AcademicPres

Available online at www.notulaebiologicae.ro

Print ISSN 2067-3205; Electronic 2067-3264

Not Sci Biol, 2011, 3(2):07-11

Notulae Scientia Biologicae

Linkage Disequilibrium between *ABO* and *Rh* Loci in Barak Valley Populations *vis-à-vis* a Few Exotic Populations

Supriyo CHAKRABORTY, Alison MOTSINGER-REIF

North Carolina State University, Bioinformatics Research Centre, 1 Lampe Drive, Raleigh, NC 27695, USA; supriyoch_2008@rediffmail.com

Abstract

From a study of 28 human populations, linkage disequilibrium (LD) between *ABO* and *Rh* blood group loci in human genome were estimated using two different statistic of population genetics *viz. D* and *Q*. In general, the haplotypes *iRh+* and *iRh-* were, on average, 2 to 3-fold and 1.1 to 1.8-fold more frequent respectively than the corresponding expected haplotype frequency in the populations under study. Expected haplotype frequencies were calculated for each population assuming equal frequencies of alleles at *ABO* and *Rh* loci. The observed frequency of I^ARh+ haplotype was more or less close to the expected frequency (0.165). The haplotypes I^ARh-, I^BRh+ and I^BRh- were less frequent than expected in the populations. The *D* value for haplotype *iRh+* was positive for all the populations suggesting preponderance of this haplotype and hence non-random association between *ABO* and *Rh* loci. *Q* is a single measure of LD and suitable for estimation of LD between loci having multiple alleles. *Q* value ranged from 2.18 (Finland) to 8.07 (Hongkong). High LD, indicated by high *Q*, between *ABO* and *Rh* loci was observed in the populations of Hongkong, BV Muslims, Saudi Arabia, BV Hindus, India and Iceland. But the populations of Finland, Austria, Estonia, Poland and Israel showed low LD revealed by low *Q* between *ABO* and *Rh* loci.

Keywords: *ABO locus*, haplotype, linkage disequilibrium, *Rh* locus

Introduction

The study of linkage disequilibrium (LD) between two or more loci in any species is a candidate of contemporary research in population genetics. LD is the non-random association of alleles at two or more loci, not necessarily on the same chromosome. It is not the same as linkage, which depicts the association of two or more loci on a chromosome with limited recombination between them. Smaller the distance between two loci on a chromosome, greater is the linkage. The phenomenon of linkage actually interferes with Mendel's law of independent segregation.

On the other hand, LD describes a situation in which some combinations of alleles or genetic markers occur more or less frequently in a population than would be expected from a random formation of haplotypes from alleles based on their frequencies. LD is also called gametic disequilibrium. Haplotype is a combination of alleles at multiple linked loci that are transmitted together. Non-random associations between polymorphisms at different loci are measured by the degree of linkage disequilibrium. Numerically, it is the difference between observed and expected (assuming random distributions) allelic frequencies.

The term LD is one of the most unfortunate terms that do not reveal its meaning. The term dates back to Lewontin and Kojima (1960). Detecting LD does not ensure either linkage or lack of disequilibrium. The term LD is somewhat misleading because the loci involved may be unlinked *i.e.* on different chromosomes and still be in LD. The level of LD between two or more loci located on same or different chromosomes is influenced by several factors *viz.* linkage, selection, rate of recombination, rate of mutation, genetic drift, non-random mating, age and sex, and population structure. Linkage disequilibrium can occur for a number of reasons namely random drift in allele frequencies, natural selection for or against a particular combination of alleles, non-random mating, mixtures of subpopulations with different allele frequencies and mutation (Thomas, 2004). It tends to decay over time at a rate that depends on the recombination rate. In fact, it is possible that there might be a population association (LD) between alleles at loci that are not linked - possibly not even on the same chromosome-as a result of admixture or chance.

