

Genetic Distance and Genetic Identity between Hindu and Muslim populations of Barak Valley for ABO and Rh genes

Supriyo CHAKRABORTY

Assam University, Department of Biotechnology, Silchar 788 011, Assam, India; supriyoch_2008@rediffmail.com

Abstract

A genetic study was carried out in two endogamous populations namely Hindus and Muslims in the Barak Valley Zone of Assam in India. Nei's genetic distance and genetic identity between two populations were calculated on the basis of estimated allele frequencies of ABO and Rh blood group genes. The genetic distance between Hindus and Muslims was 0.12% for ABO gene and 0.10% for Rh gene. The genetic identity between two populations was estimated as 99.88% for ABO gene and 99.90% for Rh gene suggesting very high genetic similarity between these two populations. Observed heterozygosity estimate was higher in Hindus (0.5598 for ABO gene and 0.2822 for Rh gene) than Muslims (0.5346 for ABO gene and 0.2408 for Rh gene) indicating lesser inbreeding in Hindus than Muslims. Fixation index was lower in Hindus (16.02% for ABO gene and 43.56% for Rh gene) than Muslims (19.80% for ABO gene and 51.84% for Rh gene). Panmictic index was higher in Hindus than Muslims for both the genes. Fixation and panmictic indices revealed that during evolutionary process the Hindus maintained more outbreeding feature than the Muslims in the valley. In this study, the concepts of genetic load of a population and genotype fitness were extended to alleles to estimate the magnitude of allele genetic load (GL) and allele fitness for 3 alleles in ABO gene and for 2 alleles in Rh gene in two populations. The genetic load for O, A and B alleles were lower in Hindus than Muslims. Similar results for genetic load were found for the alleles of Rh gene in the comparison of two populations. The fitness estimates of O, A and B alleles for ABO gene and D and d alleles for Rh gene were higher in Hindus than Muslims. A population with low allele genetic load (GL) and high allele fitness (AF) might have greater survival advantage in nature in the absence of heterozygote advantage and higher adaptive value of the allele with increased frequency.

Keywords: ABO and Rh genes, genetic distance, heterozygosity, Hindu, Muslim

Introduction

The spectacular invention of nucleic acid sequencing method in 1977 independently by Maxam and Gilbert (1977) as well as Sanger *et al.* (1977) revolutionized biology. In 1995 the first genome sequence of a living organism *Haemophilus influenzae*, a mild human pathogen, was revealed (Fleischmann and Adams, 1995) by whole genome shot-gun sequencing method. Till date more than 250 genomes have been sequenced including higher organisms like *Homo sapiens*, *Mus musculus*, and *Arabidopsis thaliana* etc. Several other genomes are standing in the queue and their genome sequencing is expected to be completed in coming years.

Sequencing is not just enough to understand the impact of evolution in a species. A genome contains a large number of genes *e.g.* human genome project (HGP) has given a draft estimate of 25,000 to 30,000 genes in human genome. The study of all the genes, each with varying number of alleles, in a genome at a time to elucidate the process of molecular evolution is very difficult even with the latest tools of bioinformatics and genomics. It is therefore essential to study one or few genes at a time using the tools of quantitative genetics, population genetics and biostatistics

for greater understanding of the ever-continuing process of molecular evolution of a genome.

Human blood group is a unique genetic marker. Among the various blood group systems, the ABO and the Rh systems are routinely analyzed in humans for medico-legal purposes. Both ABO and Rh systems are genetically controlled. Hence, the lab identification of the blood group of a person reflects the genetic make-up of the person for the particular blood group gene.

Blood typing of a large number of individuals in a human population for ABO and Rh systems generates data to determine the distribution of ABO and Rh blood groups in the population. These data are then genetically analyzed to estimate the frequencies of different alleles for the blood group systems in the population. The proportions of different alleles of a gene in a Mendelian population are known as gene (allele) frequencies (Singh, 1990). The gene frequency, in turn, determines the genotype frequencies of a particular gene in a population.

ABO blood system in human beings was established by Karl Landsteiner in 1900 on the basis of presence or absence of two antigens (A and B) on RBC but its Mendelian inheritance pattern was established by Bernstein in 1924. In ABO system three alleles A, B and O are involved

in a locus and these show intra-allelic interaction to produce four blood group phenotypes namely A, B, AB and O. The allele A produces A antigen, B produces B antigen but the O allele produces neither. Both A and B alleles are mutant forms and show codominance with each other but both are dominant over the O allele. The O allele has been found to be the most frequent allele in almost all human populations/races and may be considered as the wild type allele (Gupta, 1999).

