2884	0·1354	0·1977	0·0007	0·3015	0·2293	0·1641	0·2349	0·1096	0·2000	0·1109	0·0021
Cde	0·0198	0·0000	0·0281	0·0003	0·0000	0·0051	0·0000	0·0000	0·0246	0·0188	0·0001	0·0000	0·0616	0·0127
cdE	0·0000	0·0000	0·0000	0·0000	0·0000	0·0457	0·0000	0·0000	0·0000	0·0000	0·0276	0·0000	0·0001	0·0041
CDE	0·0000	0·0000	0·0000	0·0087	0·0000	0·0000	0·0329	0·0000	0·0119	0·0000	0·0000	0·0000	0·0000	0·0000
CdE	0·0055	0·0000	0·0000	0·0000	0·0137	0·0069	0·0000	0·0000	0·0000	0·0000	0·0000	0·0000	0·0000	0·0000
cDe	0·0304	0·0610	0·0000	0·0993	0·0752	0·1124	0·0745	0·0667	0·0469	0·0223	0·0365	0·0829	0·0200	0·1066
MN														
M	0·6277	0·6930	0·6875	0·6370	0·5679	0·5375	0·7157	0·5789	0·6442	0·5857	0·6018	0·6053	0·5357	0·7391
N	0·3723	0·3070	0·3125	0·3630	0·4321	0·4625	0·2843	0·4211	0·3558	0·4143	0·3982	0·3947	0·4643	0·2609
P														
P_1	0·3559	0·3406	0·3798	0·4510	0·4334	0·2929	0·4855	0·5864	0·4786					
p_2^2p	0·6441	0·6594	0·6202	0·5490	0·5666	0·7071	0·5145	0·4136	0·5214					
ADA														
ADA^1	0·8254	0·8841	0·8990	0·8900	0·8973	0·8188	0·8824	0·8816	0·8462	0·9190	0·9336	0·9276	0·8929	0·9130
ADA^2	0·1746	0·1159	0·1010	0·1100	0·1027	0·1813	0·1176	0·1184	0·1538	0·0810	0·0664	0·0724	0·1071	0·0870
AK														
AK^1	0·9355	0·9085	0·9192	0·9000	0·9041	0·9188	0·9216	0·9079	0·9231	0·9238	0·9063	0·8947	0·9167	0·9565
AK^2	0·0645	0·0915	0·0808	0·1000	0·0959	0·0813	0·0784	0·0921	0·0769	0·0762	0·0938	0·1053	0·0833	0·0435
PGM														
PGM^1	0·6855	0·6829	0·6768	0·7100	0·6712	0·6282	0·7255	0·6447	0·7644	0·6857	0·6770	0·7105	0·5854	0·7500
PGM^2	0·3145	0·3049	0·3232	0·2900	0·3288	0·3718	0·2745	0·3553	0·2356	0·3143	0·3142	0·2895	0·4146	0·2500
PGM^7	—	0·0122	—	—	—	—	—	—	—	—	0·0088	—	—	—
EsD														
EsD^1	0·8306	0·7250	0·7222	0·6596	0·7055	0·6923	0·7990	0·8092	0·7981	0·7381	0·7162	0·7895	0·7857	0·7273
EsD^2	0·1694	0·2750	0·2778	0·3404	0·2945	0·3077	0·2010	0·1908	0·2019	0·2619	0·2838	0·2105	0·2143	0·2727

double the frequency of the A_1 gene are clearly differentiated. The Mahajan Agarwal deviated from Jat Sikh in the B gene frequency (24% compared to 31%) and in the highest O gene frequency ever reported from Punjab (63%). The interpopulation variation in Punjab was significant ($\chi^2 = 18{\cdot}6$, d.f. $= 6$, $P < 0{\cdot}01$).

In the cis-Himalayan region of Himachal Pradesh, there is a good deal of variation in the incidence of A and B genes among the local populations. Brahmin and Chowdhury have a much higher frequency of the A than the B gene, whereas in the more interior Gaddi Rajput and Chamar, the frequency of B equals A, and in the Nepali there is a preponderance of A over B. This trend of a higher A than B gene frequency is a regular feature of mongoloid populations in this region. In spite of an appreciable frequency range of the A, B and O genes, the overall heterogeneity among populations of Himachal state failed to attain statistical significance ($\chi^2 = 19{\cdot}4$, d.f. $= 12$, N.S.).

