Comparative Study of Nei's *D* with other Genetic Distance Measures between Barak Valley Muslims and other Nations for *ABO* Locus

Supriyo CHAKRABORTY[1], Saumendra Kumar BHATTACHARJEE[2], Gauri Datta SHARMA[3]

[1]*Assam University, Department of Biotechnology, Silchar 788011, Assam, India; supriyoch_2008@rediffmail.com*

[2]*Silchar Medical College and Hospital, Department of Pathology, Silchar 788014, Assam, India*

[3]*Assam University, Department of Life Science and Bioinformatics, Silchar 788011, Assam, India*

Abstract

Quantification of the genetic distance between populations is essential in many genetic research programs. Several formulae were proposed for the estimation of genetic distance between populations using gene frequency data. But the selection of a suitable measure for estimating genetic distance between real-world human populations is a very difficult task despite the widely used measure Nei's *D*. The present study was undertaken to estimate the genetic distance between Barak Valley Muslims (BVM) and other twenty-four nations using seven different measures with *ABO* blood group gene frequency data for comparative analysis and to estimate the correlation coefficients between distance measures and to work out the linear regression equations. Seven genetic distance measures namely Nei's *D*, Nei's *Nm*, *La*, Nei's *Da*, *Dc*, *Re* and Nei's *Ne* were estimated between BVM and other 24 nations *enroute* the journey of mankind from Africa that commenced about 200,000 years ago (www.bradshawfoundation.com). Correlation coefficients between Nei's *D* with other measures were estimated to find out which other genetic distance measures were closely related to Nei's *D*. Nei's *D* showed highly significant (p=0.01) positive correlation with Cavalli-Sforza and Edwards chord distance *Dc* (0.90), Reynolds *Re* (0.90), Nei's *Da* (0.74) and Nei's *Ne* (0.63) but negative correlation with Nei's *Nm* and *La*. Linear regression equations of Nei's *D* with other distance measures were estimated as *Da* =-0.80 + 1.34*D*, *Dc* = 1.91 + 4.44*D*, *Re* =-0.51 + 0.24*D* and *Ne* =-7.60 + 1.30*D*.

Keywords: *ABO* gene, genetic distance, human populations

Introduction

Quantification of the genetic distance between populations is instrumental in many genetic research programs. A large number of formulae have been proposed for this purpose. However, the selection of an appropriate measure for assessing genetic distance between real-world human populations that diverged because of mechanisms that are not fully known can be a challenging task (Libiger *et al.*, 2009).

Nei's standard genetic distance *D* has been the most widely used genetic distance measure between populations. Since several formulae have already been proposed for genetic distance measurement, it is essential to identify which genetic measures show a close relationship with Nei's *D*. Barak Valley zone of Assam in India has a total population of about 3.21 million including Hindus, Muslims and Christians with a land area of 6,992 square kilometers. These populations have maintained distinct culture and life style for centuries despite sharing a few common features. No information is available on the genetic proximity of the Muslims of Barak Valley (BVM) with other nations/populations in the route of migration of humankind that commenced from Africa nearly 200,000 years ago.

ABO blood grouping system was established by Karl Landsteiner in 1900 on the basis of presence or absence of two antigens (A and B) on RBC and its Mendelian inheritance pattern by Bernstein in 1924 (Crow, 1993). In this system, four blood groups namely A, B, AB and O are identified by blood tests. Genetic analysis of the *ABO* blood group system revealed that three alleles namely *A* (I^A), *B* (I^B) and *O* (*i*) determine blood group phenotype. The *A* allele produces A antigen, *B* produces B antigen and 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 in diploid condition.

The nations/populations across the globe can be characterized for the distribution of *ABO* blood groups. These phenotypic data are used to estimate the frequency of different alleles of *ABO* gene using the standard formulae of population and quantitative genetics. The allele frequencies of a gene can be used to estimate the genetic distance between two populations. Human genome project (HGP) has given a draft estimate of 25000 to 30000 genes in human genome. The study of all these genes, each with varying number of alleles, at a time in a genome to elucidate the process of molecular evolution is complicated and almost impossible with the latest available state-of-the-art molecular biology technique (October, 2011). Hence it is imperative to study one or a few genes at a time to understand the evolutionary process in humankind.