In routine Rh blood group typing, human beings are classified as Rh +ve and Rh -ve groups depending on the presence or absence of Rh antigen or D-antigen. About 100 Rh antigens have been discovered so far (Saran, 2003). Out of these, the D-antigen (Rho) is clinically the most important antigen in the Rh system as it is highly antigenic. It is also the most potent immunogen followed by c and E antigens. The gene for Rh system is located on chromosome 1 of human genome. Assuming Rh blood group system as monogenically controlled character with two alleles (D, d) showing complete dominance, the Rh +ve phenotype will be exhibited by 2 genotypes *i.e.* dominant homozygote (DD) and dominant heterozygote (Dd). But the Rh -ve phenotype will be shown by the recessive homozygote (dd). In practice, the inheritance of Rh blood group system is highly complex. The role of 3 closely linked biallelic genes C, D and E in its inheritance was postulated by Fisher and Race but Weiner postulated multiple allelic series at single Rh locus to explain its inheritance. Neither of these systems has any significance in practical laboratory work, where Rh typing is routinely done using anti-Rh (D) sera. So individuals are classified as Rh +ve (RhD) and Rh -ve (Rh d) phenotypes depending on the presence or absence of D-antigen.

The present study was taken up in two major religious groups (Hindu and Muslim) of Barak Valley zone of Assam, named after the might river Barak that flows through the valley, located in North Eastern Region of India sharing its international border with Bangladesh. The Barak Valley, comprising three districts namely Cachar, Karimganj and Hailakandi, has a land area of 6992 square kilometers and inhabits a total population over 3 millions. The Hindus and the Muslims have followed endogamy and maintained distinct culture and life style but co-habited this valley for last few centuries. No information is available on the genetic identity, genetic distance and other genetic parameters of these two major religious communities. So the present study was taken up to estimate the allelic frequencies of ABO and Rh blood group genes in these two communities and to calculate the genetic identity and genetic distance between them along with the estimation of genetic load and fitness of ABO and Rh alleles in each population.

Materials and methods

Blood samples were collected by trained lab technicians following standard clinical procedure from a sample 700 individuals each in Hindus and Muslims of Barak Valley across diverse age groups of 10-70 years and across both sexes using disposable plastic syringe. ABO and Rh blood group phenotyping data of each individual was determined by appropriate reagents (Saran, 2003) on glass slides in Sonoline Diagnostics Centre and Barak Blood Bank in Silchar and properly recorded.

Considering 3 alleles (O, A and B) for ABO gene and 2 alleles (D, d) for Rh gene, the gene (allele) frequencies for ABO and Rh genes in Hindus and Muslims were estimated according to Hedrick (2005). The SE (standard error), SD (standard deviation) and the range of allele frequency at 95% CI (confidence interval) were calculated as per Harris *et al.* (2007).

The standard genetic distance (D) between two populations for ABO and Rh gene separately was estimated according to Nei (1972) as:

$$D = -\ln(I)$$

where

$$I = \text{Genetic identity} = -\ln(J_{xy} + \frac{1}{2} \ln J_x + \frac{1}{2} \ln J_y)$$

Genetic identity (I) between two populations for ABO and Rh gene separately was calculated according to Hedrick (2005) as:

$$I = \frac{J_{xy}}{\sqrt{J_x \cdot J_y}}$$

where

$$J_{xy} = \sum_{i=1}^n p_{i,x} p_{i,y}, J_x = \sum_{i=1}^n p_{i,x}^2 \text{ and } J_y = \sum_{i=1}^n p_{i,y}^2$$

and $p_{i,x}$ and $p_{i,y}$ are the frequencies of the *i*th allele in two populations x and y.

Observed heterozygosity (H_o) for ABO and Rh genes was calculated on the basis of estimated allele frequencies as

$$H_o = 1 - \sum_{i=1}^n p_i^2$$

where p_i^2 = genotype frequency of observed heterozygotes.

Equilibrium heterozygosity (H_e) at Hardy-Weinberg equilibrium was calculated assuming the occurrence of all the alleles of a gene in equal frequency ($q = 1/k$ where k is the number of alleles). It was estimated as 0.6666 for ABO gene and as 0.5000 for Rh gene considering 3 and 2 alleles, respectively.

