PRINT: ISSN 0970-9274 ONLINE: 2456-6608
J. Hum. Ecol., 11(2): 107-123 (2000)
DOI: 10.31901/24566608.2000/11.02.03

Bio-Statistical Approaches in Studying the Origin and the Parental Affinities of the Populations

Gautam Kumar Kshatriya

INTRODUCTION

Many populations seem to have evolved from genetic admixture caused by major migrations during the history and the evolution of mankind. The admixed individuals have genetic characteristics commonly found in both the populations. When the contributing parental populations are genetically very dissimilar, the admixed population/hybrid population generally resembles to each of the parental ones in some respect. Infact, most often their genetic constitution is some where in between those of the contributing parental populations. In describing the population structure of admixed populations, therefore, it is important to identify the mixing elements and the proportion of such mixture (Chakraborty, 1988). Implications of genetic admixture studies lie in understanding the magnitude of genetic affinities with the parental populations and in epidemiological context, in studying the disease-marker associations in search of candidate genes of complex diseases. It would be worth mentioning here, as an example, about the native American Indian groups, who have been shown to have extremely high prevalences of diabetes. This has been most fully documented among the Pima Indian of Arizona (Knowler et al., 1981). Genetically admixed groups having native American ancestry, such as Mexican-Americans, have been shown to have rates of diabetes that are proportional to the degree of Amerindian ancestry, and that the pattern is consistent among various Amerindian and admixed groups across widely differing environmental strata leads one to presume a genetic mechanism, although it is difficult to preclude all other environmental mechanism (Schull and Hanis, 1990). Keeping in view the aforesaid considerations, two different studies have been presented in the present paper on the genetic admixture and the study of the origin of a population as a model for futuristic approach in biological anthropological studies in India. Since India is composed of several populations sub-divided into various sub-populations and many of them being admixed, there is ample scope for studying genetic admixture in assessing the degree and magnitude of the contribution of parental population, the origin of populations *vis-a-vis* the implied genetic role in the etiology of complex diseases.

The first study deals with the estimation of genetic admixture among Mexican-American of Starr County, Texas. In this context, if one looks into historical aspects it could be seen that the first European contact with Mexico occurred in 1519 with the arrival of Hernan Cortez on the eastern coast of the Gulf of Mexico (del Hoyo, 1979). In 1583 Don Luis Carvajal founded a new administrative division, 'Nuevo Reino de Leon', in northen Mexico. Geographically this area, from 1583 to 1824, was composed of the states of Texas, New Mexico, Nuevo Leon, Coahuila, Tamaulipus, Chihuahua, San Luis Potosi, Zacatecas, Durango, Nayarit and Sinaloa (Montemayor-Hernandez, 1971). Later in 1832 the State of Texas dissociated itself from Mexico, and became a separate independent State (Republic of Texas); only to be incorporated into the United States in 1845 (Hernandez-Garza, 1973).

When the Spanish (31.7%), Portuguese and Sephardic Jews (68.3%) colonized the north eastern section of 'Nuevo Reino de Leon' (del Hoyo, 1979), there was almost no admixture with the nomadic native populations of the region (estimated at 35,000 inhabitants in Nuevo Leon and 35,000 in Coahuila). The native Indians were forced to leave Nuevo Leon and Coahuila because of the increasing pressure from the colonizers, thereby abandoning the region principally to the Europe-

ans. However, the Tlaxcaltecan Indians from Central Mexico migrated to this region as a result of an agreement with the colonizers. Subsequently, a considerable degree of admixing of the European and Tlaxcaltecan gene pools occurred (Cossio, 1925; Montemayor-Hernandez, 1971; del Hoyo, 1979; Hernandez-Garza, 1973; Crawford et al., 1974).

The Mexican populations (8,740,439) that reside in United States (Mexican-Americans) are distributed principally in the States of Texas (31.5%), New Mexico (2.7%), Arizona (4.5%), Colorado (2.4%), California (41.6%) and Illinois (4.7%) (US Census 1980). Starr County, one of the 254 counties in Texas, is at least 93.8 per cent Mexican-American. Its Mexican population originated principally from the States of Nuevo Leon and Coahuila before and after the independence of the State of Texas (Hernandez-Garza, 1973).

This study assesses the extent of heterogeneity of the genetic structure of Mexican-American as also the percentage contribution of parental populations to the present day Mexican-American.

The second study envisages the genetic affinities of Srilankan populations with special reference to the Origin of the Sinhalese. Mythological and historical sketch of the Srilankan populations indicate that they are heterogeneous and composed of diverse ethnic groups. Ancient chronicles of Srilanka relate the origin of the Sinhalese with the legend of Prince Vijaya who arrived on the north-west coast of island in 543 B.C. somewhere from north-east or north-west India. Further, as Srilanka occupies an important position on seahighways, it had received a constant influx of people frrom various parts of the world (especially from middle-east and Europe) including India (Kshatriya, 1995). Taking into consideration mythological, historical and linguistic records of Srilanka, an attempt is being made to study the degree of gene diversity and genetic admixture among the population groups of Srilanka alongwith the populations of the southern, north-east, north-west India, middle-east and Europe.

Kirk (1976) and Saha (1988) studied genetic composition of the population of Sri Lanka in order to determine the validity of Ceylonese mythology regarding the legendary origin of the Sinhalese people. While Kirk (1976) found the Sinhalese to be genetically more close to Bengalis and also with the Tamils of India to a lesser extent, Saha (1988) failed to recognise any genetic characteristics in the present day Sinhalese population that are distinct from those in the Tamils of Sri Lanka. Later, however, Tay and Saha (1989) undertook a more detailed study on gene differentiation and genetic admixture among the Sinhalese and the Tamils of Sri Lanka and found that the Sinhalese have a stronger genetic contribution from the Bengalis of India than the Tamils of Sri Lanka.

These earlier studies focussing mainly on the Sinhalese and Tamil population appear to have over simplified the historical and mytholocial records of Sri Lanka. Therefore, the degree and the magnitude of the "foreign admixture" might have been compromised. In view of the above the genetic affinities of the Sri Lankan population have been studied after considering a detailed ethno-historical account of Sri Lankan populations.

MATERIAL AND METHODS

As far as first study on Mexican-American is concerned, genetic data were collected as part of a larger investigation of the epidemiology of gallbladder disease, where 1004 randemly selected persons had complete physical examinations (Hanis, 1991a,b). From the 1004 persons examined, 993 blood samples were analysed. The following erthocyte enzymes and plasma proteins were examined: the polymorphic system ABO, Rh, MNSs, Duffy(FY), Kell (K), Kidd(JK), Adenylate kinase(AK), Haptoglobin (HP), Phosphoglucomutase 1 (PGM1), Glutamate pyruvate transaminase (GPT), Glyoxalase I (GLO I), Phosphoglycolate phosphatase (PGP), Esterase D (EST D), Acid phosphatase (ACPI), 6-phosphogluconate de hydrogenase (6-PGD), Adenosine deaminase (ADA), Group specific component (GC), Apolipoproteins E (APOE), and A-IV (APO AIV); and the monomorphic system Haemoglobin (HB) and Kell

antigen (KP). The details of sample typing, sampling protocols, and about the area and project are well documented in Hanis et al. (1983, 1985, 1991a,b).

The allele frequencies for different system other than PGMI were computed by maximum likelihood method (Reed and Schull, 1967). Allele frequencies for PGM1 were calculated by gene counting. The percentage contribution of ancestral population to the hybrid populations (the present day Mexican-American) was calculated by the method of Chakraborty (1985, 1986); each group being considered as the product of the admixture of three parental populations Spanish, Amerindan and West African. To determine whether the proportions of genes received by the sub-populations from their ancestral sources are significantly different from each other; a regrssion analysis of heterozygosity on genetic distance (Harpending and Ward, 1982) was carried out, and the significance of regression equation was assessed by the method Snedecor and Cochran (1967). Finally, examination of non-random association of alleles at different genetic loci by the methods of Brown et al. (1980) and Chakraborty (1981) was intended to show whether any residual effects of admixture remain in the current population and so make it heterogenous.

Gene frequency data on the ancestral populations was obtained from Mourant et al. (1976) and Hanis et al. (1991b).

Second study deals with the genetic affinities of Srilankan population with special reference to the Origin of the Sinhalese. In this context therefore, the available information on genetic data of various most likely contributory gene pools is presented in table 1.