The present study was taken up to estimate the genetic distance between Barak Valley Muslims (BVM) and each of other 24 nations for *ABO* blood group gene frequency data using seven different genetic distance measures namely Nei's *D*, Nei's *Nm*, Latter's *La*, Nei's *Da*, Cavalli-Sforza and Edwards *Dc* (*RE*), Reynolds *Re* and Nei's *Ne*. Genetic distance between Barak Valley Muslims and other 24 nations along the historic route of human migration as proposed by Stephen Oppenheimer (at http://www.bradshawfoundation.com) was estimated on the basis of *ABO* gene frequency and to assess the genetic proximity and the evolutionary relationship of the Barak Valley Muslims with other nations.

To identify the distance measure(s) that shows a close relationship with Nei's *D*, correlation analysis was done between the estimates of Nei's *D* and other distance measures. Linear regression equations of different distance measures on Nei's *D* were worked out to determine the value of a particular distance measure with a given value of Nei's *D*.

Materials and methods

In this study, *ABO* blood group distribution data of 25 populations excluding Barak Valley Muslims were obtained from the published literature and websites. The *ABO* blood group distribution data in Barak Valley Muslims were estimated by Chakraborty (2010). The frequencies of *O*, *A* and *B* alleles belonging to *ABO* blood group system for each population were estimated from *ABO* blood group phenotyping data using the formulae suggested by Hedrick (2005) as given below:

$$A = 1 - \sqrt{\frac{N_{22} + N_{23} + N_{33}}{N}}$$

$$B = 1 - \sqrt{\frac{N_{11} + N_{13} + N_{33}}{N}}$$

$$O = \sqrt{\frac{N_{33}}{N}}$$

Where *N*= Total individuals
$N_{11}+N_{13}$= Individuals having "A" blood group
$N_{22}+N_{23}$= Individuals having "B" blood group
N_{33}= Individuals having "O" blood group

Genetic distance
The *ABO* gene frequency data (Tab. 1) were used to estimate the genetic distance between Barak Valley Muslims and each of the remaining 24 populations using seven distance measures as given below.

Let the genetic distance for '*m*' loci with '*v*' alleles per locus be studied in populations 1 and 2 with n_1 and n_2 individuals having n as the average number of individuals. Let $\overline{P}_{lu1}$ and $\overline{P}_{lu2}$ be the frequencies of allele '*u*' at locus '*l*' in population 1 and 2, respectively and let P_{lu1} and P_{lu2} be the number of individuals that carry allele '*u*' at locus '*l*' in populations 1 and 2 respectively, then seven distance measures can be estimated as follows:

Nei's standard genetic distance (*D*) between two populations without bias correction according to Nei (1972) is estimated as:

$$D = -ln\left[\frac{\sum_l \sum_u \overline{P}_{lu1}\overline{P}_{lu2}}{\left(\sum_l \sum_u \overline{P}_{lu1}^2 \sum_l \sum_u \overline{P}_{lu2}^2\right)^{1/2}}\right]$$

Nei's minimum distance (*Nm*) is given by the following equation;

$$Nm = \tfrac{1}{2}\sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu1}\sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu2} - \sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu1}.\overline{P}_{lu2}$$

Latter's distance (*La*) according to Latter (1972) is given by:

$$La = \frac{\tfrac{1}{2}\sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu1}\sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu2} - \sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu1}.\overline{P}_{lu2}}{1 - \sum_{l=1}^{m}\sum_{u=1}^{v}\overline{P}_{lu1}.\overline{P}_{lu2}}$$

Nei's *Da* distance according to Nei *et al.* (1983) is given by:

$$Da = 1 - \frac{1}{m}\sum_{l=1}^{m}\sum_{u=1}^{v}\sqrt{\overline{P}_{lu1}\overline{P}_{lu2}}$$

Cavalli-Sforza and Edwards chord distance (*Dc* or *CE*) according to Cavalli-Sforza and Edwards (1967) is given by:

$$Dc = \tfrac{2}{3.1416m}\sum_{l=1}^{m}\left[2\left\{1 - \sum_{u=1}^{v}\left(P_{lu1}.P_{lu2}\right)^{1/2}\right\}\right]^{1/2}$$