ABO and *Rh* blood group systems in humans are two important genetic markers and routinely analyzed prior to blood transfusion and medical treatment. *ABO* blood group system is governed by a single gene with three alleles (I^A, I^B and *i*), of which I^A and I^B alleles are codominant but both of them are dominant over the recessive allele '*i*' in intra-allelic interaction in diploid condition. The gene for *ABO* blood group is located on chromosome 9 of human genome (Povey *et al.*, 1976).

Rh blood group types were discovered in 1940 by Karl Landsteiner and Alexander Weiner. This blood group system may be the most complex one genetically of all blood type systems since it involves 45 different antigens on the

Received 27 February 2011; accepted 30 April 2011

surface of red cells that are controlled by 2 closely linked genes on chromosome 1. The *Rh* system was named after rhesus monkeys since these monkeys were initially used in the research to make the antiserum for typing blood samples. If the antiserum agglutinates the red cells of a blood sample, the sample is typed as *Rh*+ and if it doesn't, the blood sample is typed as *Rh*-. Despite actual genetic complexity, the inheritance of this trait can usually be predicted by a simple conceptual model in which there are two alleles (*Rh*+ and *Rh*- and *Rh*+ is completely dominant over *Rh*-). Individuals who are either dominant homozygote (*Rh*+*Rh*+) or dominant heterozygote (*Rh*+*Rh*-) have *Rh* antigen and hence they are typed as *Rh*+ individuals. But recessive homozygotes (*Rh*-*Rh*-) do not have *Rh* antigen and hence typed as *Rh*- individuals.

Barak valley (BV), named after the mighty river Barak flowing through the region, is located in southern part of Assam state in North Eastern region of India and shares to its west the international border with Bangladesh. Geographically BV lies between 24°15' and 25°9' N latitude and between 92°16' and 93°15' E longitude. The zone is characterized by an undulating topography with wide plain area, low lying water logged area and hillocks. The climate of Barak Valley is sub-tropical, warm and humid with an average annual rainfall of 318 cm and 146 rainy days per annum. The mean minimum temperature ranges from 12.2°C in January to 25.4°C in August and the mean maximum temperature ranges from 24.3°C in January to 32°C in August. Two major communities of this region are Hindus (BV Hindus) and Muslims (BV Muslims) and they have coinhabited this region for quite a few centuries.

Study of linkage disequilibrium is of importance in evolutionary biology and human genetics because many factors affect it and are affected by it. LD provides information about past events and it constrains the potential response to both natural and artificial selection (Slatkin, 2008). It has been clear for many years that LD is prevalent across much of the human genome (Conrad *et al.*, 2006; Sved, 2009). The importance of LD to evolutionary biology and human genetics was unrecognized outside population genetics till late 1970s. But interest in LD grew rapidly in the 1980s once the usefulness of LD for gene mapping became evident and large-scale surveys of closely linked loci became feasible. LD now finds important applications in mapping gene, detecting natural selection and estimating allele age. LD throughout the genome reflects the population history, the breeding system and the pattern of geographic subdivision, whereas LD in each genomic region reflects the history of natural selection, gene conversion, mutation and other forces that cause gene-frequency evolution. These factors affect LD in a genomic region and it depends on local recombination rate. No information is available on the linkage disequilibrium in the populations of Barak valley. The present study was, therefore, undertaken to estimate the linkage disequilibrium between *ABO* and *Rh* blood group loci in Hindu and Muslim populations of Barak valley in comparison to a few exotic populations.

Materials and methods

The data on the phenotypic distribution of *ABO* and *Rh* blood group system of exotic populations (Tab. 1) were collected from http://en.wikipedia.org/wiki/ABO-blood-group-system and were used for the estimation of allelic frequencies of *ABO* and *Rh* blood group loci. The allelic frequencies for *ABO* and *Rh* loci in Hindu and Muslim populations of Barak Valley were collected from a study of the *ABO* and *Rh* blood group distribution carried out by Chakraborty (2010).

