Genetic Diversity in Sri Lanka: Some implications in Paternity and Forensic testing.

S.S. Mastana[1], R.Jayasekara[2] and S.S.Papiha[3]
1. Department of Human Sciences, Loughborough University, Loughborough, UK.
2. Human Genetics Unit, Faculty of Medicine, Colombo University, Colombo, Sri Lanka.
3. Department of Human Genetics, Newcastle upon Tyne University, Newcastle upon Tyne, UK.

INTRODUCTION:

In the Indian Subcontinent a vast number of subpopulations exist, originating from different racial stocks and their diversity is dynamically maintained by linguistic, cultural, religious and geographical barriers. The genetic markers that are polymorphic in the populations of this region, show extensive magnitude of variation in allele frequency among subpopulations in different regions (Papiha et al 1982, Mastana and Papiha 1994). In the countries of the Indian subcontinent, pooling of genetic information from different populations for medico-legal problems will lead to spurious results. The necessity to study the genetic composition of various subpopulations for paternity and forensic testing has recently been advocated by Lewontin and Hartl (1991) and others. We have studied five main endogamous groups of Sri Lanka and have seen a considerable variation in their individual allele frequencies for several genetic systems. The aim of this paper is to examine how these allele frequency differences influence parameters like PE (Power of Exclusion), PM (Probability of Match), and DP (Discriminant Probability) in each of the five endogamous subpopulations.

MATERIAL & METHODS:

Populations:

Sinhalese: The Sinhalese are the largest community on the Island and occupy most of the south, west and east of Sri Lanka. Religiously the Sinhalese follow Buddhism.

Tamils: The Tamils (Sri Lankan and Indian) are the largest *Hindu* ethnic minority in Sri Lanka. They are primarily concentrated in the North and North-eastern regions.

Burghers: The Burghers are the descendants of intermarriages between the Sinhalese and the Portuguese and Dutch traders. They practise *Christianity* and are concentrated in the Western side of the Island.

Moors: The Moors (*Muslims*) originated through the intermarriages of local people with Arab traders. Moors are also concentrated on the West Coast.

Malays: The Malays in Sri Lanka are descendants of intermarriages between the Malays (Malaysia) and local women. The Sri Lankan Malays embrace *Islam* and live practically in every part of the island.

Blood samples from Sinhalese (100), Tamils (102), Moors (100) Burghers (103) and Malays (102) were collected and transported by air to the Department of Human Genetics, Newcastle University for processing and analysis. Standard serological methods were used for analysis of blood groups (ABO, MNSs, Rhesus, Kell, FY, P, JK and Colton), while red cell enzymes (AK, ACP, ADA, ESD, GLO, GPT, PGP, PGI, 6PGD, G6PD, LDH, MDH and PGM1 subtypes) and serum proteins (C3, BF, PLG, ORM1, ORM2 and subtypes of HP, GC, TF and PI) were analysed using starch gel and iso-electric focusing electrophoresis. Power of exclusion, Probability of Match and Discriminant Probability was calculated using the methods of Garber and Morris (1983) and Jones (1972) respectively.

RESULTS AND DISCUSSION:

The allele frequency variation in all the five populations investigated was compared by Chi-sqaure analysis and overall heterogeneity was observed for 30% of the systems (Rhesus, KIDD, ACP, ESD, HP, C3, TF, GC systems and PGM1*2+, PLG*1 alleles). Sinhalese differed significantly from other populations in about 12% of the systems (11 out of 92 comparisons), and a higher number of significant differences were found in Tamils (12), Burghers (13), Moors (17) and Malays (14). This diversity among the number of genetic systems indicates that