Reynolds genetic distance (*Re*) according to Reynolds *et al.* (1983) is given by:

$$Re = \frac{\sum_l\left\{\tfrac{1}{2}\sum_u\left(\overline{P}_{lu1} - \overline{P}_{lu2}\right)^2 - \frac{1}{2(2\overline{n}-1)}\left[2 - \sum_u \overline{P}_{lu1}^2 + \overline{P}_{lu2}^2\right]\right\}}{\sum_l\left(1 - \sum_u \overline{P}_{lu1}\overline{P}_{lu2}\right)}$$

Nei's geometric distance (*Ne*) based on genotype frequency data (not gene frequency) is given by:

$$Ne = 1 - \tfrac{1}{m}\sum_{l=1}^{m}\sum_{u=1}^{v}\left(P_{lu1}P_{lu2}\right)^{1/2}$$

Correlation and regression analysis
Correlation coefficient between any two distance measures was calculated according to Harris *et al.* (2007). Correlation coefficient was tested by the t-test for significance at *p*=0.01 and 0.05. Linear regression equation of a distance measure (as dependent variable) on Nei's *D* as independent variable was estimated by the method of least squares as per Harris *et al.* (2007).

Results and discussion

Barak Valley Zone, named after the mighty river Barak flowing through the zone, is located in southern part of Assam state in North East India. The valley has inhabited one of the major endogamous religious groups, the Muslims, for several centuries. Barak Valley has a total population of about 3.21 million including Hindus, Muslims and Christians with a land area of 6,992 square kilometers. This region is characterized by undulating topography with wide

14

plain area, low lying water logged tracts and hillocks. The climate of Barak valley is sub-tropical, warm and humid with average annual rainfall of 318 cm and 146 rainy days. Nearly 80% of the total population depends on agriculture for livelihood.

Gene frequency

The frequencies of *O*, *A* and *B* alleles of *ABO* gene of different nations/populations were estimated from the *ABO* blood group distribution data of each population (Tab. 1). In general, the frequency of *O* allele was the highest in all the populations. *B* allele was not reported in Australians.

Genetic distance between populations

The estimates of various genetic distance measures between Barak Valley Muslims (BVM) and each of the twenty-four populations were calculated on the basis of *ABO* gene frequency data (Tab. 2).

Nei's *D* estimate was the lowest (0.0015) between BVM and India (in general) indicating lowest genetic distance between these two populations for *ABO* gene. On the other hand, the highest Nei's *D* value (0.0395) was found between BVM and Australia suggesting greatest genetic distance between these two populations for *ABO*

gene out of 24 combinations. Nei's geometric distance (*Ne*), except all other genetic distance measures, was calculated on the basis of genotypic data estimated from *ABO* gene frequency.

Nei's minimum genetic distance (*Nm*) ranged from the lowest value 0.0074 between BVM and West Indonesia to the highest value 0.0791 between BVM and South China irrespective of sign. Similarly, Latter's distance (*La*) ranged from 0.0146 between BVM and West Indonesia to 0.1366 between BVM and South China.

Nei's *Da* estimate ranged from 0.0009 between BVM and India to 0.1000 between BVM and Australia. Cavalli-Sforza and Edwards chord distance (*Dc*) showed the range from 0.0265 between BVM and India to 0.2847 between BVM and Australia. Reynolds genetic distance (*Re*) ranged from the lowest estimate 0.00002 between BVM and Bulgaria to the highest value 0.0073 between BVM and Australia. Nei's *Ne* estimate ranged from the lowest value 0.0169 between BVM and Sudan to the highest value 0.0808 between BVM and South China. Estimates of genetic distance between BVM and other populations using seven measures are graphically presented (Fig. 2).