Fixation index (F) for each gene in each population was estimated according to Hedrick (2005) as

$$F = \frac{H_e - H_o}{H_e}$$

and F was expressed in percentage. Panmictic index (P) was calculated as (1 - F) and expressed in percentage.

Genetic load for O allele (G_{Lo}) of ABO gene in each population was calculated as $G_{Lo} = \frac{q_{00} - q_{e0}}{q_{e0}}$

where q_{e0} and q_{00} are the observed and expected allele frequency (HW frequency) of O allele in a population. Similarly, the GL_a and GL_b for A and B alleles of ABO gene and the GL_D and GL_d for D and d alleles of Rh gene were estimated for each population.

The fitness of O allele (A_{Fo}) of ABO gene in each population was calculated as $A_{Fo} = 1 - |G_{Lo}|$ considering G_{Lo} as an absolute quantity *i.e.* a scalar. Similarly, A_{Fa}, A_{Fb} for A and B alleles of ABO gene and A_{F_D} and A_{F_d} for D and d alleles of Rh gene were estimated in each population.

Results and discussion

Barak valley, located in southern part of Assam state in North Eastern Region of India is characterized by undulating topography with wide plain area, low lying water logged tracts and hillocks. The climate of Barak valley is sub-tropical, warm and humid with average rainfall of 318 cm and 146 rainy days per annum (Tab. 1). Nearly 80% of the total population depends on agriculture for livelihood.

ABO Gene Frequency

The frequencies of O, A and B alleles in Hindus were 0.60, 0.19 and 0.21 respectively whereas in Muslims the frequencies were 0.63, 0.18 and 0.19 (Tab. 2). Little difference was observed between Hindus and Muslims with respect to ABO gene frequency. Both the populations showed the highest frequency of O allele among 3 alleles. The estimates of standard error (SE) and standard deviation (SD) of allele frequency along with the range of allele frequency for each allele at 95% confidence interval (cl) are presented. The range was much higher for O allele in Hindus compared to that of Muslim indicating greater variation for O allele in Hindus.

Rh Gene Frequency

The frequencies of RhD and Rhd alleles were 0.83 and 0.17 in Hindus and 0.86 and 0.14 in Muslims (Tab. 3) indicating more prevalence of Rh +ve phenotype in Muslims. The distribution of RhD and Rhd alleles in both the populations was similar. The estimates of SE and SD of allele frequency and the range of allele frequency at 95% cl were similar in two populations.

Genetic Distance (D) and Genetic Identity (I)

The genetic distance between Hindus and Muslims for ABO gene was estimated as 0.12% and the genetic identity was 99.88% (Tab. 4). Similarly, the genetic distance and the genetic identity for Rh gene between two populations were 0.10% and 99.90% respectively. These results suggested very high genetic similarity between these two populations.

Several studies have been carried out on genetic distance measurements across different populations. Roy *et al.* (1990) studied the genetic distance and gene diversity among 10 endogamous groups in Chattisgarh, India using the gene frequency data of three genetic loci. They observed that the gene differentiation among these population groups is only about 2 per cent. Papiha *et al.* (1982) analyzed the genetic differentiation in Indian populations and observed that genetic differentiation in populations of India was low (0.26-1.70%). In Assam, Das (1979) studied the variation among three caste populations namely Brahmin, Kalita and Kaibarta on the basis of ABO blood groups and other anthropometric characters. He observed that the Kaibarta stand apart from the Brahmin and the Kalita, who are similar to each other. Danker-Hopfe *et al.* (1988) examined 13 Assamese populations including two Muslim groups for the distribution of anthropometric, anthroposcopic and dermatoglyphic traits and reported that the Muslims were distinguished between Marias (who seemed to be more closely related to Mongoloid populations) and Sheikhs (whose phenotypic appearance was more like that of Hindu caste groups).