In the Rh system the incidence of the most frequent chromosome CDe(R_1) in India generally ranges between 50 and 70%, although an exceptionally high incidence of 85% was reported in the Oraon tribe. In the north-west zone, this wider frequency range of chromosome CDe(R_1) was also found. The highest occurred in the Mina tribe of Rajasthan (79%) and lowest in the Jat Sikh of the Punjab (47%). The cis-Himalayan populations tend to show higher incidences (range 63–78%) of CDe(R_1) than the states of Punjab and Rajasthan. The incidence of Rh-negatives is lower in the tribal populations and this phenotype is completely absent in the Mina tribe and the Nepali, in characteristic contrast to its high incidence in the rest of the populations of the northern region. The incidence of the rare chromosome Cde(r′) is greater than cdE(r″). CDE(Rz) is only present in Bhil, Jat Sikh and Mahajan Agarwal, but this haplotype is completely absent in Himachal state populations. Of particular interest is the occurrence of one individual in the Mina tribe of rhesus phenotype CcddEE, indicating the presence of the very rare chromosome combination CdE (r^y); this is not a technical artifact, as it was checked against several anti-C, -c, -E, -e antisera with consistent results. It was necessary therefore to allow for its presence in calculating gene frequencies in the other populations.

The M gene frequency in the north-west Indian populations fluctuates between 54 and 74%. In Rajasthan the highest incidence was found in Rajput and Oswal Mahajan (69%), the lowest in the Mina tribe (54%) and intermediate in Brahmin, Bhil and Meghwal, representing significant heterogeneity ($\chi^2 = 19{\cdot}1$, d.f. $= 10$, $P < 0{\cdot}05$). In the three populations of the Punjab, the Jat Sikh frequencies of the M gene significantly differ from those of Ramdasia Sikh and Mahajan Agarwal ($\chi^2 = 9{\cdot}6$, d.f. $= 4$, $P < 0{\cdot}05$). In Himachal the Nepali show the high M gene frequency of 74%, which is within the previous known range (60–78%), but the remaining four population groups show a very narrow range (54–60%).

The incidence of blood group P and Le^a antigens was studied only in Rajasthan and Punjab samples. The incidence of the P_1 gene ranges from 30 to 45% in the populations of Rajasthan, and there is a suggestion of a slight increase in P_1 frequency eastwards, for the range in Punjab was from 48 to 58%. The χ^2 analysis showed no significant heterogeneity in either state. The incidence of Le^{a+} shows a significant variability in Rajasthan due to an exceptionally low frequency of Le^{a+} in Bhil (7%) ($\chi^2 = 22{\cdot}3$, d.f. $= 5$, $P < 0{\cdot}001$); the range is much smaller in Punjab populations (16–21%). Due to the interaction of Lewis with secretor system and due to the wide range of secretor frequencies in India, the above variation should be interpreted in the light of the distribution of Se gene for which the information is unfortunately lacking in the present study.

Population structure

Despite the marked differences in gene frequencies observed in these fourteen population groups of north-west India, the extent of genetic differentiation was not of the same magnitude among the populations within each state, though there was an appreciable range of gene frequency for all the loci studied. The desert areas of Rajasthan, the fertile agricultural land of Punjab, and the mountain ranges of Himachal, provide different ecological and environmental conditions that must surely exert different selective pressures. By contrast, several populations are perhaps more subject to similar selective factors. The different conditions of the three states, however, suggest that the extent of variation may well be different, so that random genetic drift, inbreeding or other non-selective processes may contribute differentially to genetic variation among the groups of each state.

Wahlund (1982) described a variance of gene frequency f of any given allele as an index of deviation from panmixia.

$$f = \frac{\sigma_p^2}{\bar{p}(1-\bar{p})}$$

The value of f has been computed for each locus over all the caste and tribal groups of each state.