After careful scrutiny, two data sets were selected, and genetic distance analysis was performed - first on the basis of 11 population groups: the Sinhalese, the Tamils of Sri Lanka, the Veddahs, the Tamils of India, the South Indian Muslims (Andhra Pradesh, Kerala), the Gujaratis, the Punjabis, the West Indian Muslims (Gujarat, Bombay), the Bengalis, the populations of Middle east (Saudi Arabia, Kuwait, Jordian and Iran) and Europe (U.K., Holland, Denmark) with 40 alleles controlled by 13 polymorphic loci; and the second on the basis of 8 population groups (the Sinhalese, the Veddahs, the Tamils of India, the South Indian Muslims, the Gujaratis, the Punjabis, the West Indian Muslims and the Bengalis) with 43 alleles controlled by 15 polymorphic loci (Kshatriya, 1995).

Genetic distances among these population were computed by Nei's standard genetic distance (Nei, 1972), and their standard errors (SE) by Nei and Roychoudhury's (1974) method. To determine the significance of genetic distances among the different population the gene frequency data were compared pairwise by the Chi-square statistic (Nei and Roychoudhary, 1974). The distance matrix was then used to construct a phylogenetic tree based on unweighted pair-group method with arithmetic mean (Li Jin, 1988).

The percentage contribution of ancestral populations to the hybrid populations (the Sinhalese, the Tamils of Sri Lanka) was calculated by the method of Chakraborty (1985, 86), each being considered the product of the admixture of three parental population. The parental population for the Sinhalese have been considered as the Bengalis and the Tamils of India, and the Veddahs of Sri Lanka and for the Tamils as the Sinhalese of Sri Lanka, the Bengalis and the Tamils of India.

RESULTS AND DISCUSSION

Allele Frequency (Mexican-American of Starr Country)

The allele frequency estimates for the 21 loci (Table 2) were used for a goodness-of-fit chi-square test to determine whether the phenotype and genotype frequencies in the Mexican-Americans, and their sex and birthplace subgroups, depart from the Hardy-Weinberg proportions (Table 3), omitting the cases where sample sizes were too small (<10 individuals). The phenotype (genotype) frequencies for most of the loci are in reasonable agreement with their respective Hardy-Weinberg expectations.

Only 11 chi-squares values out of 189 are significant (at<0.05). Some of these involves

Table 1: Genetic loci and samples sizes[a] for the study of origin of the Sinhalese

	Sri Lanka[b]					*India*[b]										*Middle*	
Locus	*SIN*	*TAM*	*MUS*	*BUR*	*VED*	*BEN*	*ASSM*	*ORI*	*GUJ*	*PUN*	*TEL*	*TAM*	*KER*	*WIM*	*SIM*	*EAST*	*Europe*
A1A2BO	1225	660	25	80	254	936	0	146	2155	3398	3265	30	205	820	273	2660	795
MN	1327	788	252	80	138	956	0	0	1238	476	426	128	203	754	0	435	0
RH	1484	788	252	80	305	861	0	146	1955	1600	1085	128	200	820	273	2692	795
P	415	128	0	0	118	396	0	146	93	281	211	121	0	40	87	2505	795
LU	36	0	0	0	63	0	0	142	27	0	211	20	0	61	87	1866	795
LE	137	76	0	0	98	0	0	146	27	419	211	23	0	61	87	0	0
Se	0	0	0	0	0	269	0	0	1609	5411	495	200	0	638	186	0	0
KEL	157	0	0	0	67	398	0	146	517	293	21	48	0	65	87	2312	795
FY	157	0	0	0	106	491	0	146	1217	293	211	48	0	494	87	2362	795
DI	72	0	0	0	39	504	0	0	0	168	0	49	0	0	0	464	0
JK	56	0	0	0	0	0	0	146	3	0	211	0	0	466	87	529	0
INA	0	0	0	0	0	0	0	0	267	0	0	0	192	0	0	0	0
HB	1477	765	306	80	275	967	485	0	1151	835	602	1310	270	762	278	1212	3000
KM	149	0	0	0	0	329	225	0	0	229	0	0	0	0	0	0	0
HP	654	241	49	0	124	712	479	137	1238	1471	1431	647	242	1512	468	2601	801
TF	730	201	47	0	124	784	286	142	848	1468	2487	523	362	1732	342	2750	790
ACP	407	89	47	0	54	569	136	145	753	1109	3261	268	270	527	205	990	782
AK	341	58	26	0	74	452	136	124	610	1386	495	246	306	527	125	1562	801
ADA	278	0	0	0	54	0	136	144	195	1230	296	0	0	61	86	446	798
ESD	396	90	54	0	54	196	179	123	375	951	1914	175	206	153	436	519	6658
GLO	41	56	25	0	0	268	0	0	229	421	1754	50	0	339	350	0	757
CP	0	0	0	0	0	465	0	0	264	649	461	0	250	395	125	0	0
ALB	106	105	49	0	0	0	0	0	267	649	2956	0	250	396	224	540	0
PGD	357	62	16	0	45	271	136	144	812	289	294	267	269	109	211	2147	801
PGM1	397	105	49	0	47	348	136	146	608	1362	545	218	393	507	210	1068	801
PGM2	261	105	49	0	0	0	0	0	610	289	296	0	270	506	210	0	795
LDHA	262	105	49	0	0	1272	0	0	815	289	743	1480	270	527	211	782	795
MDH	106	105	49	0	0	108	0	0	1449	149	733	171	270	527	211	150	795
PEPA	0	0	0	0	0	176	0	0	436	149	85	0	269	415	125	0	795
PEPB	0	0	0	0	0	264	0	0	555	149	85	0	269	466	124	0	795
PEPC	0	0	0	0	0	0	0	0	95	0	85	0	267	0	109	0	795
PEPD	0	0	0	0	0	0	0	0	106	0	0	0	270	0	125	0	0
SOD	106	105	49	0	0	269	0	0	610	149	650	0	270	506	211	0	0
G6PD	1333	160	120	0	51	388	185	0	787	729	294	761	1372	287	86	1431	116
ESC5	0	0	0	0	0	95	74	0	0	784	3134	272	0	247	895	0	798
PGI	0	0	0	0	0	0	0	124	49	640	0	0	0	61	0	0	0
PHI	217	105	49	0	54	0	0	0	265	249	211	0	270	0	211	0	795
DIA	0	0	0	0	0	0	0	0	0	149	85	0	0	0	0	0	801
GC	103	102	47	0	0	214	383	0	0	888	0	347	0	249	0	1573	791
PI	0	0	0	0	0	290	0	0	0	0	0	0	0	396	0	357	0
C3	0	0	0	0	0	0	0	102	0	276	1837	0	41	0	0	359	790
PGK	74	0	0	0	34	0	0	0	265	149	0	0	267	0	125	0	795
GPT	111	0	0	0	54	0	0	0	106	0	131	136	0	0	28	359	778
GOT	111	0	0	0	54	0	0	0	264	0	0	0	0	0	0	0	795
IDH	111	0	0	0	54	0	0	0	0	0	0	0	242	0	12	0	795
NP	60	0	0	0	18	0	0	0	0	0	0	0	241	0	12	0	0
GM	159	108	0	0	52	161	0	0	0	0	0	122	46	0	0	0	0
UMPK	74	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0

a.Number of individuals sampled.

b.Abbreviations for populations : SIN, sinhalese; TAM, Tamils; MUS, Muslims; BUR, Burghers; VED, Veddahs; BEN, Bengalis; ASSM, Assamese; ORI, Oriyas; GUJ, Gujaratis; PUN, Punjabis; TEL, Telgu; KER, Keralites; WIM, West Indian Moslims; SIM, South Indias Muslims.

small observed frequencies (<5 individuals) of specific phenotypes (*e.g.* KK phenotype of the Kell blood group in total; 2^+2^+,2^-2^-,2^+2^-, phenotype of PGM1 in Texas males). The overall pattern of phenotype (genotype) distributions at these 21 loci is in accordance with the Hardy-Weinberg expectations.