Since *ABO* system is governed by multiple allelic system and *Rh* by biallelic system, the LD between these two loci in all the populations were estimated as per Hedrick (2000) as:

$$\hat{D}ij = \hat{x}_{ij} - \hat{p}_i\hat{q}_i$$

where $\hat{x}_{ij}$ is the estimate of the frequency of gamete A_iB_j in the population and $\hat{p}_i$ and $\hat{q}_i$ are the estimates of the frequencies of alleles A_i and B_i, respectively. The LD between these two loci for all the populations were also calculated by estimating the Q statistic proposed by Hill (1975) as:

$$Q = n\sum_{i=1}^{k}\sum_{j=1}^{m}\left(\hat{D}^2{}_{ij}/\hat{p}_i\hat{q}_j\right)$$

where k and m are number of alleles at loci A and B, respectively and n is the number of gametes. If all Dij=0, then Q approximates a chi-square distribution with (k-1) (m-1) degrees of freedom.

Linkage disequilibrium (D) varies in size as a function of the frequencies of the constituent alleles. D has a maximum and a minimum value of 0.25 and -0.25, respectively only when all the alleles have equal frequencies *i.e.* 0.5 (in the case of biallelic loci). When the allelic frequencies are not equal, the D value is smaller and falls between 0.25 and -0.25.

Results and discussion

Estimates of the allelic frequency for *ABO* locus in different populations revealed that *i* allele was the most frequent one among the three alleles in all the populations (Tab. 1). The frequency of *i* allele was more than 3-fold and 5-fold the frequency of dominant alleles I^A and I^B, respectively. Highest frequency of *i* allele was found in the populations of Iceland (0.75) followed by Ireland (0.74) and Australia (0.70). The lowest frequency of *i* allele was found in Finland (0.56), Israel (0.58) and Turkey (0.58) amongst all the populations. The frequency of I^A allele ranged from 0.18 (BV Muslims, Iceland, Hongkong) to 0.31 (Finland). On the other hand, the frequency of I^B allele varied from the lowest value 0.06 (Portugal) to the highest value 0.22 (India).

Tab. 1. Estimates of allele frequency of *ABO* and *Rh* loci in human populations

Population	Allele frequency at *ABO* locus				Allele frequency at *Rh* locus		
	i	I^A	I^B	Total	*Rh+*	*Rh-*	Total
Australia	0.70	0.23	0.07	1.00	0.56	0.44	1.00
Austria	0.61	0.27	0.12	1.00	0.56	0.44	1.00
Belgium	0.67	0.25	0.08	1.00	0.61	0.39	1.00
Brazil	0.67	0.26	0.07	1.00	0.56	0.44	1.00
Canada	0.68	0.26	0.06	1.00	0.61	0.39	1.00
Denmark	0.64	0.28	0.08	1.00	0.6	0.4	1.00
Estonia	0.59	0.24	0.17	1.00	0.64	0.36	1.00
Finland	0.56	0.31	0.13	1.00	0.64	0.36	1.00
France	0.65	0.28	0.07	1.00	0.61	0.39	1.00
Germany	0.64	0.28	0.08	1.00	0.61	0.39	1.00
Hongkong	0.63	0.18	0.19	1.00	0.92	0.08	1.00
Iceland	0.75	0.18	0.07	1.00	0.61	0.39	1.00
India	0.62	0.16	0.22	1.00	0.8	0.2	1.00
Ireland	0.74	0.19	0.07	1.00	0.6	0.4	1.00
Israel	0.58	0.27	0.15	1.00	0.68	0.32	1.00
New Zealand	0.68	0.24	0.08	1.00	0.58	0.42	1.00
Norway	0.63	0.31	0.06	1.00	0.61	0.39	1.00
Poland	0.61	0.27	0.12	1.00	0.61	0.39	1.00
Portugal	0.65	0.29	0.06	1.00	0.62	0.38	1.00
Saudi Arabia	0.72	0.16	0.12	1.00	0.73	0.27	1.00
South Africa	0.68	0.23	0.09	1.00	0.61	0.39	1.00
Spain	0.67	0.26	0.07	1.00	0.56	0.44	1.00
Sweden	0.62	0.29	0.09	1.00	0.6	0.4	1.00
Turkey	0.58	0.30	0.12	1.00	0.67	0.33	1.00
UK	0.67	0.26	0.07	1.00	0.6	0.4	1.00
USA	0.66	0.27	0.07	1.00	0.61	0.39	1.00
BV Hindus	0.6	0.19	0.21	1.00	0.83	0.17	1.00
BV Muslims	0.63	0.18	0.19	1.00	0.86	0.14	1.00