Several studies were carried out on genetic distance measurements across different populations. Genetic distance and gene diversity studies by Roy *et al.* (1990)

Tab. 1. Estimates of allele frequencies of *ABO* gene in Barak Valley Muslims (BVM) and other 24 populations/nations

| Sl. No.	Population	Allele Frequency				Reference*
		O	A	B	Total	
1	Kenya	0.69	0.17	0.14	1.00	Anees and Mirza, 2005
2	Sudan	0.81	0.11	0.08	1.00	www.bloodbook
3	Saudi Arabia	0.58	0.21	0.21	1.00	-do-
4	India (Overall)	0.62	0.16	0.22	1.00	-do-
5	Sri Lanka	0.69	0.16	0.15	1.00	-do-
6	West Indonesia	0.69	0.10	0.21	1.00	Breguet *et al.*, 1986
7	Borneo (Malaysia)	0.62	0.22	0.16	1.00	Kamil *et al.,* 2010
8	South China	0.53	0.23	0.24	1.00	www.bloodbook
9	Australia	0.78	0.22	-	1.00	-do-
10	Bulgaria	0.57	0.31	0.12	1.00	-do-
11	Hungary	0.60	0.27	0.13	1.00	-do-
12	Austria	0.60	0.30	0.10	1.00	-do-
13	Pakistan	0.74	0.12	0.14	1.00	-do-
14	Central Asia (Uzbekistan)	0.56	0.25	0.19	1.00	Revavov *et al.*, 1983
15	Eastern Europe(Poland)	0.57	0.28	0.15	1.00	www.bloodbook
16	Siberia	0.57	0.16	0.27	1.00	-do-
17	Russia	0.57	0.25	0.18	1.00	-do-
18	Alaska	.62	0.29	0.09	1.00	-do-
19	USA (Whites)	0.67	0.25	0.08	1.00	-do-
20	Britain	0.69	0.26	0.05	1.00	-do-
21	Norway	0.62	0.32	0.06	1.00	-do-
22	Sweden	0.62	0.31	0.07	1.00	-do-
23	Iceland	0.74	0.19	0.07	1.00	-do-
24	Denmark	0.64	0.27	0.09	1.00	-do-
25	Barak Valley Muslims	0.63	0.18	0.19	1.00	Chakraborty, 2010

*Detailed reference in text

Tab. 2. Estimates of seven different genetic distance measures between BVM and other nations for *ABO* gene

Sl. No.	Combination	Nei's D	Nei's Nm	La	Nei's Da	Cavalli-Sforza Dc	Reynolds Re	Nei's Ne
1	BVM-Kenya	0.0045	0.0081	0.0159	0.0027	0.0464	-0.0038	-0.0549
2	BVM-Sudan	0.0268	-0.0453	-0.0996	0.0216	0.1325	0.0051	-0.0169
3	BVM-Saudi Arab	0.0032	0.0569	0.1021	0.0013	0.0330	-0.0042	-0.0762
4	BVM-India	0.0015	0.0388	0.0720	0.0009	0.0265	-0.0047	-0.0691
5	BVM-Sri Lanka	0.0038	0.0080	0.0157	0.0022	0.0418	-0.0039	-0.0552
6	BVM-West Indonesia	0.0083	0.0074	0.0146	0.0068	0.0741	-0.0029	-0.0505
7	BVM-Borneo	0.0027	0.0394	0.0730	0.0017	0.0368	-0.0045	-0.0689
8	BVM-South China	0.0138	0.0791	0.1366	0.0051	0.0646	-0.0021	-0.0808
9	BVM-Australia	0.0395	-0.0310	-0.0661	0.1000	0.2847	0.0073	0.0239
10	BVM-Bulgaria	0.0276	0.0623	0.1108	0.0135	0.1047	-0.00002	-0.0662
11	BVM-Hungary	0.0136	0.0487	0.0888	0.0076	0.0783	-0.0025	-0.0669
12	BVM-Austria	0.0253	0.0490	0.0893	0.0150	0.1101	-0.0004	-0.0599
13	BVM-Pakistan	0.0113	-0.0144	-0.0297	0.0071	0.0761	-0.0013	-0.0412
14	BVM-Central Asia	0.0093	0.0661	0.1168	0.0039	0.0562	-0.0031	-0.0773
15	BVM-Eastern Europe (Poland)	0.0161	0.0620	0.1103	0.0074	0.0776	-0.0020	-0.0722
16	BVM-Siberia	0.0106	0.0608	0.1084	0.0045	0.0607	-0.0029	-0.0740
17	BVM-Russia	0.0084	0.0617	0.1098	0.0037	0.0547	-0.0033	-0.0758
18	BVM-Alaska (Eskimo)	0.0235	0.0401	0.0742	0.0158	0.1131	-0.0006	-0.0559
19	BVM-USA Whites	0.0175	0.0177	0.0342	0.0149	0.1098	-0.0013	-0.0483
20	BVM-Britain	0.0261	0.0090	0.0177	0.0269	0.1476	0.0009	-0.0349
21	BVM-Norway	0.0379	0.0404	0.0748	0.0282	0.1513	0.0022	-0.0447
22	BVM-Sweden	0.0328	0.0403	0.0746	0.0235	0.1379	0.0012	-0.0489
23	BVM-Iceland	0.0185	-0.0137	-0.0282	0.0170	0.1172	0.0004	-0.0339
24	BVM-Denmark	0.0187	0.0311	0.0586	0.0138	0.1058	-0.0014	-0.0543