Genetic Admixture

Table 4 presents the estimated values of admixture based on 17 polymorphic genetic loci, fitting a trihybrid model using the ancestral frequencies shown in the appendix1. There is little difference among the Mexican-Amer-

Table 2: Allele frequencies Among Mexican-American of Starr County by sex and place of birth

System	Birthplace					
	Texas			Mexico		
	Males	Females	Total	Males	Females	Total
ABO						
A1	0.153	0.170	0.164	0.141	0.136	0.138
A2	0.031	0.038	0.036	0.018	0.027	0.024
B	0.070	0.063	0.065	0.093	0.076	0.081
O	0.747	0.730	0.735	0.748	0.761	0.757
n	156	362	518	129	302	431
RH						
DCE	0.006	0.015	0.012	0.011	0.022	0.019
DCe	0.440	0.433	0.435	0.414	0.461	0.447
DcE	0.187	0.168	0.173	0.204	0.167	0.178
Dce	0.077	0.059	0.065	0.050	0.057	0.055
DCE	0.000	0.000	0.000	0.000	0.000	0.000
dCe	0.000	0.000	0.000	0.000	0.000	0.000
dcE	0.000	0.008	0.006	0.000	0.000	0.000
dce	0.291	0.318	0.309	0.321	0.293	0.301
n	156	362	518	128	302	430
MNSs						
MS	0.297	0.294	0.296	0.280	0.296	0.291
Ms	0.383	0.382	0.381	0.349	0.344	0.345
NS	0.072	0.096	0.088	0.068	0.066	0.067
Ns	0.249	0.227	0.235	0.304	0.294	0.297
n	156	360	516	128	301	429
Duffy						
*A	0.463	0.399	0.417	0.369	0.474	0.450
*B	0.438	0.504	0.484	0.538	0.455	0.480
*QO	0.099	0.097	0.099	0.066	0.070	0.070
n	151	353	504	125	294	419
Kell						
*K	0.010	0.014	0.014	0.016	0.003	0.007
*k	0.990	0.986	0.986	0.984	0.997	0.993
n	156	362	518	129	302	431
Kidd						
*A	0.532	0.504	0.513	0.395	0.454	0.436
*B	0.468	0.496	0.487	0.605	0.546	0.564
n	155	359	514	129	3-2	431
AK						
*1	0.962	0.977	0.972	0.977	0.977	0.977
*2	0.038	0.023	0.028	0.023	0.023	0.023
n	156	363	519	129	302	431
ADA						
*1	0.968	0.967	0.967	0.961	0.962	0.941
*2	0.032	0.033	0.033	0.039	0.038	0.059
n	156	364	520	129	302	451
EST D						
*1	0.889	0.851	0.863	0.840	0.892	0.877
*2	0.111	0.149	0.138	0.160	0.108	0.133
n	157	363	520	128	302	430

ican subgroups. The Spanish contribution varies from 55.9% for Texas males to 66.2% for Texas females, that from Amerindians varies from 27.6% in Texas females to 34.2% in Mexico females, and the African contribution varies from 5.9% in the Mexico total to 11.7% in Texas males. The Mexican-Americans of Starr County, Texas appear to be hybrid pop-

Table 2: Contd.....

	Birthplace					
	Texas			Mexico		
	Males	*Females*	*Total*	*Males*	*Females*	*Total*
ACP						
*A	0.229	0.256	0.248	0.295	0.276	0.282
*B	0.752	0.737	0.741	0.671	0.712	0.700
*C	0.019	0.007	0.011	0.035	0.012	0.019
n	157	363	520	129	302	431
GPT						
*1	0.539	0.458	0.483	0.480	0.455	0.463
*2	0.461	0.542	0.517	0.520	0.545	0.537
n	153	348	501	125	290	415
GLO						
*1	0.385	0.385	0.385	0.395	0.396	0.396
*2	0.615	0.615	0.615	0.605	0.604	0.604
n	157	364	521	129	302	431
PGD						
*A	0.974	0.983	0.981	0.992	0.980	0.984
*C	0.026	0.017	0.019	0.008	0.020	0.016
n	155	362	517	128	301	429
PGP						
*1	0.914	0.859	0.875	0.895	0.887	0.890
*2	0.073	0.136	0.117	0.101	0.108	0.106
*3	0.013	0.005	0.008	0.004	0.005	0.005
n	156	365	522	129	302	431
HP						
*1	0.404	0.437	0.427	0.388	0.465	0.442
*2	0.596	0.563	0.573	0.612	0.535	0.558
n	156	365	522	129	299	428
GC						
1S	0.510	0.521	0.517	0.547	0.493	0.509
1F	0.186	0.203	0.198	0.213	0.222	0.219
*2	0.304	0.277	0.285	0.240	0.285	0.272
n	153	365	518	129	300	429
APOE						
*2	0.050	0.033	0.038	0.052	0.032	0.038
*3	0.864	0.867	0.866	0.868	0.847	0.853
*4	0.086	0.101	0.096	0.080	0.121	0.109
n	151	353	504	125	298	423
APOIV						
*1	0.884	0.938	0.918	0.926	0.943	0.939
*2	0.105	0.062	0.076	0.074	0.057	0.061
*3	0.012	0.000	0.006	0.000	0.000	0.000
n	43	121	165	34	123	157
PGMI						
*1+	0.621	0.575	0.589	0.601	0.553	0.567
*1-	0.197	0.216	0.211	0.201	0.229	0.221
*2+	0.115	0.137	0.130	0.132	0.150	0.144
*2-	0.067	0.071	0.070	0.066	0.068	0.067
n	157	365	522	129	301	430

Kell-Kp and HB loci were monomorphic for Kpb and Hb-a allele, respectively.

ulations with Spanish, Amerindian and West African admixture, with a predominantly Spanish contribution followed by Amerindian and a small West African contribution.

Table 3: Chi-square test for estimating Hardy-Weinberg equilibrium for the Mexican-Americans of Starr County, Texas by sex and birthplace

	Texas			*Mexico*			*Total*		
System	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*
ABO									
χ^2	2.99	2.32	2.90	2.99	1.57	4.06	1.83	0.18	0.19
n	156	362	518	129	302	431	298	691	989
RH									
χ^2	2.39	3.94	1.95	6.74	6.07	8.57	2.13	4.65	3.25
n	156	362	518	128	302	430	297	691	988
MNSs									
χ^2	6.57	2.15	2.99	1.58	0.70	0.54	4.35	1.52	2.54
n	156	360	516	128	301	429	297	687	984
Duffy									
χ^2	3.66	1.99	0.00	0.92	3.35	1.25	0.89	0.04	0.11
n	151	353	504	125	294	419	288	674	962
Kell									
χ^2	0.01	10.36†	8.90†	0.03	0.00	0.02	0.05	12.42†	6.44†
n	156	362	518	129	302	431	298	691	989
Kidd									
χ^2	0.00	0.07	0.04	0.09	0.00	0.04	0.21	0.12	0.00
n	155	359	514	129	302	431	297	688	985
AK									
χ^2	0.25	0.21	0.43	0.07	0.17	0.24	0.32	0.46	0.77
n	156	363	519	129	302	431	298	692	990
ADA									
χ^2	0.17	0.99	0.38	0.21	0.47	1.76	0.36	0.05	0.02
n	156	364	520	129	302	451	298	693	991
EST D									
χ^2	0.59	0.16	0.00	0.22	0.09	0.04	0.01	0.46	0.25
n	157	363	520	129	302	431	299	692	991
ACP									
χ^2	3.25	3.57	0.61	1.95	2.03	2.12	3.69	0.41	2.84
n	157	363	520	128	302	430	298	692	990
GPT									
χ^2	1.29	0.39	0.00	0.01	0.93	0.57	0.03	0.03	0.02
n	153	348	501	125	290	415	291	665	956
GLO									
χ^2	0.19	0.72	0.91	1.36	0.00	0.48	1.71	0.36	1.48
n	157	364	521	129	302	431	299	692	991
PGD									
χ^2	0.11	0.10	0.20	0.01	0.12	0.12	0.09	0.23	0.32
n	155	362	517	128	301	429	296	690	986
PGP									
χ^2	0.38	0.66	1.21	0.21	0.15	0.83	0.58	2.16	2.89
n	156	365	522	129	302	431	299	694	993
HP									
χ^2	0.31	0.24	0.02	0.05	4.76†	4.27†	0.00	2.55	2.02
n	156	365	518	129	300	429	295	629	987
GC									
χ^2	0.38	3.35	3.05	0.12	1.20	0.53	0.40	2.88	2.28
n	153	365	518	129	299	428	299	691	990
APOE									
χ^2	1.31	1.53	0.21	0.44	1.76	0.33	1.32	1.66	0.35
n	151	353	504	125	298	423	287	676	963
APO IV									
χ^2	0.74	0.53	6.13	0.21	0.45	0.65	0.88	0.95	7.59†
n	43	121	165	34	123	157	78	250	329
PGM I									
χ^2	14.12†	10.72	18.74†	2.75	10.43	11.48	9.07	12.66†	20.97†
n	157	365	522	129	301	430	299	692	991

d.f.: ABO=2; Rh=10; MNSs=5; ACP,PGP, Gc, APOE, APOIV=3, Duffy, Kell, Kidd, AK, ADA, ESTD, GPT, GLO, PGD, Hp=1; PGM1=6
† χ^2 significant at $p<0.05$.