Tab. 1. Climatic and other features of the Barak Valley Zone

Climate	Temp (°C)		Mean Annual Rainfall (cm)	No. of rainy days/year	Bright Sunshine hours	Soil p ^H	Population (M)	Major diseases
	min.	max.						
Subtropical warm and humid (hot summer and cool winter)	12.2	36.2	318	146	3.8 (July) to 8.4 (Dec.)	4.6-5.7 (acidic)	3.21	Diabetes, water borne diseases, cancer, sinusitis and cool and gastroenteritis

Tab. 2. Estimates of allele frequency and descriptive statistics of ABO alleles in Hindu and Muslim populations of Barak Valley

Population	O allele (i)				A allele (I ^A)				B allele (I ^B)			
	p (Freq)	SE	SD	Range at 95% cl	q (Freq)	SE	SD	Range at 95% cl	r (Freq)	SE	SD	Range at 95% cl
Hindu	0.60	0.0380	0.0130	0.52-0.67	0.19	0.0004	0.0104	0.18-0.19	0.21	0.0004	0.0109	0.20-0.21
Muslim	0.63	0.0005	0.0001	0.62-0.63	0.18	0.0004	0.0102	0.17-0.18	0.19	0.0003	0.0104	0.18-0.19

Tab. 3. Estimates of allele frequency and descriptive statistics of Rh alleles in Hindu and Muslim populations of Barak Valley

Population	RhD allele				Rhd allele			
	p	SE	SD	Range at 95% cl	q	SE	SD	Range at 95% cl
Hindu	0.83	0.0004	0.0110	0.82-0.83	0.17	0.0003	0.0100	0.16-0.17
Muslim	0.86	0.0003	0.0092	0.85-0.86	0.14	0.0003	0.0093	0.13-0.14

Tab. 4. Estimates of genetic distance, identity, fixation index and panmictic index between Hindus and Muslims of Barak Valley for ABO and Rh gene

Population	Genetic Distance (D)	D (%)	Genetic Identity (I)	I (%)	Observed H	Equilibrium H	Fixation Index (F)	F (%)	Panmictic Index (P)	P (%)
<i>ABO gene</i>										
Hindu	-	-	-	-	0.5598	0.6666	0.1602	16.02	0.8398	83.98
Muslim	0.0012	0.12	0.9988	99.88	0.5346	0.6666	0.1980	19.80	0.8020	80.20
<i>Rh gene</i>										
Hindu	-	-	-	-	0.2822	0.5000	0.4356	43.56	0.5644	56.44
Muslim	0.0010	0.10	0.9990	99.90	0.2408	0.5000	0.5184	51.84	0.8416	48.16

Roychoudhury (1982) analyzed the genetic distance between Jews and Non-Jews using gene frequency data of nine blood groups and protein loci. They observed that the Yemenite Jews have a high degree of genetic affinity to the Israeli Arabs and the Iranian Jews to the Iranians. Triantaphyllidis *et al.* (1983) studied the genetic distance between the inhabitants of nine Mediterranean countries and the three major human races using the gene frequency data of several genetic markers. Their results suggested that Algerians were closer to Negroids while the other Mediterraneans were closer to Caucasoids.

Sokal (1988) computed genetic and taxonomic distances among 3466 samples of human populations in Europe based on 97 allele frequencies and 10 cranial variables. The findings demonstrated that speakers of different language families in Europe differ genetically and that this difference remains even after geographic differentiation.

Observed Heterozygosity

The observed heterozygosity for ABO gene was 0.5598 in Hindus and 0.5346 in Muslims (Tab. 4). Both the estimates were much lower than the Hardy-Weinberg equilibrium heterozygosity (0.6666) calculated on the basis of occurrence of 3 alleles in equal frequency ($1/k$) where k = number of alleles. Lower values of observed heterozygosity as compared to equilibrium heterozygosity in both the populations indicated inbreeding in Hindus and Muslims during evolution. But inbreeding was more in Muslims than Hindus.

Similarly, for Rh gene the observed heterozygosity was 0.2822 in Hindus and 0.2408 in Muslims. Both estimates were lower than HW equilibrium heterozygosity (0.50). These estimates revealed inbreeding in both the populations but inbreeding was more in Barak Valley Muslims than Hindus.

Aarzo and Mohammad (2005) carried out a study on the gene frequencies of ABO, Rh, PTC taste ability and

two other systems in 6 endogamous groups namely Sheikh, Syed, Pathan, Ansari, Saifi and Hindu Bania. Heterozygosity estimate ranged from 0.048 to 0.617 among Sheikhs, 0.149 to 0.599 among Pathans, 0.105 to 0.585 among An-saris, 0.250 to 0.869 among Syeds, 0.107 to 0.565 among Saifis and 0.100 to 0.492 among Hindu Banias.