among 10 endogamous groups in Chattisgarh, India using the gene frequency data of three genetic loci revealed that the gene differentiation among these population groups is only about 2 per cent.

Genetic differentiation studies in Indian populations by Papiha *et al.* (1982) revealed that genetic differentiation in India populations was low (0.26-1.70%). In Assam, genetic variation studies by Das (1979) among three caste populations namely Brahmin, Kalita and Kaibarta on the basis of ABO blood groups and other anthropometric characters revealed that the Kaibarta stand apart from the Brahmin and the Kalita, who are similar to each other. Genetic study by Danker-Hopfe *et al.* (1988) among 13 Assamese populations including two Muslim groups for the distribution of anthropometric, anthroposcopic and dermatoglyphic traits revealed that the Muslims in As-
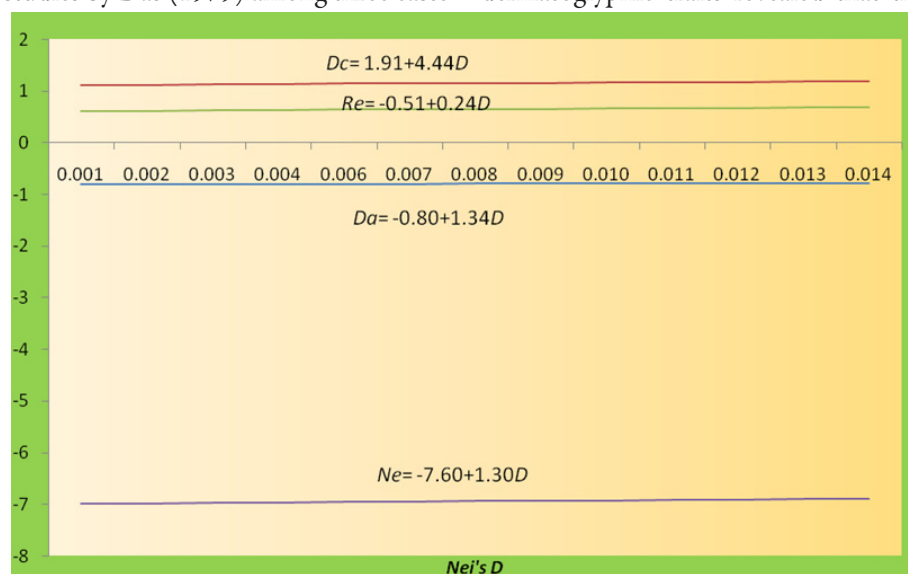
Fig. 1. Linear regression equations showing relationship between Nei's *D* with other distance measures
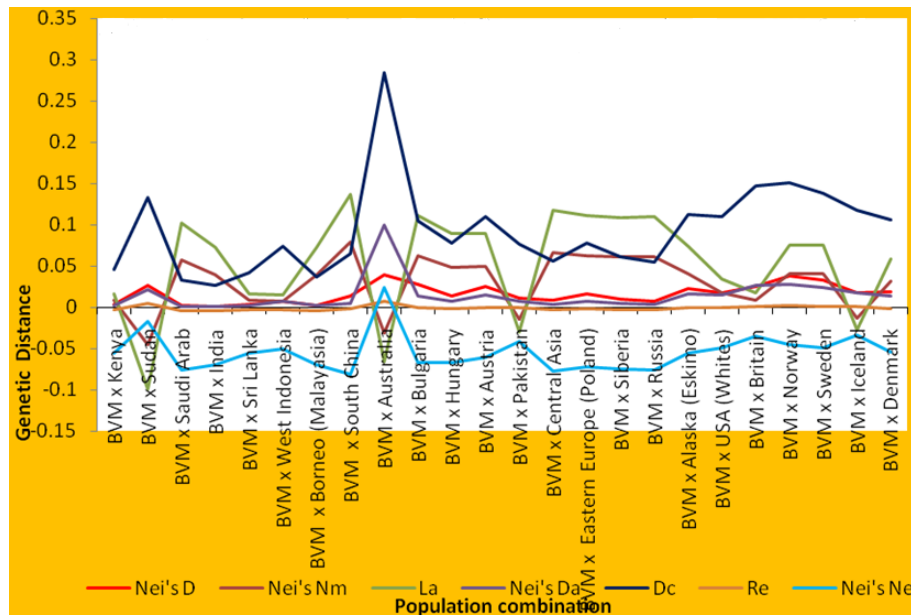
16