Although the standard errors of these estimates are provided, no rigorous test of homogeneity is possible because of the unknown sampling distribution of the estimated admixture proportions to determine whether the gene-flow from outside is homogeneous. The procedure of Harpending and Ward (1982) was therefore applied. The genetic distance (r_g) of subpopulation from a hypothetical centroid of all subpopulations is related to the average heterozygosity ($H_i=b(1-r_g)$, with absolute value of b being equal to $\hat{H}$, the average heterozygosity in the pooled population. Using 18 genetic loci (excluding the two monomorphic Hb and Kp loci, and the APOA IV locus for which a large number of indiviuals were not typed), analysis of the regression (Table 5) of heterozygosity on genetic distance shows it to be consistent with linearity. $\hat{H}$ in the pooled populations (34.06%) does not differ significantly from the regression coefficient (33.97). The various subpopulations of the Mexican-American population of Starr County are therefore similar in the proportion of the genes they have received from the ancestral proportions, which is consistent with the similarity of admixture proportions estimated in the previous section.

Non-random Association Among Genetic Loci

It is well known that the mixture of populations with disparate allele frequencies can produce non-random association of alleles at two or more unlinked loci (Li, 1955; Nei and Li, 1973; Chakraborty and Weiss, 1988). Employing the procedure suggested by Brown et al. (1980), from the available genotype data on each individual, we defined a multi-locus genotype for each individual excluding the monomorphic Hb and Kp loci and the APOA IV genotype, for which too few data were available. With respect to the remaining 18 loci, the number of loci was determined for which each of 862 individuals was heterozygous. Comparing the observed distribution with that expected, under the assumption of random association of alleles at different loci using Chakraborty's (1981) algorithm we find (table 6) that the observed distribution is in fair agreement with the observed one (goodness-of-fit χ^2 with 9d.f. = 19.96, $p>0.01$). Since the expected distribution involves the observed data at least partially (locus-specific observed heterozygosigy values), there are some technical difficulties for determing the degree of freedom of the above goodness-of-fit statistic, detailed in Chakarborty (1984). However, Brown et al. (1980) showed that the expectations of mean and variance of the number of heterozygous loci can be written as functions

Table 4: Percentage contribution from Spanish, Amerindian and West African gene pools to the contemporary Mexican-Americans of Starr County by sex and birth place

Mexican -American	*Spanish*	*Amerindian*	*West African*
Males			
Texas	55.95±2.74	32.33±2.36	11.72±1.24
Mexico	64.08±1.42	29.33±1.23	6.95±0.65
Total	58.62±2.47	31.69±2.13	9.68±1.12
Females			
Texas	66.25±0.21	27.57±0.19	6.17±0.10
Mexico	59.30±1.75	34.19±1.51	6.51±0.79
Total	63.77±0.66	30.16±0.57	6.08±0.30
Total			
Texas	63.25±0.94	28.88±0.81	7.88±0.43
Mexico	62.71±1.04	31.37±0.90	5.92±0.47
Total	62.30±1.19	30.55±1.03	7.16±1.00

The computations are done with 17 polymorphic loci (ABO, Rh, MNSs, Duffy, Kell, Kid, AK, ADA, ESTD, ACP, GPT, GLO, PGD, PGP, Hp, Gc and PGM1). APOE locus data are not used for admixture estimation because allele frequencies at this locus are not available for the appropriate ancestral populations.

Total 5: Average heterozygosity (H_i) and genetic distance from acentroid (r_g) among the Mexican-Americans of Starr County based on 18 polymorphic loci

Population	$r_g \pm SE$	$H_i S \pm E$
Males		
Texas	0.0038±0.0010	0.3343±0.0550
Mexico	0.0048±0.0015	0.3386±0.0549
Females		
Texas	0.0012±0.0003	0.3342±0.0553
Mexico	0.0016±0.0004	0.3390±0.0559

Regression Analysis: $H_i = b(1-r_g)$; H_i plotted against $1-r_g$ through the origin has $t = -0.901$; d.f. = 2, $p>0.05$
Regresion coefficient through origin (b) = 0.3397±0.0033
Average heterozygosity in pooled population ($\hat{H}$) = 0.3406±0.0554

of locus-specific heterogzygosities under the assumption of random association of alleles, and the 95% confidence limit of the observed variance of the number of heterozygous loci can also be calculated. In the Mexican-American data the mean number of heterozygous loci was 6.63 and the variance was 3.54. Their expected values (under random association) are 6.84 and 3.33, respectively. The 95% confidence limits of the variance are 3.02-3.64. Clearly, these suggest no evidence of non-random association of allele among the 18 polymorphic loci in this population.

The admixture results are largely consistent with report for other Mexican-American groups in the United States. Reed (1974), using the Rh blood group system, estimated 32.0 ± 5.6% Amerindian ancestry among the Mexican-Americans of California. Gottlieb and Kimberling's (1979) findings, from a small admixture study in Colorado, showed 60.0% Spanish contribution to the population, indicating a somewhat larger Amerindian component than seen in other studies. Population admixture estimates computed for the Mexican-Americans of San Antonio,Texas, based on skin reflectance (Reletl ford et al., 1983), showed the Amerindian contribution to be 46.0%, 27.0% and 18.0%, respectively among the Barrio, Transition and suburban Mexican-American neighbourhoods. Based on gene frequency data on 18 genetic loci Chakraborty et al. (1986) estimated 43.8%, 30.0% and 18.7% Amerindian ancestry, respectively, in the same three social classes.

Table 6: Observed and expected distribution of the number of heterozygous loci in the Mexican-Americans of Starr County Texas

Number of heterozygous loci	*Number of individuals*	
	Observed	*Expected*
0-2	6	5.72
3	27	20.73
4	81	58.15
5	125	115.96
6	180	169.61
7	167	185.10
8	137	151.80
9	86	93.48
10	31	42.89
11-18	22	18.56
Total	862	862.00
Mean	6.630	6.836
Variance	3.540	3.331

Goodness-of-fit χ^2=19.96, d.f.=9, p>0.01
95% Confidence interval for variance (3.024, 3.638)

The result obtained from the trihybrid model in the present investigation indicated around 62.0%, 30.0% and 8.0% contribution from the Spanish, Amerindian and West African gene pools, respectively. Although various studies show some regional as well as social class variation regarding the contribution of Amerindians to the Mexican-American of the United States, there was no remarkable heterogeneity in genetic admixture among subgroups of the Starr County population. Our estimates of admixture are similar to those obtained by Garza-Chapa (1983), Cerda-Flores et al. (1987), Cerda-Flores and Garza-Chapa (1989) and Cerda-Flores et al. (1991) for the Mexicans of Nuevo Leon in northeastern Mexico. The Chisquare statistic for the 21 genetic marker systems to fit the Hardy-Weinberg equilibrium indicates that intermixing is old enough to have eliminated any early non-random association of genes.

In summary, on the basis of the genetic data presented, it is concluded that the Mexican-Americans of Starr County, Texas, classified by sex and birthplace, are not genetically distinguishable. The findings on genetic admixture indicate a predominant influence from the Spanish, and lesser contributions from Amerindians and West Africans. The history of admixture is apparently old enough to have brought the entire Mexica-American gene pool to Hardy-Weingberg equilibrium. The multilocus heterozygosity distribution also supports the inference of genetic homogeneity.

Coming back to the dynamics of genetic relationships between Sri Lankan Populations with special reference to the Origin of the Sinhalese the allele frequencies of 40 alleles in 11 populations and 43 alleles in 8 populations are presented in table 7.

Average heterozygosity and genetic differences along with their standard errors among all the 11 population groups were estimated using Nei's standard genetic distance among