For *Rh* locus, the frequency of *Rh+* allele varied from the lowest value 0.56 (Australia, Austria) to the highest value 0.86 (BV Muslims). But the frequency of *Rh-* allele ranged from 0.14 (BV Muslims) to 0.44 (Australia, Austria, Belgium and Spain). In general, *Rh+* allele was more frequent than *Rh-* allele in all the populations.

Haplotype frequency

Considering 3 alleles for *ABO* locus and 2 alleles for *Rh* locus, the frequency of 6 haplotypes (*iRh+, iRh-,* I^ARh+, I^ARh-, I^BRh+ and I^BRh-) for each population were calculated from the estimated allele frequencies (Tab. 2). In our study, the expected frequencies of 6 haplotypes for each population were calculated assuming the occurrence of all alleles in equal frequencies for *ABO* and *Rh* loci. All the populations, therefore, had the same expected frequencies for 6 haplotypes *i.e. iRh+*(0.170*), iRh-*(0.170), I^ARh+(0.165), I^ARh-(0.165), I^BRh+(0.165) and I^BRh-(0.165). The comparison of observed and expected haplotype frequencies for all the populations , in general, revealed that the frequencies of observed *iRh+* and *iRh-*

Tab. 2. Estimates of haplotype frequency for *ABO* and *Rh* loci based on observed allelic frequency in human populations

Population	Observed haplotype frequency						Total
	iRh+	*iRh-*	I^ARh+	I^ARh-	I^BRh+	I^BRh-	
Australia	0.392	0.308	0.129	0.101	0.039	0.031	1.00
Austria	0.342	0.268	0.151	0.119	0.067	0.053	1.00
Belgium	0.409	0.261	0.153	0.098	0.049	0.031	1.00
Brazil	0.375	0.295	0.146	0.114	0.039	0.031	1.00
Canada	0.415	0.265	0.159	0.101	0.037	0.023	1.00
Denmark	0.384	0.256	0.168	0.112	0.048	0.032	1.00
Estonia	0.378	0.212	0.154	0.086	0.109	0.061	1.00
Finland	0.358	0.202	0.198	0.112	0.083	0.047	1.00
France	0.397	0.254	0.171	0.109	0.043	0.027	1.00
Germany	0.390	0.250	0.171	0.109	0.049	0.031	1.00
Hongkong	0.580	0.050	0.166	0.014	0.175	0.015	1.00
Iceland	0.458	0.293	0.110	0.070	0.043	0.027	1.00
India	0.496	0.124	0.128	0.032	0.176	0.044	1.00
Ireland	0.444	0.296	0.114	0.076	0.042	0.028	1.00
Israel	0.394	0.186	0.184	0.086	0.102	0.048	1.00
New Zealand	0.394	0.286	0.139	0.101	0.046	0.034	1.00
Norway	0.384	0.246	0.189	0.121	0.037	0.023	1.00
Poland	0.372	0.238	0.165	0.105	0.073	0.047	1.00
Portugal	0.403	0.247	0.180	0.110	0.037	0.023	1.00
Saudi Arabia	0.526	0.194	0.117	0.043	0.088	0.032	1.00
South Africa	0.415	0.265	0.140	0.090	0.055	0.035	1.00
Spain	0.375	0.295	0.146	0.114	0.039	0.031	1.00
Sweden	0.372	0.248	0.174	0.116	0.054	0.036	1.00
Turkey	0.389	0.191	0.201	0.099	0.080	0.040	1.00
UK	0.402	0.268	0.156	0.104	0.042	0.028	1.00
USA	0.403	0.257	0.165	0.105	0.043	0.027	1.00
BV Hindus	0.498	0.102	0.158	0.032	0.174	0.036	1.00
BV Muslims	0.542	0.088	0.155	0.025	0.163	0.027	1.00