Fig. 2. Comparison of various genetic distance measures between BVM and other nations for *ABO* gene

sam 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).

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

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

Correlation analysis

The estimates of correlation coefficients between any two distance measures (Tab. 3) revealed that Nei's *D* showed highly significant (p=0.01) positive correlation with Cavalli-Sforza and Edwards chord distance *Dc* (0.90), Reynolds *Re* (0.90), Nei's *Da* (0.74) and Nei's *Ne* (0.63). This indicated great similarity among these four distance measures and anyone of these measures could be used instead of all the four measures in genetic analysis. But due to very high magnitude of positive correlation of Nei's *D* with Cavalli-Sforza and Edwards chord distance *Dc* and Reynolds *Re*, the use of anyone measure out of these three measures would be more effective in genetic analysis. Nei's *D* showed non-significant negative correlation with Nei's minimum distance *Nm* and Latter's distance *La*.

Regression analysis

Nei's *D* is the most widely used genetic distance measure in research programs. Assuming Nei's *D* as a dependent variable and anyone of the remaining distance measures (*Da, Dc, Re* or *Ne*) as independent variable, the linear regression equations of the latter on Nei's *D* were estimated (Tab. 4; Fig. 1). Since Nei's minimum distance *Nm* and Latter's distance *La* did not show significant correlation with Nei's *D*, hence *Nm* and *La* were not used as dependent variables in determining the linear regression equation with Nei's *D*.

These regression equations could be used to estimate the magnitude of the particular genetic distance measure with a given value of Nei's *D* between two populations. But the accuracy of the particular genetic estimates calculated from a given estimate of Nei's *D* using the above linear regression equations would decrease with the decreasing value of correlation coefficients. In the regression equation $y = A+Bx$, the *B* estimate represents the regression coefficient (slope) for linear regression and the regression constant *A* represents the magnitude of the *y*-intercept *i.e.* the distance from the origin to the point where the straight line intersects the *y*-axis.

Tab. 3. Correlation coefficients of *Nei's D* with other genetic distance measures

Distance measure	Nei's *Nm*	*La*	Nei's *Da*	*Dc*	*Re*	Nei's *Ne*
Nei's *D*	-0.24	-0.25	0.74**	0.90**	0.90**	0.63**

** Significant at p=0.01

Tab. 4. Linear regression equations of different distance measures on Nei's *D*

Dependent variable (*y*)	Independent variable (*x*)	Regression equation $y = A+Bx$
Nei's *Da*	Nei's *D*	$Da = -0.80 + 1.34D$
Cavalli-Sforza *Dc*	Nei's *D*	$Dc = 1.91 + 4.44D$
Reynolds *Re*	Nei's *D*	$Re = -0.51 + 0.24D$
Nei's *Ne*	Nei's *D*	$Ne = -7.60 + 1.30D$