Table 7: Allele frequencies among the selected population Groups

		Sri Lanka			India						Middle East	Europe
Locus		SIN	TAM	VED	BEN	GUJ	PUN	TAM	WIM	SIM		
ABO	*A1	0.1013	0.1190	0.0402	0.1769	0.1693	0.1467	0.0515	0.1634	0.1270	0.1051	0.1919
	*A2	0.0260	0.0359	0.0103	0.0229	0.0163	0.0194	0.0361	0.0214	0.0126	0.0449	0.0602
	*B	0.1815	0.2182	0.2765	0.2531	0.2061	0.2845	0.2943	0.2342	0.1950	0.1427	0.0548
	*O	0.6912	0.6269	0.6730	0.5471	0.6083	0.5494	0.6181	0.5810	0.6654	0.7073	0.6931
n		612	660	254	242	191	485	30	150	186	532	795
M		0.6636	0.6789	0.4484	0.6458	0.5958	0.6335	0.6172	0.6055	0.6390	0.7805	0.5120
N		0.3364	0.3211	0.5516	0.3542	0.4042	0.3665	0.3828	0.3945	0.3610	0.2195	0.4880
n		442	394	69	191	177	238	128	189	87	415	795
CDE		0.0123	0.0	0.0012	0.0051	0.0047	0.0045	0.0	0.0	0.0	0.0052	0.0
CDe		0.6614	0.6694	0.7635	0.7336	0.5808	0.5678	0.6360	0.5677	0.5670	0.4381	0.4338
Cde		0.0165	0.0140	0.0262	0.0297	0.0067	0.0192	0.0260	0.0192	0.0110	0.0274	0.0109
cDE		0.0712	0.0951	0.0488	0.0751	0.0848	0.1250	0.0550	0.0932	0.1156	0.1375	0.1513
cDe		0.0281	0.0286	0.1045	0.0460+	0.0478	0.0303	0.0150	0.0362	0.0668	0.1216	0.0132
cde		0.2105	0.1929	0.0558	0.1105	0.2752	0.2532	0.2680	0.2837	0.2396	0.2702	0.3798
cdE		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0110
n		371	394	153	172	191	119	128	164	186	525	795
P	*1	0.4342	0.3813	0.4592	0.3047	0.4815	0.5168	0.4545	0.4299	0.3935	0.4848	0.5189
	*2	0.5658	0.6187	0.5408	0.6953	0.5185	0.4832	0.5455	0.5701	0.6065	0.5152	0.4811
n		276	256	212	396	98	188	242	80	174	626	795
HBA		0.9935	0.9976	0.8474	0.9766	0.9946	0.9909	0.9950	0.9973	1.0	0.8886	0.9993
HBE		0.0049	0.0024	0.1526	0.0196	0.0	0.0	0.0038	0.0	0.0	0.0	0.0
Other		0.0	0.0	0.0	0.0038	0.0054	0.0056	0.0012	0.0011	0.0	0.0002	0.0007
HBS		0.0016	0.0	0.0	0.0	0.0	0.0035	0.0	0.0016	0.0	0.1112	0.0
n		738	766	184	276	328	556	2620	304	306	606	6000
HP	*1	0.1720	0.1540	0.2416	0.1644	0.2093	0.2021	0.1332	0.2017	0.2444	0.3378	0.4206
	*2	0.8280	0.08460	0.7584	0.8356	0.7907	0.7979	0.8668	0.7983	0.7556	0.6622	0.5794
n		276	240	120	356	300	416	320	368	266	644	1600
TF	*C	0.9955	0.9975	0.9354	0.9976	0.9958	0.9991	0.9988	0.9989	1.0	0.9960	0.9880
	*D	0.0029	0.0025	0.0646	0.0024	0.0	0.0	0.0012	0.0	0.0	0.0040	0.0019
	*B	0.0016	0.0	0.0	0.0	0.0042	0.0009	0.0	0.0011	0.0	0.0	0.0101
n		292	202	124	392	340	420	348	434	306	786	1580
ACP	*A	0.2447	0.2247	0.4167	0.2456	0.3006	0.3696	0.2484	0.2959	0.2059	0.2358	0.3651
	*B	0.7510	0.7753	0.5833	0.7534	0.6979	0.6263	0.7516	0.6945	0.7941	0.7539	0.5876
	*C	0.0043	0.0	0.0	0.0010	0.0015	0.0041	0.0	0.0096	0.0	0.0103	0.0473
n		272	178	108	284	502	444	268	352	238	396	1564
AK	*1	0.9137	0.9569	0.9054	0.9391	0.8982	0.9066	0.9105	0.9434	0.8800	0.9640	0.9663
	*2	0.0863	0.0431	0.0946	0.0609	0.1018	0.0934	0.0895	0.0566	0.1200	0.0360	0.0337
n		228	116	148	302	406	346	246	352	250	520	1602
ESD	*1	0.7172	0.7222	0.5278	0.7830	0.8379	0.7966	0.7343	0.7727	0.7400	0.8041	0.8751
	*2	0.2828	0.2778	0.4722	0.2170	0.1621	0.2034	0.2657	0.2273	0.2600	0.1959	0.1249
n		236	180	108	196	250	272	350	154	700	520	3330
PGD	*A	0.9792	0.9597	0.9444	0.9797	0.9789	0.9809	0.9791	0.9959	0.9840	0.9386	0.9794
	*C	0.0208	0.0403	0.0556	0.0203	0.0211	0.0191	0.0209	0.0041	0.0160	0.0614	0.0206
n		238	124	90	542	540	290	268	110	250	458	1602
PGM1	*1,1	0.7238	0.6827	0.7761	0.7646	0.6851	0.7084	0.6780	0.7034	0.6720	0.7212	0.7846
	*1,2	0.2709	0.3125	0.2239	0.2347	0.3143	0.2901	0.3157	0.2938	0.3280	0.2773	0.2154
	*1,3	0.0	0.0	0.0	0.0	0.0006	0.0009	0.0	0.0	0.0	0.0006	0.0
	*1,6	0.0	0.0	0.0	0.0007	0.0	0.0002	0.0063	0.0	0.0	0.0005	0.0
	*1,7	0.0053	0.0048	0.0	0.0	0.0	0.0004	0.0	0.0028	0.0	0.0004	0.0
n		264	208	94	232	406	340	218	338	250	428	1602
G6PD	*N	0.9638	1.0	1.0	0.9614	0.9851	0.9528	0.9816	0.8790	0.9884	0.9331	1.0
	*D,E,F	0.0362	0.0	0.0	0.0386	0.0149	0.0472	0.0184	0.1210	0.0116	0.0669	0.0
n		666	320	102	156	224	486	1522	192	172	716	232
k		0.0032	-	0.0777	0.1534	0.0040	0.0132	0.0	0.0233	0.0173	-	-
k		0.9968	-	0.9223	0.8466	0.9960	0.9868	1.0	0.9767	0.9826	-	-
n		157	-	67	80	103	147	48	65	87	-	-
FY	*A	0.5931	-	0.6929	0.3483	0.4481	0.4448	0.5436	0.4103	0.6670	-	-
	*B	0.4069	-	0.3071	0.6517	0.5519	0.5552	0.4564	0.5897	0.3330	-	-
n		157	-	106	82	135	147	48	143	87	-	-

*Kshatriya, 1995.

all the pairs of population (Table 8). The average heterozygosity varies between 27.4% (Bengalis) and 32.6% (Veddhas). The genetic distances show no significant differentation as examined pairwise by the chi square statistics.

However, the dendrogram generated from genetic distance matrix (Fig. 1) revealed an absorbing pattern of clustering. It can be seen that the Sinhalese, the Tamils of Sri Lanka and India and the South Indian Muslims form one cluster, while the Gujaratis, the Punjabis and the West Indian Muslims form the other. These two clusters are quite distinct and do not show much affinities with the Bengalis, the Veddahs, the populations of Middle East and Europe. Infact, the Veddahs stand out to be way apart from all the populations.

The results of a more detailed analysis on the basis of 43 alleles in 8 population groups for average heterozygosity and the genetic distance among all the pairs of populations together with their standard errors are presented in table 9. The average heterozygosity varies between 27.9% (Sinhalese) and 32.2% (Veddahs). The genetic distances do not reveal significant differentation as examined pairwise by the chi square statistic.

Nevertheless, dendrogram produced from the genetic distance matrix (Fig. 2) and the clustering obtained, further strengthens earlier observations. It can be seen that the Sinhalese, the Tamils of India and South Indian Muslims form one cluster, while the Gujaratis, the Punjabis and the West Indian Muslims the other, an almost identical clustering observed in the earlier dendrogram. Here too, the Bengalis and the Veddahs are farthest from the Sinhalese.

Both the dendrograms reveal close similarities between the Sinhalese and the Tamils of Sri Lanka; the Sinhalese and the Tamils of India and between the Tamils of Sri Lanka and the Tamils of India.