Table 5 shows the estimate of weighted variance of gene frequencies and Wahlund variance f, for the alleles M, C, E, P_1, ADA^1, AK^1, PGM^1 and EsD^1. The mean f for the total north-west Indian population is $1{\cdot}54 \times 10^{-2}$ with variance $(S_f^2)\ 9{\cdot}1 \times 10^{-5}$. For the state of Rajasthan the mean f is $1{\cdot}154 \times 10^{-2}$, which is higher than those of Punjab and Himachal ($6{\cdot}461 \times 10^{-3}$ and $8{\cdot}358 \times 10^{-3}$ respectively), with variances over these eight alleles respectively $5{\cdot}5 \times 10^{-5}$, $3{\cdot}5 \times 10^{-5}$ and $3{\cdot}3 \times 10^{-5}$. These observed variances of f are quite low and indicate that there is no significant diversity in the f value for the different alleles considered. These values can be compared with the expected variance (σ_f^2). The expected variance can be calculated according to Lewontin and Krakauer (1973), which assumes that mean frequency $\bar{p}$ and the mean $\bar{f}$ are the estimates from the random samples taken from the subpopulations, the number of which (n) is fairly large

$$\sigma_f^2 = \frac{2\bar{f}^2}{(n-1)}$$

where 2 is a constant value used for the distribution pattern of gene frequency.

The expected variance found for Rajasthan is $5{\cdot}3 \times 10^{-5}$, Punjab $4{\cdot}2 \times 10^{-5}$, Himachal $3{\cdot}5 \times 10^{-5}$ and for the total north-west Indian population $3{\cdot}6 \times 10^{-5}$. The expected variance was not significantly different from the observed variance (Rajasthan $F = 1{\cdot}03$, d.f. $= 7/\infty$, $P > 0{\cdot}05$; Punjab $F = 0{\cdot}84$, d.f. $= 7/\infty$, $P > 0{\cdot}05$; Himachal $F = 0{\cdot}94$, d.f. $= 6/\infty$, $P > 0{\cdot}05$ and north-west India $F = 2{\cdot}53$, d.f. $= 7/\infty$, $0{\cdot}01 < P < 0{\cdot}05$).

The homogeneous distribution of f over all loci in the three state population groups suggests that the variation in the gene frequency within each state is not due to selection, but to some non-selective process which contributes to the present structure of these populations. The heterogeneity among the total north-west India may be due to the differences in the effect of the non-selective processes in different states, or it may involve a selective element.

In India, the different caste and tribal groups are largely inbreeding isolates. A member of a particular caste group has a greater probability of mating with another from the same group than with an individual from outside. The contribution of this

Table 5. Comparison of Wahlund's variance for eight alleles in population groups of three states and total north-west India.

Mean frequency, Variance of p, Wahlund's f	Allele								Mean
	M	C	E	P_1	ADA^1	AK^1	PGM^1	EsD^1	
Rajasthan (6)									
$\bar{p}$	0·63187	0·64351	0·10603	0·37220	0·87025	0·91480	0·67421	0·72323	
σ_p^2	0·00329	0·00623	0·00136	0·00257	0·00114	0·00012	0·00057	0·00229	0·00220
f	0·01410	0·02714	0·01439	0·01100	0·01009	0·00151	0·00260	0·01142	0·01154
Punjab (3)									
$\bar{p}$	0·65248	0·56560	0·15603	0·51027	0·86879	0·91844	0·71809	0·80142	
σ_p^2	0·00293	0·00380	0·00003	0·00216	0·00030	0·00004	0·00226	0·00002	0·00144
f	0·01291	0·01545	0·00021	0·00864	0·00263	0·00055	0·01117	0·00014	0·00646
Himachal Pradesh (5)									
$\bar{p}$	0·59889	0·70613	0·09053	—	0·92201	0·91341	0·68169	0·74719	
σ_p^2	0·00179	0·00390	0·00124	—	0·00016	0·00024	0·00154	0·00091	0·00140
f	0·00745	0·01880	0·01510	—	0·00218	0·00307	0·00712	0·00479	0·00836
North-west India (14)									
$\bar{p}$	0·62679	0·64394	0·11322	0·41905†	0·88695	0·91529	0·68814	0·75162	
σ_p^2	0·00315	0·00757	0·00163	0·00671†	0·00120	0·00014	0·00166	0·00222	0·00304
f	0·01348	0·03302	0·01618	0·02754†	0·01200	0·00185	0·00774	0·01190	0·01547

Number of population groups in each region is given in parentheses.
† For nine population groups tested for P.

inbreeding to population structure can be further examined by using Wright's (1943, 1965) *F*-statistics:

F_{IS} is the coefficient of local inbreeding, and measures the probability that homologous genes in a given individual are identical by descent from a common ancestor within the subpopulation. F_{IS} is obtained from the mean heterozygote proportions within all subpopulations.