haplotypes were, on average, 2 to 3-fold and 1.1 to 1.8-fold higher than corresponding expected haplotype frequency (*i.e.* 0.170). It suggested that the alleles of *ABO* and *Rh* loci, although located on separate chromosomes, did not show random association in gametes resulting in linkage disequilibrium between two loci during evolution.

The observed frequency of haplotype I^ARh+ was more or less close to expected frequency (0.165) whereas the observed frequency of haplotype I^ARh- was lower than the expected haplotype frequency (0.165). On the other hand, the observed frequency of haplotypes I^BRh+ and I^BRh- were much lower than the expected frequency (0.165).

Linkage disequilibrium

Linkage disequilibrium between *ABO* and *Rh* loci was estimated for all haplotypes using the statistic *D* on the basis of observed and expected haplotype frequency in each population (Tab. 3). The *D* estimate for the haplotype *iRh+* was positive for all the populations and it varied from the lowest value of 0.172 (Austria) to the highest value of 0.410 (Hongkong). *D* normally ranges from -0.25 to 0.25. In this study, seven populations namely Hongkong,

Tab. 3. Estimates of linkage disequilibrium (*D* & *Q*) for pair of alleles between *ABO* and *Rh* loci

Population	*D* between inter-allelic pair						Q
	iRh+	*iRh-*	I^ARh+	I^ARh-	I^BRh+	I^BRh-	
Australia	0.222	0.138	-0.036	-0.064	-0.126	-0.134	3.84
Austria	0.172	0.098	-0.014	-0.046	-0.098	-0.112	2.27
Belgium	0.239	0.091	-0.013	-0.068	-0.116	-0.134	3.62
Brazil	0.205	0.125	-0.019	-0.051	-0.126	-0.134	3.37
Canada	0.245	0.095	-0.006	-0.064	-0.128	-0.142	3.91
Denmark	0.214	0.086	0.003	-0.053	-0.117	-0.133	3.12
Estonia	0.208	0.042	-0.011	-0.079	-0.056	-0.104	2.32
Finland	0.188	0.032	0.033	-0.053	-0.082	-0.118	2.18
France	0.227	0.084	0.006	-0.056	-0.122	-0.138	3.40
Germany	0.220	0.080	0.006	-0.056	-0.116	-0.134	3.19
Hongkong	0.410	-0.120	0.001	-0.151	0.010	-0.150	8.07
Iceland	0.288	0.123	-0.055	-0.095	-0.122	-0.138	5.12
India	0.326	-0.046	-0.037	-0.133	0.011	-0.121	5.06
Ireland	0.274	0.126	-0.051	-0.089	-0.123	-0.137	4.83
Israel	0.224	0.016	0.019	-0.079	-0.063	-0.117	2.67
New Zealand	0.224	0.116	-0.026	-0.064	-0.119	-0.131	3.56
Norway	0.214	0.076	0.024	-0.044	-0.128	-0.142	3.24
Poland	0.202	0.068	0.000	-0.060	-0.092	-0.118	2.55
Portugal	0.233	0.077	0.015	-0.055	-0.128	-0.142	3.57
Saudi Arabia	0.356	0.024	-0.048	-0.122	-0.077	-0.133	5.97
South Africa	0.245	0.095	-0.025	-0.075	-0.110	-0.130	3.72
Spain	0.205	0.125	-0.019	-0.051	-0.126	-0.134	3.37
Sweden	0.202	0.078	0.009	-0.049	-0.111	-0.129	2.80
Turkey	0.219	0.021	0.036	-0.066	-0.085	-0.125	2.74
UK	0.232	0.098	-0.009	-0.061	-0.123	-0.137	3.61
USA	0.233	0.087	0.000	-0.060	-0.122	-0.138	3.54
BV Hindus	0.328	-0.068	-0.007	-0.133	0.009	-0.129	5.21
BV Muslims	0.372	-0.082	-0.010	-0.140	-0.002	-0.138	6.53