Genetic Admixture

Table 10 presents the estimated values of admixture for the two hybrid populations *i.e.*

Table 8: Standard genetic distance and average heterozygosity among the population groups of Sri Lanka, India, the Middle East, and Europe

Population	*Sinhalese*	*Sri Lankan Tamils*	*Veddahs*	*Bengalis*	*Gujaratis*	*Punjabis*	*South Indian Tamils*	*West Indian Muslims*	*South Indian Muslims*	*Middle East*	*Europe*
Sinhalese	28.44 ±5.48										
Sri Lankan Tamils	0.14 ±5.45	27.57 ±5.86									
Veddahs	16.80 ±6.84	19.64 ±7.34	32.64 ±6.87								
Bengalis	3.92 ±2.46	2.11 ±2.69	20.37 ±8.21	27.41 ±5.44							
Gujaratis	2.98 ±1.65	4.14 ±2.71	22.76 ±11.40	7.05 ±3.89	29.69 ±6.08						
Punjabis	4.54 ±2.38	5.92 ±2.71	20.92 ±6.94	8.56 ±5.40	0.40 ±1.70	30.89 ±6.13					
South Indians	0.42	0.69	17.44	5.13	1.81	2.60	29.64				
Tamils	±0.71	±0.95	±6.59	±2.91	±1.42	±1.57	±5.92				
West Indian	2.23	3.13	21.45	4.93	0.61	0.93	1.50	30.85			
Muslims	±1.26	±1.53	±8.39	±3.13	±1.24	±2.82	±1.11	±6.06			
South Indian	1.02	1.26	21.81	5.45	2.15	5.40	1.02	2.37	29.59		
Muslims	±0.71	±1.06	±8.17	±2.61	±1.34	±3.31	±1.37	±1.45	±6.01		
Middle East	10.20 ±4.55	11.93 ±4.95	38.63 ±17.05	18.38 ±8.11	10.09 ±4.26	11.45 ±4.24	13.85 ±5.34	10.17 ±4.14	8.02 ±3.45	31.32 ±6.55	
Europe	21.45 ±8.72	25.91 ±9.97	37.50 ±19.40	20.58 ±12.79	11.58 ±5.57	14.42 ±3.94	25.53 ±10.35	14.87 ±6.12	17.16 ±5.84	16.89 ±8.47	29.35 ±6.68

Values on the diagonal are the average heterozygosities (%) (*e.g.*, Sinhalese = 28.44); below the diagonal are standard genetic distances in 10^{-1} codon differences per locus.

All computations are based on 13 polymorphic loci: ABO, MN, RH, P, HB, HP, TF, ACP, AK, ESD, PGD, PGM and G6PD.

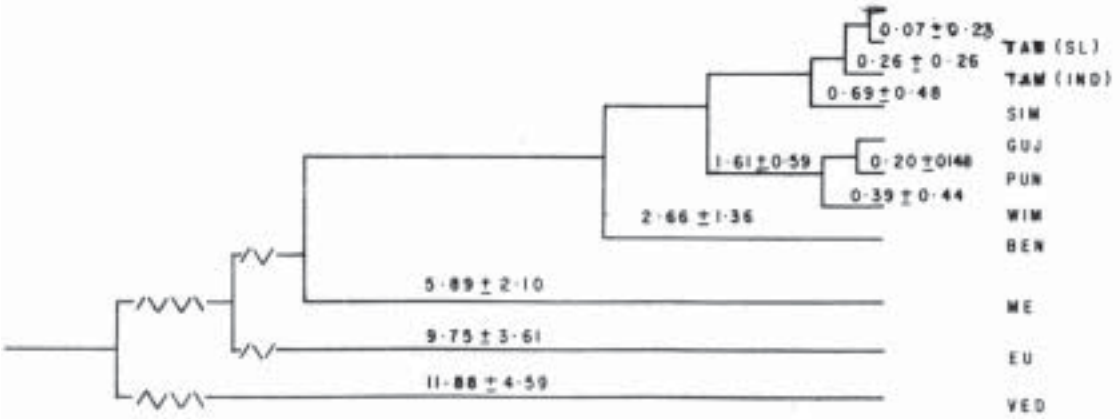

Fig. 1. Genetic relationships of Sri Lankan populations (Sihhalese, Tamils and Veddahs with Indian populations (Tamils, South Indian Muslims, Gujaratis, Punjabs, West Indian Muslims and Bangalis) and populations from the Middle East and Europe based on the genetic distance matrix (40 alleles controlled by 13 loci)

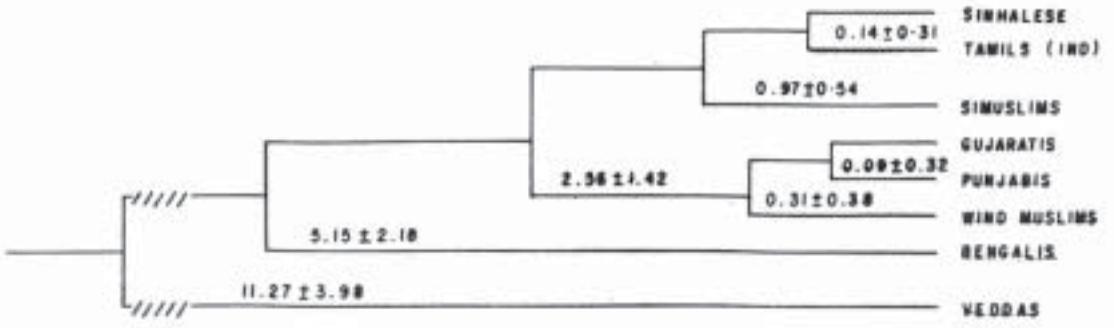

Fig. 2. Genetic relationships of the Sihhalese with the Tamils of India, South Indian Muslims, Gujaratis, Punjabis, West Indian Muslims, Bengalis and the Veddahs based on the genetic distance matrix (43 alleles controlled by 15 loci)

the Sinhalese and the Tamils based on 13 polymorphic loci, fitting a trihybrid model using the ancestral frequencies shown in appendix 2.

The Bengalis, the Tamils and the Veddahs are considered parental populations for the Sinhalese. The Bengalis contribution is found to be 25.41%, the Tamil (India) contribution to be 69.86% and that of the Veddahs only 4.73%. Thus, the Sinhalese have predominantly Tamil (India) contribution followed by the Bengalis and the Veddahs. It may be highlighted here that the fusion of the Veddahs and the Sinhalese has been recorded in the ancient chronicles of Sri Lanka (Dipavamsa and Mahavamsa) as early as 543 B.C., but, the Veddahs were subsequently pushed to inhospitable dry zone for a long period of time under the pressure of early colonizers.

While studying the Tamils of Sri Lanka; the Sinhalese, the Bengalis and the Tamils of India are considered as ancestral populations. The contribution of the Sinhalese to the Tamils of Sri Lanka is found to be 55.20%. Similarly, the Bengali contribution is found to be 28.17% and that of the Tamils (India) 16.63%. The results indicate a predominant influence of the Sinhalese (who already have a very high contribution from the Tamils of India) and the Bengalis to a lesser extent.

In conclusion, it may be said that the original inhabitants of Sri Lanka are the Veddahs who have had little admixture with the Sinhalese and possibly none with the Tamils. The Veddahs are quite distinct because they were confined to inhospitable dry zones and hardly influenced by their neighboring inhabitants. Further, the Sinhalese and the Tamils of Sri Lanka are genetically an admixed populations. The Sinhalese who first came from north-west India under the leadership of Prince Vijaya in 543 B.C. and occupied the Island, have received and exchanged a substantial amount of their genes from and with the populations of north-east and south India. The Sinhalese and the Tamils have no contribution from the population groups of north-west India. In fact, the contribution made by Prince Vijaya and his small band of 700 companions to the original pool of the Sinhalese must have been washed out by the long standing contribution for over 2000 years from the population groups of north-east and south India.

Table 9: Standard genetic distance and average heterozygosity among the selected population groups of Sri Lanka and India

Population	*Sinhalese*	*Veddahs*	*Bengalis*	*Gujaratis*	*Punjabis*	*Tamils*	*West Indian Muslims*	*South Indian Muslims*
Sinhalese	27.93 ±5.28							
Veddahs	15.43 ±5.75	32.12 ±4.44						
Bengalis	10.58 ±5.80	29.45 ±12.94	28.57 ±4.85					
Gujaratis	4.21 ±2.16	25.43 ±11.02	8.70 ±3.79	29.11 ±5.77				
Punjabis	5.66 ±2.68	23.95 ±9.82	9.80 ±4.94	0.00 ±0.69	30.26 ±5.78			
Tamils	0.05 ±0.75	16.89 ±5.80	9.36 ±4.07	1.74 ±1.23	2.51 ±1.37	28.20 ±5.64		
West Indian Muslims	4.76 ±3.06	26.10 ±10.12	5.82 ±3.06	0.32 ±1.09	0.55 ±0.75	2.32 ±1.39	30.29 ±5.66	
South Indian Muslims	1.01 ±0.64	18.49 ±7.18	15.50 ±9.45	5.94 ±4.28	8.89 ±5.10	2.27 ±1.36	7.84 ±6.01	28.87 ±5.58

Values on the diagonal are the average heterozygosities (%) (*e.g.*, Sinhalese = 27.93); below the diagonal are standard genetic distances in 10^{-3} codon differences per locus.

All computations are based on 15 polymorphic loci: ABO, MN, RH, P, HB, HP, TF, ACP, AK, ESD, PGD, PGM and G6PD, Kell and Duffy.

Table 10: Contribution (%) from the gene pools of Sinhalese, Bengalis, Indian Tamils and Veddahs to the contemporary population groups of Sri Lanka

Population	*Sinhalese*	*Bengalis*	*Tamils*	*Veddahs*
Sinhalese	-	25.41±0.51	69.86±0.61	4.73±0.36
Tamils	55.20±9.47	28.17±5.70	16.63±8.78	-

All computations are based on 13 polymorphic loci: ABO, MN, RH, P, HB, HP, TF, ACP, AK, ESD, PGD, PGM, and G6PD.