F_{IT} is the coefficient of individual inbreeding and indicates the correlation of uniting gametes relative to the gametes of the total population. F_{IT} can be calculated from the observed proportion of heterozygotes in the total population.

F_{ST} is the coefficient of genetic differentiation among sub-populations and can be defined as the correlation between two gametes derived at random within subpopulations relative to the genes of the total population.

Table 6 gives the values of these three statistics. In the total sample of north-west India the mean F_{ST} value found is 0·013; the AK allele gives the lowest estimate, the Rh C allele the highest. The estimates of F_{ST} among the subpopulations of the different states indicate that genetic differentiation is slightly higher in Rajasthan compared to the Punjab and Himachal, which are quite similar to each other. The negative value of F_{ST} for esterase D in Punjab is most probably an artefact of sample size, for all other estimates are positive. The F_{IT} value in Punjab and Himachal are distributed around a positive mean of 0·034 and 0·006, but in Rajasthan F_{IT} around a negative mean of 0·002. Within each state the F_{IT} estimates from different alleles are fairly similar. The samples are small, so that there is a relatively large random element in the deviation from equilibrium of heterozygosity levels in each. There may be selective pressures for or against heterozygosity, and these pressures may vary from population and more so from state to state. Hence interpretation of these *F* values is tentative and can be attempted only in the broadest terms.

In Punjab there is an appreciable tendency to overall inbreeding due mainly to such a tendency within each local population, but reinforced by the inbreeding effects of subdivision. In Himachal Pradesh there is less tendency to overall inbreeding, due mainly to the effects of subdivision and counteracted by a very slight inbreeding tendency within local populations. In Rajasthan there is a greater tendency to outbreeding within local populations, but this is largely, though not fully, counteracted by the effects of population subdivision, so that overall there is a tendency to outbreeding.

Another measure of genetic differentiation among these sub-divided caste and tribal groups of the different states, which may point to their relationship along the evolutionary time-scale, comes from their absolute gene frequency differences. For all the fourteen populations of north-west India, genetic distance matrices were computed using the *E* statistic of Edwards (1971) and standardized genetic distance measure *D* of Nei (1973). These matrices were calculated from the common systems studied for all populations. Edward's *E* matrix allows comparison of distances among populations envisaged as situated on the surface of a sphere in multidimensional space. If p_i and p'_i are the allele frequencies in two random populations and k is the number of alleles at a given locus, the matrix E is given by

$$E^2 = \frac{8\{1 - \sum\sqrt{(p_i p'_i)}\}}{(1 + \sum\sqrt{p_{i/k}})(1 + \sum\sqrt{p'_{i/k}})}.$$

Table 6. Comparison of a set of *F*-statistics for seven codominant alleles in population groups of three states and total north-west India.

F-statistic	Locus							Mean
	MN	Cc	Ee	ADA	AK	PGM	EsD	
Rajasthan (6)								
F_{IT}	−0·00783	−0·02005	−0·00289	−0·01042	−0·09314	0·08329	0·03839	−0·00181
F_{IS}	−0·02337	−0·04351	−0·00525	−0·01206	−0·09329	0·07930	0·02515	−0·01043
F_{ST}	0·01519	0·02249	0·00235	0·00163	0·00014	0·00433	0·01358	0·00853
Punjab (3)								
F_{IT}	0·01475	−0·03195	0·11134	0·19117	−0·08880	−0·02474	0·06416	0·03371
F_{IS}	−0·00295	−0·04699	0·11125	0·18872	−0·08886	−0·03237	0·06432	0·02759
F_{ST}	0·01765	0·01436	0·00011	0·00303	0·00005	0·00739	−0·00018	0·00606
Himachal Pradesh (5)								
F_{IT}	0·09554	0·01338	0·00196	0·10909	−0·09480	−0·02556	−0·05581	0·00626
F_{IS}	0·08507	−0·01073	−0·00495	0·11316	−0·09511	−0·03450	−0·05989	−0·00099
F_{ST}	0·01144	0·02385	0·00687	−0·00460	0·00029	0·00864	0·00386	0·00719
North-west India (14)								
F_{IT}	0·03122	−0·00225	0·04001	0·08337	−0·09255	0·02184	0·01770	0·01419
F_{IS}	0·01428	−0·03443	0·02249	0·08130	−0·09274	0·01267	0·00730	0·00155
F_{ST}	0·01718	0·03111	0·01792	0·00225	0·00017	0·00929	0·01049	0·01263

Number of population groups in each region is given in parentheses.