Table 1. PE, PM and DP estimates in five endogamous populations of Sri Lanka

System	Sinhalese			Tamils			Burghers			Moors			Malays		
	PE	PM	DP	PE	PM	DP	PE	PM	DP	PE	PM	DP	PE	PM	DP
ABO	0.253	0.346	0.654	0.296	0.295	0.705	0.330	0.377	0.623	0.307	0.285	0.715	0.330	0.310	0.690
RHESUS	0.260	0.338	0.662	0.229	0.340	0.660	0.345	0.252	0.748	0.325	0.299	0.701	0.183	0.454	0.546
MNSs	0.455	0.183	0.817	0.410	0.203	0.797	0.411	0.186	0.814	0.453	0.179	0.821	0.413	0.193	0.807
KELL	0.000	1.000	0.000	0.010	1.000	0.000	0.027	0.886	0.114	0.000	1.000	0.000	0.000	1.000	0.000
DUFFY	0.183	0.367	0.633	0.172	0.394	0.606	0.178	0.401	0.599	0.169	0.406	0.594	0.165	0.418	0.582
JK	0.184	0.374	0.626	0.172	0.392	0.608	0.187	0.336	0.664	0.187	0.339	0.661	0.181	0.364	0.636
COB	0.005	0.981	0.019	0.000	1.000	0.000	0.032	0.873	0.127	0.000	1.000	0.000	0.005	0.981	0.019
P1	0.187	0.611	0.389	0.187	0.597	0.403	0.187	0.632	0.368	0.183	0.543	0.457	0.183	0.558	0.442
ACP	0.143	0.507	0.493	0.169	0.426	0.574	0.188	0.428	0.572	0.297	0.261	0.739	0.177	0.410	0.590
ADA	0.106	0.604	0.396	0.094	0.662	0.338	0.077	0.699	0.301	0.096	0.644	0.356	0.063	0.775	0.225
ESD	0.140	0.541	0.459	0.074	0.752	0.248	0.134	0.585	0.415	0.137	0.511	0.489	0.091	0.683	0.317
GLO	0.146	0.480	0.520	0.140	0.496	0.504	0.173	0.391	0.609	0.147	0.479	0.522	0.155	0.457	0.543
GPT	0.182	0.377	0.623	0.191	0.397	0.603	0.180	0.407	0.593	0.180	0.373	0.627	0.187	0.361	0.639
AK	0.061	0.759	0.241	0.084	0.692	0.308	0.067	0.761	0.239	0.057	0.792	0.208	0.085	0.687	0.313
PGI	0.010	0.959	0.041	0.010	0.961	0.039	0.000	1.000	0.000	0.008	0.967	0.033	0.005	0.981	0.019
PGP	0.045	0.820	0.180	0.030	0.885	0.115	0.051	0.805	0.195	0.014	0.938	0.062	0.005	0.943	0.057
PGMS	0.353	0.211	0.789	0.376	0.199	0.801	0.340	0.235	0.765	0.361	0.199	0.801	0.367	0.200	0.800
PGD	0.010	0.961	0.039	0.000	1.000	0.000	0.000	1.000	0.000	0.007	0.974	0.026	0.014	0.943	0.057
BF	0.175	0.459	0.541	0.222	0.402	0.598	0.170	0.429	0.571	0.183	0.397	0.603	0.179	0.389	0.611
C3	0.037	0.862	0.138	0.079	0.712	0.288	0.105	0.599	0.401	0.107	0.592	0.408	0.091	0.697	0.304
GC-S	0.305	0.239	0.761	0.247	0.335	0.665	0.260	0.307	0.693	0.283	0.282	0.718	0.336	0.226	0.774
PLG	0.110	0.629	0.371	0.061	0.757	0.243	0.118	0.606	0.394	0.109	0.591	0.409	0.101	0.624	0.376
PI-S	0.227	0.366	0.634	0.276	0.273	0.727	0.249	0.328	0.672	0.247	0.341	0.659	0.192	0.445	0.555
TF-S	0.160	0.531	0.469	0.188	0.461	0.540	0.202	0.471	0.529	0.189	0.438	0.562	0.158	0.520	0.480
HP-S	0.217	0.396	0.604	0.190	0.452	0.548	0.287	0.286	0.714	0.231	0.364	0.636	0.227	0.351	0.649
ORM1	0.174	0.428	0.572	0.168	0.439	0.561	0.176	0.436	0.564	0.167	0.437	0.563	0.155	0.461	0.539
ORM2	0.035	0.868	0.132	0.039	0.849	0.151	0.014	0.943	0.057	0.010	0.958	0.042	0.039	0.849	0.151
Combined Estimates															
Blood groups	**0.838**	**1.76E-03**	**9.98E-01**	**0.823**	**1.87E-03**	**9.98E-01**	**0.868**	**1.16E-03**	**9.99E-01**	**0.859**	**1.14E-03**	**9.99E-01**	**0.821**	**2.26E-03**	**9.98E-01**
Enzymes	**0.738**	**3.63E-03**	**9.96E-01**	**0.734**	**4.88E-03**	**9.95E-01**	**0.743**	**3.99E-03**	**9.96E-01**	**0.776**	**2.13E-03**	**9.98E-01**	**0.727**	**4.30E-03**	**9.96E-01**
Serum Proteins	**0.801**	**1.69E-03**	**9.98E-01**	**0.807**	**1.54E-03**	**9.98E-01**	**0.832**	**8.69E-04**	**9.99E-01**	**0.820**	**8.91E-04**	**9.99E-01**	**0.810**	**1.21E-03**	**9.99E-01**
16 systems	**0.984**	**1.11E-07**	**1.00E+00**	**0.983**	**1.32E-07**	**1.00E+00**	**0.988**	**5.64E-08**	**1.00E+00**	**0.989**	**2.64E-08**	**1.00E+00**	**0.983**	**1.27E-07**	**1.00E+00**
All systems	**0.992**	**1.09E-08**	**1.00E+00**	**0.991**	**1.41E-08**	**1.00E+00**	**0.994**	**4.05E-09**	**1.00E+00**	**0.994**	**2.17E-09**	**1.00E+00**	**0.991**	**1.18E-08**	**1.00E+00**