It is further observed that although there is a noteworthy contribution from the Bengalis (North-east India) to the present day Sinhalese, the Tamil (South India) contribution is predominant. The contribution from the populations of North-east India got diluted probably because the Indian influence on the Sinhalese after 8th century A.D. became predominantly South Indian (Raghavan, 1964). Thus, the Sinhalese of Sri Lanka are genetically more similar to the Tamils of Sri Lanka and India who were always in close proximity with each other historically, linguistically, geographically as well as culturally.

In nutshell, the findings from these two studies have a number of important implications with respect to the utility of such population in anthropogenetic and epidemiologic contexts. First, the ethno-historical, mythological and other related records in association with the genetic data can be used to demonstrate the origin of a population as also the genetic relationships between populations. Secondly, the demonstration of homogeneity in an admixed population can suggest that this population is suitable for studying disease-marker association in search of candidate genes of complex diseases, such an association cannot possibly arise from population mixture alone (Chakraborty and Weiss, 1988). Thirdly, in spite of polyphyletic origin of the admixed population, multi-locus genotypic distribution can be demonstrated to satisfy the premises of random segregation of unlinked loci. Therefore, the probability of finding a specific multi-locus genotype in such a population can be determined by the product rule of locus-specific genotype probabilities. The 862 Mexican-American individuals on which 18-locus genotypes, data were available constitute 862 different multiple-locus genotypes, *i.e.* no repeat of any multiple-locus genotype was observed in the sample. Based on allele frequencies the most probable multiple-locus genotype in this population would be encountered once in every 885,764 individuals. This shows that the discretized genotypic infomation in such a population is sufficient to determine the identity of the individual even when the population is of admixed origin.

KEY WORDS Mexican-American. Srilankan. Sinhalese. Genetic Admixture. Genetic Loci. Chi-square. Centroid.

ABSTRACT A number of bio-statistical approaches are now available to understand the origin of a population, the genetic relationships between the populations and the relative contribution of ancestral populations to the present day admixed populations. The present paper provides the findings of two different studies; one on the estimation of genetic admixture among the Mexican-American of Starr County, Texas and the other on genetic affinities of Srilankan populations with special reference to origin of the Sinhalese. The results, in aggregate, suggest that the populations under study are homogeneous inspite of their admixed nature and point towards the utility of such populations for genetic and epidemiological studies. The findings also suggest the utility of bio-statistically methods in understanding the origin of a population in association with other related facts.

REFERENCES

Brown, A.H.D., Feldman, M.W. and Nevo, E.: Multilocus structure of natural populations of Hordeumspontaneum. *Genetics*, **96**: 249-261 (1980).

Cerda-Flores, R.M., and Garza-Chapa, R.: Variation in gene frequencies of three generations of humans from Monterrey, Nuevo Leon, Mexico. *Human Biology*, **61**: 249-261 (1987).

Cerda-Flores, R.M., Ramirez-Fernandez, E. and Garza-Chapa, R.: Genetic admixture and distances between populations from Monterrey, Nuevo Leon, Mexico and their putative ancestral populations. *Human Biology*, **59**: 31-39 (1987).

Cerda-Flores, R.M., Kshatriya, G.K., Barton, S.A., Leal-Garza, C.H., Garza-Chapa, R., Schull, W.J. and Chkraborty, R.: Genetic structure of the immigrant populations of San Luis Postosi and Zacatecas to Nuevo Leon in Mexico. *Human Biology*, **63**: 309-327 (1991).

Chakraborty, R.: The distribution of the number of heterozygous loci in natural populations. *Genetics*, **98**: 461-166 (1981).

Chakraborty, R.: Detection of non-random association of alleles from the distribution of the number of heterozygous loci in a sample. *Genetics*, **108**: 719-731 (1984).

Chakraborty, R.: Gene identity in a racial hybrids and estimation of admixture rates, pp. 171-180. In: *Ge-*

netic Micro-differentiation - Human and Other populations. Y.R. Ahuja and J.V. Neel (Eds.). Indian Anthropology Association, New Delhi (1985).

Chakraborty, R.: Gene admixture in human populations: Models and Predictions. *Yearbook of Physical Anthropology*, **29**: 1-23 (1986).

Chakraborty, R.: Genetic admixture: Concepts and methods, pp. 242-275. In: *Statistical Methods in Human Population Genetics*. K.C. Malhotra (Ed.). Indian Statistical Institute, Calcutta (1988).

Chakraborty, R. and Weiss, K.M.: Admixture as a tool for finding linked genes and detecting that differences from allelic association between loci. *Proc. Acad. Sci., USA*, **85**: 9119-9123 (1988).

Chakraborty, R., Ferrell, R.E., Stern, M.P., Haffne, S.M., Hazuda, S.P. and Rosenthal, M.: Relationship of prevalence of non-insulin dependent diabetes mellitus with Amerindian admixture in Mexican-Americans of San Antonio, Texas. *Genetic Epidemiology*, **3**: 435-454 (1986).

Cossio, D.A.: Historiade Nuevu Leon, Evolucion Politica Y Social. S. Cantu leal (Ed.). Monterrey, Nueeo Leon, Mexico (1925).

Crawford, M.H., Leyshon, W.C., Brown, K., Less, F. and Taylor, L.: Human Biology in Mexico. I. A comparison of blood group, serum and red cell enzyme frequencies and genetic distances of Indian populations of Mexico. *Am. J. Phy. Anthrop.*, **41**: 251-268 (1974).

Del Hoyo, H.: Historia del Nuevo Reino de Leon (1577-1723). Editorial Libros de Mexico, Mexico, (1979).

Garza-Chapa, R.: Genetic distances for ABO and Rh (d) blood groups in the State of Nuevo Leon, Mexico. *Social Biology*, **30**: 24-31 (1983).

Gottlieb, K. and Kimberling, W.J.: Admixture estimates for the gene pool of Mexican-Amercian in Colorado. Abs. p. 444, *Forty-eighth Annual Meeting of the American Association of Physical Anthropologists*, San Francisco, California (1979).

Hanis, C.L., Ferrell, R.E., Barton, S.A., Aguilar, L., Garzalbara, A., Tullock, B.R., Garcia, C.A. and Schull, W.J.: Diabetes among Mexico Americans in Starr County. *Amer. J. Epidem.*, **118**: 659-672 (1983).

Hanis, C.L., Ferrell, R.E. and Schull, W.J.: Gallbladder disease epidemiology in Mexican Amercians in Starr County. *Amer. J. Epidem.*, **122**: 820-829 (1985).

Hanis, C.L., Douglas, T.C., and Hewett-Emmett, D.: Apolipoprotein A-IV protein polymorphism: Frequency and effects of lipids, lipoproteins and apolipoproteins among Mexican Amercians in Starr County, Texas. *Human Genetics*, **86**: 323-325 (1991a).

Hanis, C.L., Hewett-Emmett, D., Bertin, T.K., and Schull, W.J.: The origins of U.S. Hispanics: Implications of diabetes. *Diabetes Care*, **14**: 618-627 (1991b).

Harpending, H.C. and Ward, R.H.: Chemical systematics and human evolution, pp. 213-256. In: *Biochemical Aspects of Evolutionary Biology*. M. Nitecki (Ed.). University of Chicago Press, Chicago (1982).

Hernandez-Garza, T.L.: *Breve Historia de Nuevo Leon*. Editorial Trillas, Mexico (1973).

Knowler, W.C., Pettitt, D.J., Savage, P.J. and Bennett, P.H.: Diabetes incidence in Pima Indians: Contribution of obesity and parental diabetes. *Amer. J. Epidem.*, **113**: 144-156 (1990).

Kirk, R.L.: The legend of Prince Vijaya: A study of Sinhalese Origins. *Am. J. Phys. Anthrop.*, **45**: 91-100 (1976).

Kshatriya, G.K.: Genetic affinities of Srilankan populations. *Human Biology*, **67**: 843-866 (1995).

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

Montemayor-Hernandez, A.: *Historia de monterrey*. Asociacion de Editores de Monterrey, Mexico (1971).

Mourant, A.E., Kopec, A.C. and Domaniewska-Sobczek, K.: *The Distribution of Human Blood Groups and Other Polymorhisms*. Oxford University Press, London (1976).

Nei, M.: Genetic distance between populations. *American Naturalist*, **106**: 283-292 (1972).

Nei, M. and Roychoudhary, A.K.: Sampling variances of heterozygosity and genetic distance. *Genetics*, **76**: 379-390 (1974).