Iceland, India, Saudi Arabia, Ireland, BV Hindus and BV Muslims outranged the maximum value of 0.25. It could be due to very high frequency of the *Rh+* allele in these populations. In Barak valley zone, the Muslims (0.542) showed higher frequency of observed haplotype *iRh+* than the Hindus (0.498). For the haplotype *iRh-*, the *D* estimate was positive for almost all the populations except Hongkong, India, BV Hindus and BV Muslims. Negative *D* estimate indicated that expected haplotype frequency was more than observed frequency. For the remaining three haplotypes *i.e.* I^ARh-, I^BRh+ and I^BRh-, the *D* estimate was negative for all the populations suggesting that these haplotypes occurred fewer than expected in all the populations. This further indicated non-random association of alleles between *ABO* and *Rh* loci and hence occurrence of linkage disequilibrium between the two loci.

The statistic *Q* is a single, combined measure of linkage disequilibrium unlike *D* and it takes into account all the *D* values of all haplotypes in a population. In our study, the sum of all *D* values for each population was zero; hence calculated *Q* was compared with chi-square value (5.99) for 2 degrees of freedom at p=0.05 for significance. Only two populations namely Hongkong (8.07) and BV Muslims (6.53) had *Q* estimate higher than the table value of chi-square (5.99) indicating that the observed haplotype frequency was significantly different (p=0.05) from expected haplotype frequency in these populations.

In general, the *Q* estimate ranged from 2.18 (Finland) to 8.07 (Hongkong) (Tab. 3). Higher the value of *Q*, greater is the magnitude of linkage disequilibrium. High linkage disequilibrium indicated by high *Q* between *ABO* and *Rh* loci was observed in the populations of Hongkong, BV Muslims, Saudi Arabia, BV Hindus, India and Iceland. But populations like Finland, Austria, Estonia, Poland, Israel showed low linkage disequilibrium *i.e.* *Q* between *ABO* and *Rh* loci. In Barak valley, the Muslims (6.53) recorded higher linkage disequilibrium than the Hindus (5.21) between *ABO* and *Rh* loci. In the present study, the LD observed between *ABO* and *Rh* blood group genes in different populations could be due to natural selection of *i* and *Rh+* allele resulting in higher observed frequency of the haplotype *iRh+* than expected. Further study could be taken up to understand the cause behind high LD between *ABO* and *Rh* genes in human populations.

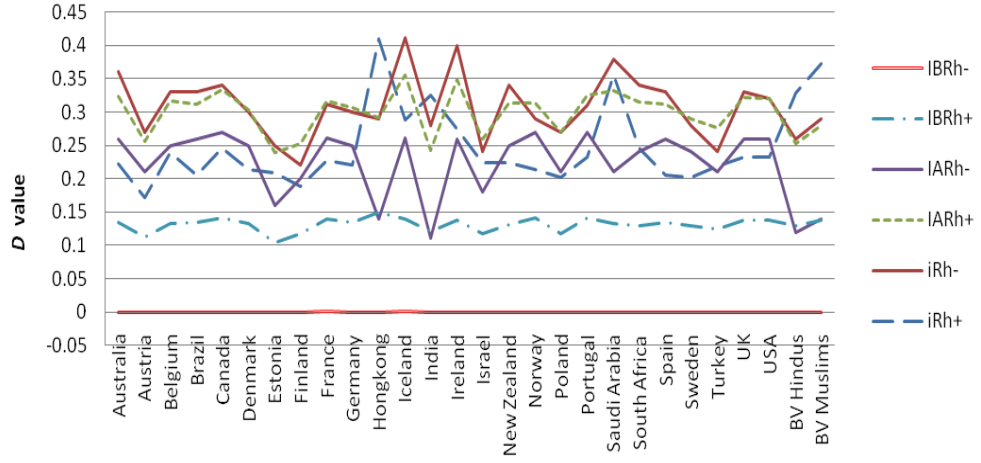

Fig. 1. Linkage disequilibrium (*D*) for each haplotype between *ABO* and *Rh* loci in different populations