the five populations of Sri Lanka constitute different subpopulations, maintained by cultural religious and ethnic barriers. To understand the effect of these genetic differences will have on paternity and forensic testing , we calculated the Power of exclusion (PE), Probability of Match (PM), Discriminant Probability (DP) for each system (Table 1). There is a considerable spread of PE values in Sri Lankan populations. It ranges from 0% in Kell to 45% in MNSs. The MNSs system is the most informative system followed by PGM subtypes and GC subtypes in all the five populations. Using only the blood groups, one can exclude around 83% of all falsely accused fathers in three populations (Sinhalese, Tamils and Moors), while this exclusion power of the blood groups is slightly better in Burghers (86%) and Moors (87%). Using the battery of red cell enzymes, cumulative power of exclusion (CPE) range between 73-78%, the highest found in the Moors. For serum proteins, CPE shows a narrow range 80-83%. Overall using these conventional systems more than 99% of wrongly alleged fathers could be excluded from each of these populations. Sixteen systems have been found to be very informative in all the five populations (ABO, Rhesus, MNSs, FY, JK, ACP, ESD, GLO, GPT, BF, ORM1 and subtypes of PGM, GC, TF, PI and HP) excluding 99% of alleged fathers in Burghers and Moors and 98% in Sinhalese, Tamils and Malays.

A further analysis was attempted to calculate the probability of match (PM) and discriminant probability (DP) of these twenty seven conventional systems. A trend similar to that observed in PE values was observed in PM and DP estimates. In all populations MNSs shows lowest PM values, and therefore has the highest probability of being able to discriminate between blood samples. Other systems with low PM and high DP are PGM and GC subtypes. For blood groups, there is a very similar range of PM estimates in all the five populations with the lowest in the Moors (1.14×10^{-3}) and highest in the Malays (2.26×10^{-3}). A similar trend was observed for red cell enzymes and serum proteins. Moors and Burghers showed the lowest PM values for both red cell enzymes and serum proteins. The combined probability that two blood samples match on all twenty seven genetic markers varies from 1.09×10^{-8} (Sinhalese) to 2.64×10^{-9} (Moors). Using the battery of sixteen systems, PM values varied from from 1.11×10^{-7} (Sinhalese) to 2.64×10^{-8} (Moors).

In conclusion, it seems that in spite of the wide variation in allele frequencies among the five populations, the CPE values for each group are very similar, thus the substructure of the population may have very little effect in paternity testing. Due to considerable variation in each genetic system in different populations, appropriate genetic database for traditional markers is still required. The cumulative probability of match is quite different in the investigated groups therefore, for forensic testing individual subpopulation analysis is essential and warranted.

REFERENCES:

Garber RA, Morris JW (1983). General equations for Average power of exclusion. In: Walker RH (ed), Inclusion probabilities in parentage testing. Am. Ass. Blood Banks. Arlington pp 277-280.

Jones DA (1972). Blood samples: Probability of Discrimination. J.Forens.Sci. Soc. 12:355-359.

Lewontin RC, Hartl DL (1991). Population genetics in forensic DNA typing. Science 254:1745-1750.

Mastana SS, Papiha SS (1994) Genetic structure and Microdifferentiation among four endogamous groups of Maharashtra, Western India. Ann.Hum. Biol. (in press)

Papiha SS, Mukherjee BN, Chahal SMS, Malhotra, KC, Roberts DF (1982) Genetic heterogeneity and population structure in north-west India. Ann. Hum. Biol. 9: 235-251.