Raghavan, M.D.: *India in a Ceylonese History, Society and Culture*. Indian Council of Cultural Relation. Asia Publishing House, New Delhi (1964).

Reed, T.E.: Ethnic classification of Mexican-American. *Science*, **185**: 283 (1974).

Reed, T.E. and Schull, W.J.: A general maximum likelihood estimation programme. *Amer. J. Hum. Genet.*, **20**: 579-580 (1968).

Relethford, J.H., Stern, M.P., Gaskill, S.P. and Hazuda, H.P.: Social class, admixture, and skin colour variation in Mexican-Amercians and Anglo-Americans living in San Antonio, Texas. *Amer. J. Phys. Anthrop.*, **61**: 97-102 (1983).

Saha, N.: Blood genetic markers in Srilankan populations: Reappraisal of the legend of Prince Vijaya. *Am. J. Phys. Anthrop.*, **76**: 217-225 (1988).

Schull, W.J. and Hanis, C.L.: Genetics and public health in the 1990's. *Ann. Rev. Publ. Hlth.*, **11**: 105-125 (1990).

Snedecor, G.W. and Cochran, W.G.: *Statistical Methods*. Iowa State University Press, Iowa (1967).

Tay, J. and Saha, N.: Gene differentiation and admixture among the Sinhalese and the Tamils of Srilanka Unpublished (1989).

US Bureau of the Census: *General Population characteristics: Texas. Census of population, Vol.1 Part 45*, PC80-1-B45. US Government Printing Office, Washington (1980).

Author's Address: **Gautam Kumar Kshatriya,** *Reader,* Department of Anthropology, University of Delhi, Delhi 110 007, India.

Appendix 1: Allele frequencies for 17 genetic loci in Mexican-Americans of Starr County and putative ancestral populations†

	Mexican-American			*Spanish*	*Amerindian*	*West African*
Allele	*Males*	*Females*	*Total*			
**A*	0.174	0.184	0.181	0.310	0.063	0.191
**B*	0.077	0.073	0.074	0.067	0.003	0.213
**O*	0.749	0.743	0.745	0.623	0.934	0.596
DCE	0.011	0.020	0.017	0.048	0.022	0.002
DCe	0.429	0.443	0.438	0.418	0.626	0.028
DcE	0.191	0.166	0.174	0.090	0.330	0.084
Dce	0.068	0.064	0.065	0.049	0.000	0.637
dCE	0.000	0.000	0.000	0.002	0.000	0.000
dCe	0.000	0.000	0.000	.011	0.000	0.007
dcE	0.000	0.004	0.003	0.001	0.000	0.000
dce	0.302	0.303	0.303	0.380	0.022	0.242
MS	0.287	0.299	0.296	0.243	0.346	0.088
Ms	0.365	0.360	0.361	0.311	0.444	0.432
NS	0.067	0.080	0.075	0.057	0.080	0.138
Ns	0.282	0.262	0.268	0.389	0.130	0.342
**A*	0.436	0.434	0.435	0.365	0.820	0.002
**B*	0.484	0.482	0.483	0.635	0.180	0.000
**QO*	0.080	0.084	0.083	0.000	0.000	0.998
**K*	0.013	0.010	0.011	0.038	0.000	0.003
**k*	0.987	0.990	0.989	0.962	1.000	0.997
**A*	0.468	0.480	0.476	0.537	0.360	0.433
**B*	0.532	0.520	0.524	0.463	0.640	0.567
*AK*1*	0.968	0.975	0.973	0.978	1.000	0.993
*AK*2*	0.032	0.025	0.027	0.022	0.000	0.007
*ADA*1*	0.966	0.966	0.966	0.951	1.000	1.000
*ADA*2*	0.034	0.034	0.034	0.049	0.000	0.000
*ESD*1*	0.867	0.868	0.868	0.879	0.716	0.938
*ESD*2*	0.133	0.132	0.132	0.121	0.284	0.062
*ACP*A*	0.254	0.265	0.262	0.298	0.203	0.180
*ACP*B*	0.721	0.725	0.724	0.666	0.797	0.820
*ACP*C*	0.025	0.009	0.014	0.036	0.000	0.000
*GPT*1*	0.514	0.456	0.473	0.505	0.442	0.863
*GPT*2*	0.486	0.544	0.527	0.495	0.558	0.137
*GLO*1*	0.390	0.391	0.391	0.423	0.211	0.309
*GLO*2*	0.610	0.609	0.609	0.577	0.789	0.691
*PGD*A*	0.983	0.982	0.982	0.975	0.997	0.937
*PGD*C*	0.017	0.018	0.018	0.025	0.003	0.063
*PGP*1*	0.908	0.868	0.880	0.927	0.690	1.000
*PGP*2*	0.082	0.126	0.113	0.047	0.310	0.000
*PGP*3*	0.010	0.006	0.007	0.026	0.000	0.000
*HP*1*	0.396	0.448	0.432	0.429	0.452	0.654
*HP*2*	0.604	0.552	0.568	0.571	0.548	0.346
*GC*1*	0.725	0.720	0.721	0.641	0.859	0.916
*GC*2*	0.275	0.280	0.279	0.359	0.141	0.084
*PGM*1†*	0.610	0.566	0.579	0.621	0.485	0.787
*PGM*1-*	0.199	0.223	0.215	0.114	0.342	0.039
*PGM2*2†*	0.124	0.144	0.138	0.211	0.020	0.147
*PGM2*2-*	0.067	0.068	0.068	0.054	0.153	0.027

† The allele frequencies for the Mexican-Americans are from the present survey (summing over all individuals) and those for the ancestral populations are compiled from the literature (see Hanis et al., 1991b for exact sources from Mourant et al.'s, 1976 compilation).

Appendix 2: Allele frequency 13 genetic loci among the selected Sri Lankan populations and putative ancestral populations

Allele		*Sri Lankan Tamils*	*Sinhalese*	*Bengalis*	*Indian Tamils*	*Veddahs*
ABO	*A1	0.1190	0.1013	0.1769	0.0515	0.0402
	*A2	0.0359	0.0260	0.0229	0.0361	0.0103
	*B	0.2182	0.1815	0.2531	0.2943	0.2765
	*O	0.6269	0.6912	0.5471	0.6181	0.6730
M		0.6789	0.6636	0.6458	0.6172	0.4484
N		0.3211	0.3364	0.3542	0.3828	0.5516
DCE		0.0	0.0123	0.0051	0.0	0.0012
DCe		0.6694	0.6614	0.7336	0.6360	0.7635
dCe		0.0140	0.0165	0.0297	0.0260	0.0262
DcE		0.0951	0.0712	0.0751	0.0550	0.0488
Dce		0.0286	0.0281	0.0460	0.0150	0.1045
dce		0.1929	0.2105	0.1105	0.2680	0.0558
P	*1	0.3813	0.4342	0.3047	0.4545	0.4592
	*2	0.6187	0.5658	0.6953	0.5455	0.5408
HbA		0.9976	0.9935	0.9766	0.9950	0.8474
HbE		0.0024	0.0049	0.0196	0.0038	0.1526
Other		0.0	0.0	0.0038	0.0012	0.0
HBS		0.0	0.0016	0.0	0.0	0.0
HP	*1	0.1540	0.1720	0.1644	0.1332	0.2416
	*2	0.8460	0.8280	0.8356	0.8668	0.7584
TF	*C	0.9975	0.9955	0.9976	0.9988	0.9354
	*D	0.0025	0.0029	0.0024	0.0012	0.0646
	*B	0.0	0.0016	0.0	0.0	0.0
ACP	*A	0.2247	0.2447	0.2456	0.2484	0.4167
	*B	0.7753	0.7510	0.7534	0.7516	0.5833
	*C	0.0	0.0043	0.0010	0.0	0.0
AK	*1	0.9569	0.9137	0.9391	0.9105	0.9054
	*2	0.0431	0.0863	0.0609	0.0895	0.0946
ESD	*1	0.7222	0.7172	0.7830	0.7343	0.5278
	*2	0.2778	0.2828	0.2170	0.2657	0.4722
PGD	*A	0.9597	0.9792	0.9797	0.9791	0.9444
	*C	0.0403	0.0208	0.0203	0.0209	0.0556
PGM1	*1,1	0.6827	0.7238	0.7646	0.6780	0.7761
	*1,2	0.3125	0.2709	0.2347	0.3157	0.2239
	*1,3	0.0	0.0	0.0	0.0	0.0
	*1,6	0.0	0.0	0.0007	0.0063	0.0
	*1,7	0.0048	0.0053	0.0	0.0	0.0
G6PD	*N	1.0	0.9638	0.9614	0.9816	1.0
DEF		0.0	0.0362	0.0386	0.0184	0.0

* Kshatriya (1995)