Table 7. Estimates of Nei's genetic distance *D* (above diagonal) and Edwards's genetic distance *E* (below diagonal) among the fourteen population groups of north-west India.

	PB	RR	OM	BT	MR	MT	JS	RS	MA	BH	CH	GR	CH	NH
PB		0·0055	0·0072	0·0137	0·0077	0·0250	0·0054	0·0070	0·0041	0·0059	0·0169	0·0129	0·0133	0·0195
RR	0·0789		0·0031	0·0061	0·0043	0·0218	0·0042	0·0112	0·0046	0·0049	0·0117	0·0135	0·0167	0·0116
OM	0·0893	0·1130		0·0064	0·0080	0·0232	0·0075	0·0110	0·0092	0·0034	0·0087	0·0095	0·0141	0·0176
BT	0·1044	0·0758	0·1288		0·0064	0·0121	0·0158	0·0130	0·0095	0·0047	0·0038	0·0083	0·0119	0·0091
MR	0·0817	0·0623	0·1269	0·0780		0·0146	0·0113	0·0100	0·0066	0·0033	0·0103	0·0142	0·0109	0·0167
MT	0·1888	0·1828	0·2131	0·1468	0·1638		0·0391	0·0281	0·0192	0·0170	0·0095	0·0253	0·0103	0·0223
JS	0·0909	0·0779	0·1282	0·0955	0·0913	0·2085		0·0113	0·0078	0·0103	0·0235	0·0158	0·0272	0·0202
RS	0·0780	0·0762	0·1263	0·0845	0·0780	0·1902	0·0900		0·0136	0·0056	0·0142	0·0049	0·0104	0·0233
MA	0·0648	0·0807	0·1106	0·0856	0·0858	0·1679	0·0819	0·0980		0·0073	0·0135	0·0153	0·0159	0·0103
BH	0·0605	0·0770	0·0653	0·0841	0·0797	0·1811	0·0999	0·0755	0·0750		0·0054	0·0056	0·0071	0·0163
CH	0·1257	0·1030	0·1244	0·0861	0·1128	0·1320	0·1416	0·1153	0·1219	0·0969		0·0074	0·0069	0·0146
GR	0·1028	0·0929	0·1221	0·0709	0·0952	0·1774	0·1050	0·0629	0·1114	0·0780	0·0935		0·0113	0·0198
CH	0·1001	0·1307	0·1077	0·1180	0·1206	0·1568	0·1497	0·1157	0·1015	0·0778	0·1205	0·1154		0·0215
NH	0·1732	0·1630	0·2068	0·1369	0·1561	0·1352	0·1826	0·1665	0·1390	0·1648	0·1564	0·1644	0·1456	

Population identification codes are given in table 1.

For several loci information is combined as follows:

$$\sum_{\text{loci}} E^2 / \sum_{\text{loci}} (k-1).$$

Nei's D matrix measures the number of mutant or net codon differences per locus and is given as

$$D = -\log_e I,$$

where I is the identity of two populations over all loci and is

$$I = J_{XY}/\sqrt{(J_X J_Y)},$$

in which J_X and J_Y are the average homozygosity per locus in population X and Y respectively and J_{XY} is the average identity of the genes between X and Y.

The distance matrices E and D thus calculated are given in table 7. For easier understanding of the major differences and phenotypic relationships of the various castes and tribal groups, both these genetic matrices were reduced by the construction of dendrograms (figures 2 and 3) by furthest neighbourhood or complete linkage method as given in Everitt (1974). There was a definite correlation between the two dendrograms in the hierarchical similarities. In both dendrograms the Brahmin groups of Rajasthan and Himachal come out in one cluster near to the Mahajan Agarwal of Punjab. The mongoloid population of Nepali does not show any association, but differentiates quite early from the cluster of the populations of Chamar, Mina and Chowdhury. These three populations by Nei's distance show close association with the Bhil tribe of Rajasthan. The differentiation of these tribal and the mongoloid Nepali groups from the remainder could well be due to a considerable mongoloid admixture in these tribes, as demonstrated by the presence of a large proportion of mongoloid Gm