Fixation Index (F) and Panmictic Index (P)

Fixation index is a good measure of the extent of inbreeding. For ABO gene, Muslims (19.80%) showed higher fixation index than Hindus (16.02%). Similar trend was also observed for Rh gene. Since humans are naturally outbreeding, the magnitude of inbreeding *i.e.* F should be as small as possible to suppress the expression of harmful recessive genes in homozygous condition.

Panmictic index (P) measures the extent of outbreeding in a population and is the opposite of fixation index. For ABO gene the panmictic index was higher in Hindus (83.98%) than Muslims (80.20%). Similarly, for the Rh gene the panmictic index was higher in Hindus (56.44%) than Muslims (48.16%). Fixation and panmictic indices for ABO and Rh genes revealed that Hindus maintained more outbreeding nature than Muslims during evolutionary period. Outbreeding hides the expression of recessive genes in heterozygous genotypes whereas inbreeding opens up the expression of recessive gene in homozygous condition.

Genetic Load (GL) and Allele Fitness (AF)

Genetic load of a population is defined as the reduction in fitness from the maximum possible in a population (Muller, 1950). In other words, it is the relative difference between the actual average fitness of a population and the average fitness that would exist if the fittest genotype presently in the population were to become ubiquitous. According to Crow (1958), genetic load is denoted as

Tab. 5. Estimates of genetic load and fitness of ABO and Rh alleles in Hindus and Muslims of Barak Valley

Population	Genetic load and fitness of ABO alleles						Genetic load and fitness of Rh alleles			
	GLo	AFo	GLa	AFa	GLb	AFb	GLD	AFD	GLd	AFd
Hindu	0.4333	0.5667	-0.7368	0.2632	-0.3636	0.6364	0.1566	0.8434	-0.7647	0.2353
Muslim	0.4603	0.5397	-0.8333	0.1667	-0.4242	0.5758	0.1860	0.8140	-1.1428	Very low

$$GL = \frac{W_{max} - \bar{W}}{W_{max}}$$

where W_{max} and \bar{W} are the fitness of the fittest genotype and the actual average fitness of the population, respectively. Fitness is the relative ability of an organism/genotype to survive and transmit its genes to the next generation. But Kirkpatrick and Jarne (2000) defined genetic load of a population as a reduction in population fitness from the optimum possible.

The deviation ($W_{max} - \bar{W}$) is not necessarily a burden or load to the population. Rather it is the basis for natural selection operating in the population and extremely necessary for evolution to occur in the population.

In the present study, the concept of genetic load applicable to a population has been extended to the genes (alleles) on the given reasoning. Since the origin of mankind in Africa about 200,000 years ago, the frequency of O allele in ABO system increased and that of A and B alleles decreased from the initial frequency ($q=1/k=0.33$ where k is number of alleles *i.e.* 3) under Hardy-Weinberg equilibrium condition to reach the present levels in the existing world populations. But B allele frequency decreased more than that of A. Several evolutionary forces like genetic drift, migration, mutation etc. might have played a major role in changing the allele frequencies differentially in various human populations. These forces resulted in a genetic load on the allele in the existing population compared to the allele in the initial generations.

The genetic load for O allele in Hindus (0.4333) was positive and lower than that in Muslims (0.4603) (Tab. 5). Positive value indicated an increase in allele frequency in existing population over the equilibrium condition. Similarly, the genetic loads for A and B alleles in Hindus were lower in magnitude than those of Muslims but negative in both the populations suggesting a decrease in allele frequency over the equilibrium condition. In Rh system, the genetic load of RhD and Rhd alleles in Hindus was lower than that in Muslims.

Like allele genetic load, the concept of genotype fitness was extended to allele fitness (AF). Allele fitness and allele genetic load are related to each other as $AF=1-|GL|$ where GL is considered as a scalar quantity. The fitness of O, A and B alleles in Hindus was higher than that of Muslims. For Rh system too, similar results were observed in the comparison of allele fitness between Hindus and Muslims. Generally, a population with lower allele genetic load and higher allele fitness may have greater survival advantage

in nature in the absence of both heterozygote advantage (maintaining all alleles in population) and higher adaptive value of the allele with increased frequency.