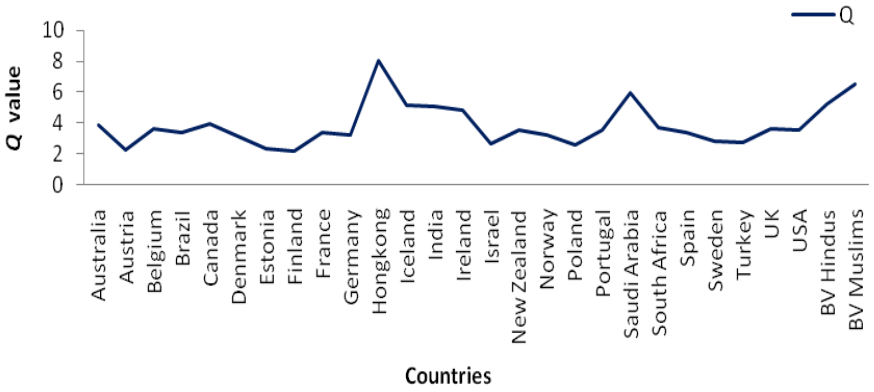

Fig. 2. Linkage disequilibrium (*Q*) for *ABO* and *Rh* loci in different populations

Conclusions

The results of the present study revealed that *ABO* and *Rh* loci, although genetically unlinked being located on chromosome 9 and 1 of human genome, showed substantial linkage disequilibrium in several populations during evolution. The haplotypes *iRh+* and *iRh-* were, on an average, 2 to 3-fold and 1.1 to 1.8-fold more frequent than corresponding expected haplotype frequency calculated on the basis of equal allelic frequency of *ABO* and *Rh* loci. The haplotypes I^ARh-, I^BRh+ and I^BRh- were less frequent than the expected haplotype frequency in the populations under study. Linkage disequilibrium (*D*) for the haplotype *iRh+* was positive for all the populations indicating non-random association of two loci. The estimates of *Q*, a single measure of linkage disequilibrium and suitable for estimation of linkage disequilibrium between loci having multiple alleles, were high in some populations namely Hongkong, BV Muslims, Saudi Arabia, BV Hindus, India and Iceland but low in some other populations like Finland, Austria, Estonia, Finland, Israel and Poland.

Acknowledgements

The first author is thankful to the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi for providing him the DBT Overseas Associateship for North Eastern Region (No.BT/05/NE/2009) to undertake the present research work.

References

Chakraborty S (2010). Genetic analysis on frequency of alleles for *Rh* and *ABO* blood group systems in the Barak valley populations of Assam. Not Sci Biol 2(2):31-34.

Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD, Rosenberg NA, Pritchard JK (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat Genet 38:1251-1260.

Hedrick PW (2000). Genetics of Populations, 2nd Ed. Jones and Bartlett Publishers, Sudbury, Massachusetts, USA, p. 395-444.

Hill WG (1975). Tests for association of gene frequencies at several loci in random mating diploid populations. Biometrics 31:881-888.

Lewontin RC, Kojima K (1960). The evolutionary dynamics of complex polymorphisms. Evolution 14:458-472.

Povey S, Slaughter CA, Wilson DE, Gormley IP, Buckton KE, Perry PA, Bobrow M (1976). Evidence for the assignment of the loci AK_1, AK_3 and $Acon_S$ to chromosome 9 in man. Ann Hum Genet 38:1-5.

Slatkin M (2008). Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477-485.

Sved JA (2009). Linkage disequilibrium and its expectation in human populations. Twin Res Hum Genet 12(1):35-43.

Thomas DC (2004). Statistical Methods in Genetic Epidemiology. Oxford University Press, UK, p. 235-252.