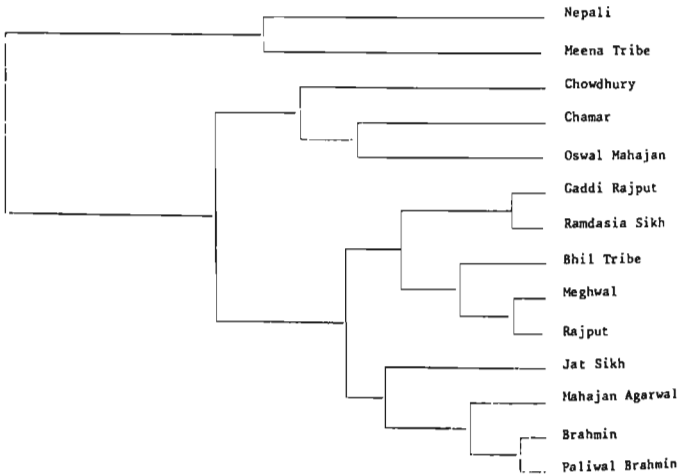

Figure 2. Dendrogram of the fourteen population groups of north-west India constructed from Edwards's distance matrix.

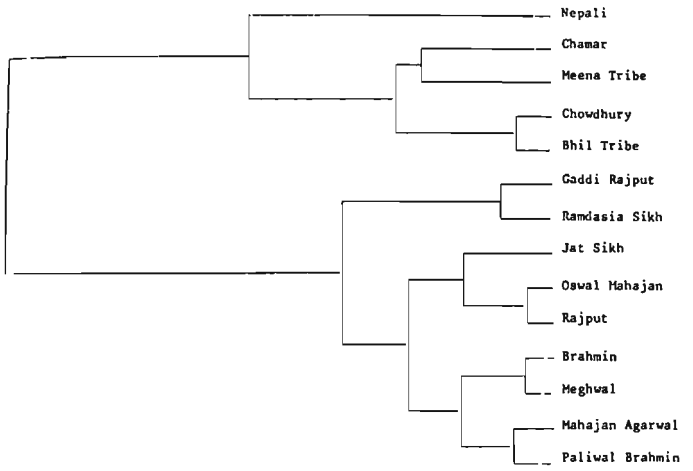

Figure 3. Dendrogram of the fourteen population groups of north-west India constructed from Nei's distance matrix.

haplotypes in Bhils (unpublished observation). The overall pattern of the clusters indicates that, within a given geographical area, it is the internal structure that regulates genetic differentiation of the various subpopulations. At the less local level, both geographical distance and major ethnic affiliation exert a primary influence.

4. Discussion

Studies of the pattern of genetic heterogeneity and the infrastructure of inter- and intra-regional populations of the Indian subcontinent are few. With respect to inter-regional differences, genetic data on the seven different state populations were examined by Roychoudhury (1977), who concluded that genetic differentiation between different state populations is very small; due to the heterogeneous nature of the samples considered in that study the results should be treated with caution. Balakrishnan (1978), using data on ABO and Rh subtype D for 102 endogamous groups, demonstrated considerable homogeneity among populations of north-west India. In a review article Balakrishnan (1981) further pointed out that homogeneity was increased in certain regions by excluding the tribal populations. In the present results, the genetic differentiation (F_{ST}) among all the samples from the three states of the north-west is higher than the coefficient of genetic differentiation of seven states of India given by Roychoudhury (1977), a difference which, it may be argued, could be due to the greater diversity of ethnic composition of the present population, consisting as it does of elements of autochthonous, caucasoid and mongoloid origin. More comparable is another study of six caste groups of north India, where the extent of microdifferentiation is similar to the present total sample of north-west India (Roychoudhury 1974). These results indicate that genetic differentiation of tribal groups and

hierarchical caste groups of different states is of similar magnitude. The extent of gene diversity due to genetic differences in subdivided populations of each state is also similar (F_{ST} 0·006–0·009).

The negative values (F_{IS}) for inbreeding within state populations in Himachal and Rajasthan have no simple explanation. They may show selection in favour of heterozygotes, avoidance of consanguinity, differential fertility, linkage disequilibrium between genes, and sex-related differences in gene frequencies. Should they be attributable solely to inbreeding *per se*, these estimates would possibly suggest a decline in the tendency of close mating in Rajasthan and Himachal, whereas Punjab caste groups seem to be strictly endogamous.