Conclusions

The results of the present study indicated that the genetic distance between Hindus and Muslims of Barak Valley in Assam was 0.12% for ABO gene and 0.10% for Rh gene. The genetic identity between these two populations was 99.88% for ABO gene and 99.90% for Rh gene suggesting very high genetic similarity between these two populations. Observed heterozygosity was higher in Hindus (0.5598 for ABO gene and 0.2822 for Rh gene) than Muslims (0.5346 for ABO gene and 0.2408 for Rh gene) indicating lesser inbreeding in Hindus than Muslims. Fixation index was lower in Hindus (16.02% for ABO gene and 43.56% for Rh gene) than Muslims (19.80% for ABO gene and 51.84% for Rh gene). Panmictic index was higher in Hindus than Muslims for both the genes. Fixation and panmictic indices revealed that during evolutionary process the Hindus outbred more than the Muslims. The genetic load for O, A and B alleles were lower in Hindus than Muslims. Similarly, the genetic load for D and d alleles of Rh gene were lower in Hindus than Muslims. The fitness estimates of O, A and B alleles for ABO gene and D and d alleles for Rh gene were higher in Hindus than Muslims. A population with low allele genetic load (GL) and high allele fitness (AF) might have greater survival advantage in nature in the absence of heterozygote advantage and higher adaptive value of the allele with increased frequency.

References

- Aarzo, S. S. and A. Mohammad (2005). Gene diversity in some Muslim populations of North India. *Hum Biol.* 77(3):DOI:10.1353/hub.2005.0046.
- Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Hum. Biol.* 30(1):1-13.
- Danker-Hopfe, H., B. M. Das, H. Walter, P. B. Das and R. Das (1988). Anthropological studies in Assam, India - Differentiation processes among Assamese populations. *Anthropol. Anz.* 46(2):159-184.
- Das, B. M. (1979). Physical variation in three Assamese castes. *Anthropol. Anz.* 37(3):204-210.
- Fleischmann, R. D. and M. D. Adams (1995). Whole genome

- random sequencing and assembly of *Haemophilus influenzae* Rd. Science 406:496-512.
- Gupta, P. K. (1999). Quantitative inheritance. In: Genetics. Rastogi Publications, Meerut, India.
- Harris, M., G. Taylor and J. Taylor (2007). Maths and Stats for the Life and Medical Sciences. Viva Books Pvt Ltd, New Delhi.
- Hedrick, P. W. (2005). Measures of genetic variation. Genetics of Populations (3rd Ed). Jones and Bartlett Publishers, Sudbury, Massachusetts, USA.
- Kirkpatrick, M. and P. Jarne (2000). The effects of a bottleneck on inbreeding depression and the genetic load. Am. Nat. 155:154-167.
- Maxam, A. M. and W. Gilbert (1977). A new method for sequencing DNA. Proc Natl Acad Sci, USA 74:560-564.
- Muller, H. J. (1950). Our load of mutations. Am. J. Hum. Genet. 2:111-176.
- Nei, M. (1972). Genetic distance between populations. Am. Nat. 106:283-292.
- Papiha, S. S., B. N. Mukherjee, S. M. S. Chahal, K. C. Malhotra and D. F. Roberts (1982). Genetic heterogeneity and population structure in north-west India. Ann. Hum. Biol. 9(3):235-251.
- Roy, M., U. Datta, M. Mitra and C. S. Singhrol (1990). Genetic distance and gene diversity among ten endogamous groups in Chattisgarh, Central India. Int. J. Anthropol. 5(2):109-115.
- Roychoudhury, A. K. (1982). Genetic distance between Jews and Non-Jews of four regions. Hum. Hered. 32(4):259-263.
- Sanger, F., S. Nicklen and A. R. Coulson (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- Saran, R. K. (2003). Transfusion Medicine Technical Manual. Directorate of Health Services, Ministry of Health and Family Welfare, Govt. of India, New Delhi.
- Singh, B. D. (1990). Population genetics. In: Fundamentals of Genetics. Kalyani Publishers, New Delhi.
- Sokal, R. R. (1988). Genetic, geographic and linguistic distances in Europe. Proc. Natl. Acad. Sci. USA 85:1722-1726.
- Triantaphyllidis, C. D., A. Kouvatzi and L. Kaplanoglou (1983). The genetic distances between the inhabitants of nine Mediterranean countries and the three major human races. Hum. Hered. 33(2):137-139.