Conclusions

The results of the present study revealed that the Muslims of Barak Valley in Assam showed the highest genetic distance (Nei's *D*) with Australians but the lowest with Indians in general. Correlation analysis between Nei's *D* and other distance measures revealed that Nei's *D* had highly significant ($p=0.01$) positive correlation with Cavalli-Sforza and Edwards chord distance *Dc* (0.90), Reynolds *Re* (0.90), Nei's *Da* (0.74) and Nei's *Ne* (0.63). This indicated great similarity among these four distance measures. Nei's *D* is the most widely used measure in research programs. Assuming Nei's *D* as a dependent variable and any one of the remaining distance measures (*Da, Dc, Re* or *Ne*) as independent variable, the linear regression equations of the latter on Nei's *D* in the form $y = A+Bx$ were estimated. These regression equations were $Da = -0.80 + 1.34D$, $Dc = 1.91 + 4.44D$, $Re = -0.51 + 0.24D$ and $Ne = -7.60 + 1.30D$.

References

Anees M, Mirza MS (2005). Distribution of ABO and Rh blood group alleles in Gujrat region of Punjab, Pakistan. Proc Pak Acad Sci 42(4):233-238.

Breguet G, Ney R, Gerber H, Garner MF (1986). Treponemal serology and blood groups on Bali Island, Indonesia. Genitourin Med 62:298-301.

Cavalli-Sforza LL, Edwards AWF (1967). Phylogenetic analysis: models and estimation procedure. Am J Hum Genet 19:233-257.

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.

Crow JF (1993). Felix Bernstein and the first human marker locus. Genet 133(1):4-7.

Danker-Hopfe H, Das BM, Walter H, Das PB, Das R (1988). Anthropological studies in Assam, India-Differentiation processes among Assamese populations. Anthropol Anz 46(2):159-184.

Das BM (1979). Physical variation in three Assamese castes. Anthropol Anz 37(3):204-210.

Harris M, Taylor G, Taylor J (2007). Maths and Stats for the life and medical sciences. New Delhi: Viva Books Pvt Ltd, 157-167 p.

Hedrick PW (2005). Genetics of populations (3rd Ed). Sudbury: Jones and Bartlett Publishers, 63-106 p.

http://www.bradshawfoundation.com

Kamil M, Han Al-Jamal, Yusoff NM (2010). Association of ABO blood groups with diabetes mellitus. Libyan J Med 5 at http://www.journals.sfu.ca/coactions/index.php/ljm/article/view article/4847/5365.

Landsteiner K (1900). Note the antifermantativen, lytic and agglutinating activity of blood serum and lymph. Centralblatt f. Bacteriology, Infectious Diseases and Parasites Customer. 27:357-362 (http://en.wikipedia.org/wiki/Karl_Landsteiner).

Latter BD (1972). Selection in infinite populations with multiple alleles. III. Genetic divergence with centripetal selection and mutation. Genetics 70:475-490.

Libiger O, Nievergelt CM, Schork NJ (2009). Comparison of genetic distance measures using human SNP data. Hum Biol 81(4):389-406.

Nei M (1972). Genetic distance between populations. Am Nat 106:283-292.

Nei M, Tajima F, Yateno Y (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J Mol Evol 19:150-173.

Papiha SS, Mukherjee BN, Chahal SMS, Malhotra KC, Roberts DF (1982). Genetic heterogeneity and population structure in north-west India. Ann Hum Biol 9(3):235-251.

Reynolds J, Weir BS, Cockerham CC (1983). Estimation of the coancestry coefficient: basis for a short term genetic distance. Genetics 105:767-779.

Roy M, Datta U, Mitra M, Singhrol CS (1990). Genetic distance and gene diversity among ten endogamous groups in Chattisgarh, Central India. Int J Anthropol 5(2):109-115.

Roychoudhury AK (1982). Genetic distance between Jews and Non-Jews of four regions. Hum Hered 32(4): 259-263.

Revavov AA, Asanov A, Lunga IN, Bakhramov SM (1983). Frequencies of ABO system blood groups and haptoglobins in Uzbekistan-the problems of sampling studies. Genetika 19(7):1193-1197.

Sokal RR (1988). Genetic, geographic and linguistic distances in Europe. Proc Natl Acad Sci, USA 85:1722-1726.

Triantaphyllidis CD, Kouvatsi A, Kaplanoglou L (1983). The genetic distances between the inhabitants of nine Mediterranean countries and the three major human races. Hum Hered 33(2):137-139.

www.bloodbook.com/world-abo.html.