Differences that exist in the genetic constitution of tribal and caste groups of north-west India are particularly apparent from the genetic distance studies. By both Edwards' and Nei's indices, the tribals and low-caste groups (Chamar) showed closer genetic affinity and form a different cluster from the high-caste Brahmin and other non-tribal middle castes. Slight differences in the alignment within clusters of non-tribal groups may be the consequence of sampling error attaching to each genetic distance estimate or, if real, of relatively recent separation and adjustment of social structure. Similar results have been obtained from nine endogamous groups of western India using Nei's index by Mukherjee, Majumder, Malhotra, Das, Kate and Chakraborty (1979).

In conclusion, it seems that there is considerable genetic differentiation between populations of the Indian north-west with ancestral and immigrant elements of caucasoid and mongoloid. There is little possibility of disruptive selection by the restricting mechanisms of the population structure on gene flow. There are as yet few studies of genetic composition and estimates of genetic affinities to allow identification of forces which maintain the differences in genetic constitution of the populations of India.

Acknowledgements

We are grateful to Dr R. S. Mann, Dr I. J. S. Bansal and Dr L.S. Sidhu who kindly helped in establishing rapport with the subjects. Shri Badal Dey helped in collection of the blood samples, we thank him. Our thanks are also due to Mr C. K. Creen for laboratory assistance and Mrs Anthea Stephenson for statistical analysis.

References

Balakrishnan, V., 1978, A preliminary study of genetic distances among some populations of the Indian sub-continent. *Journal of Human Evolution*, **7**, 67–75.

Balakrishnan, V., 1981. Genetic distance studies in India. *Acta Anthropogenetica*, **5**, 57–65.

Bhalla, V., 1966, Blood-group distribution pertaining to ABO, MNSs and Rh-Hr systems in the Indian sub-continent (an ethno-geographical variation). *Anthropologie Brno*, **4**, 67–86.

Blake, N. M., Kirk, R. L., McDermid, E. M., Omoto, K., and Ahuja, Y. R., 1971, The distribution of serum protein and enzyme group systems among north Indians. *Human Heredity*, **21**, 440–457.

Das, S. K., Mukherjee, B. N., Malhotra, K. C., and Majumder, P. P., 1978, Seriological and biochemical investigations among five endogamous groups of Delhi, India. *Annals of Human Biology*, **5**, 25–31.

Edwards, A. W. F., 1971, Distances between populations on the basis of gene frequencies. *Biometrics*, **27**, 873–881.

Everitt, B. S., 1974, *Cluster Analysis* (London: Heinemann).

Harris, H., and Hopkinson, D. A., 1976, *Handbook of Enzyme Electrophoresis in Human Genetics* (Amsterdam: North-Holland).

Lewontin, R. C., and Krakauer, J., 1973, Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

Mourant, A. E., Kopeč, A. C., and Domaniewska-Sobczak, K., 1976, *The Distribution of the Human Blood Groups and other Polymorphisms* (London: Oxford University Press), 2nd edition.

MUKHERJEE, B. N., MAJUMDER, P. P., MALHOTRA, K. C., DAS, S. K., KATE, S. L., and CHAKRABORTY, R., 1979. Genetic distance analysis among nine endogamous population groups of Maharashtra, India. *Journal of Human Evolution*, **8**, 567–570.

NEGI, R. S., and DAS, A., 1962. The blood groups (ABO, MN and Rh), ABH secretion in saliva, and colour-blindness in the Rajput of western U.P. and Dholpur. *Bulletin of Anthropological Survey of India*, **11**, 221–231.

NEI, M., 1973. The theory and estimation of genetic distance. In *Genetic Structure of Populations*, edited by N. E. Morton (Honolulu: University of Hawaii Press), p. 45.

PAPIHA, S. S., 1982. Distribution of some polymorphic red cell enzymes in the Indian subcontinent. In *Medical Genetics in India*, edited by I. C. Verma, volume 3 (Pondicherry: Auroma Enterprises), pp. 65–96.

PAPIHA, S. S., ROBERTS, D. F., WIG, N. N., and SINGH, S., 1972. Red cell enzyme polymorphisms in Punjabis in North India. *American Journal of Physical Anthropology*, **37**, 293–299.

PAPIHA, S. S., ROBERTS, D. F., and GULATI, P. D., 1976. Some erythrocyte enzyme and serum protein systems in two communities of north-west India. *Human Biology*, **48**, 323–336.

PAPIHA, S. S., ROBERTS, D. F., MUKERJEE, D. P., SINGH, S. D., and MALHOTRA, M., 1978. A genetic survey in the Bhil tribe of Madhya Pradesh, Central India. *American Journal of Physical Anthropology*, **49**, 179–186.

PAPIHA, S. S., CHAHAL, S. M. S., ROBERTS, D. F., and SINGH, I. P., 1980. Genetic studies among Kanet and Koli district in Himachal Pradesh, India. *America Journal of Physical Anthropology*, **53**, 275–283.

ROYCHOUDHURY, A. K., 1974. Gene differentiation among caste and linguistic populations of India. *Human Heredity*, **24**, 317–322.

ROYCHOUDHURY, A. K., 1977. Gene diversity in Indian populations. *Human Genetics*, **40**, 99–106.

SINGH, S., SAREEN, K. N., and GOEDDE, H. W., 1974. Investigation of some biochemical genetic markers in four endogamous groups in Punjab (N.W. India) II. Red cell enzyme polymorphisms. *Humangenetik*, **22**, 133–138.

VYAS, G. N., BHATIA, H. M., SUKUMARAN, P. M., BALAKRISHNAN, V., and SANGHVI, L. D., 1962. Study of blood groups, abnormal hemoglobins and other genetic characters in some tribes of Gujarat. *American Journal of Physical Anthropology*, **20**, 255–265.

WAHLUND, S., 1928. Zusammensetzung von Populationen und Korrelationserscheinungen von Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65–106.

WRIGHT, S., 1943. Isolation by distance. *Genetics*, **28**, 114–138.

WRIGHT, S., 1965. The interpretation of population structure by *F*-statistics with special regard to systems of mating. *Evolution*, **19**, 395–420.

Address correspondence to: Professor D. F. Roberts, Department of Human Genetics, University of Newcastle upon Tyne, 19 Claremont Place, Newcastle upon Tyne NE2 4AA or in India to Dr K. C. Malhotra, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035.

Zusammenfassung. Genetische Marker, namlich 11 Blutgruppen- und Erythrozytenenzymsysteme wurden bei 14 endogamen Gruppen Nordwestindiens untersucht. Die genetische Differenzierung zwischen den Stichproben nach F_{ST} ist betrachtlich, worin sich die charakteristische ethnische Vielseitigkeit dieser Region niederschlagt. Die lokale Variation innerhalb eines jeden Staates ist niedriger, was auf einen geographischen Bestandteil an der gesamten Variation hinweist. Diese Variation wird genauer bestimmt durch Berechnungen der genetischen Abstände, die zeigen, daß die Stammesgruppen und die niedrigen Kasten einander naher sind, aber klar getrennt von den Brahmanen der hohen Kasten und von anderen Nichtstammesbevolkerungen mittlerer Kasten. Es gibt eine schwache Möglichkeit einer auflösenden Selektion, aber die Analyse legt eher nahe, daß die Unterschiede in der genetischen Struktur von Nordwestindien eher bedingt sind durch Heiratsstruktur, differenzielle Wanderung und ethnische Abstammung.

Resume. Des marqueurs genetiques consistant en 11 systemes de groupes sanguins et d'enzymes érythrocytaires ont ete etudies dans 14 groupes endogames d'Inde du Nord-Ouest. La différenciation genetique entre les echantillons, telle que calculée par F_{ST}, est appreciable, ce qui réflète la diversite ethnique caracteristique de cette region. La variation locale à l'intérieur de chaque état est plus faible, ce qui indique une composante géographique dans la variation totale. Cette variation est raffinée par le calcul de distances genetiques, qui montrent que les groupes tribaux et de basse caste sont plus proches les uns des autres, mais bien separés des Brahmines de caste superieure et d'autres castes moyennes non tribales. Il y a une legere possibilite de selection disruptive, mais les analyses suggèrent que les différences de structure génétique en Inde du Nord-Ouest sont plus vraisemblablement dues à leur structure des unions, la migration différentielle et l'affiliation ethnique.