

Trans-National Genetic Distance and Genetic Identity of Barak Valley Hindus en Route the Journey of Mankind from Africa for ABO Gene

Supriyo CHAKRABORTY¹, Saumendra Kumar BHATTACHARJEE²

¹Assam University, Department of Biotechnology, Silchar 788 011, Assam, India; supriyoch_2008@rediffmail.com (*corresponding author)

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

Abstract

The present study aimed at estimating the genetic distance and genetic identity between Barak Valley Hindus and other twenty four nations for ABO blood group gene along the route of historic journey of mankind from Africa as proposed by Stephen Oppenheimer to gain insights on the evolutionary relationship and genetic closeness of the Hindus with other nations. Barak Valley Zone, located in southern part of Assam state in North East India, has inhabited the major endogamous group, the Hindus, for several centuries. Over the last few decades, they have maintained distinct culture and life style. This study used ABO gene frequency data of these populations to estimate Nei's standard genetic distance and genetic identity of population genetics between Barak Valley Hindus and other nations. The historic journey of mankind commenced from Africa about 200,000 years ago (www.bradshawfoundation.com). Genetic distance estimate ranged from 0.07 to 5.18%. Barak Valley Hindus (BVH) showed relatively low genetic distance for ABO gene with the populations of Saudi Arabia (0.07%), India (0.13%), Borneo (0.40%), Russia (0.59%), Central Asia (0.60%), Siberia (0.60%), South China (0.71%) and Sri Lanka (0.93%) suggesting high genetic identity and possible evolutionary relationship of BVH during migration with these nations. But the BVH showed highest genetic distance with Australia (5.18%) followed by Norway (4.13%), Sudan (3.89%) and Sweden (3.60%) indicating low genetic identity of BVH with these nations. Migration was not the key determining factor in changing the ABO gene frequency in human populations.

Keywords: ABO gene, genetic distance, Hindu, genetic identity, other nations

Introduction

Genetic variation is central to all evolutionary change. Hence, the study of genetic variation spans plant and animal breeding, conservation biology, human genetics and comparative genomics. Despite its fundamental importance in several areas of genetics, there has been a long period of struggle to measure genetic variation. It was a struggle because markers used to measure genetic variation, such as morphological and biochemical markers often were limited and influenced by environmental factors. These limitations also impeded progress toward the understanding of genetic architecture and its influence on phenotype, including human disorders (Lewontin, 1974). Fortunately, advances in population genetics, molecular genetics and bioinformatics since 1980s, specifically our knowledge of genetic mechanisms governing the easily detectable phenotypes have revolutionized our ability to describe variation at all levels of genetic organization: genes, individuals, populations, species and genera.

ABO blood group system is one of the most significant blood factors in clinical applications involving blood transfusion. The four blood groups in ABO system also serve as good genetic markers for various scientific studies. With our recent ability to rapidly sequence genes, the ABO and Rh blood groups are also proving to be valuable assets for determining the human migration patterns and origins across the globe.

The ABO gene, a multiple allelic gene located on chromosome 9 of human genome, determines the blood group of a person in ABO system. ABO blood group typing is now routinely done in human beings for medical, legal and identification purposes. Human blood group is a good genetic marker as it does not change in the life time of a person.

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 disomic condition.

Different 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 and genetic identity 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. Hence it is imperative to study one or a few genes at a time to understand the evolutionary process in mankind.

Barak Valley Zone, named after the mighty river Barak flowing through the zone, located in southern part of Assam state in North East India, has inhabited the major endogamous religious group, the Hindus, 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. 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 Hindus of Barak Valley with other nations/populations in the route of migration of mankind that commenced from Africa nearly 200,000 years ago. The present study was, therefore, taken up to estimate the genetic distance and the genetic identity of Barak Valley Hindus with 24 nations in the historic route of human migration as proposed by Stephen Openheimer (at <http://www.bradshawfoundation.com>) on the basis of *ABO* gene frequency and to assess the genetic proximity and the evolutionary relationship of the Barak Valley Hindus with other nations.

Several studies have been carried out on genetic distance based on ABO system in different human populations. Zhu *et al.* (2009) studied the distribution of ABO blood group allele in the Chinese Han population and identified three novel alleles in that population. Spitsyn *et al.* (2009) analyzed the genetic position of Chuvashes in the system of Finno-Ugric and Turkic-speaking peoples. Serological and biochemical polymorphisms in marker genes of the ABO, MN, RH, FY,HP,TF, ACP1, ESD and GLO1 systems were studied in 369 individuals of Chuvash ethnicity from Morgaushskii, Mariinsko-Posadskii and Yadrinskii districts of Chuvash Republic. They observed a linear relationship between genetic and geographical distances by examining 11 ethno territorial groups in Northeastern Europe and Western Siberia.

Ara *et al.* (2008) analyzed the gene diversity among some Muslim populations of western Uttar Pradesh, India and studied the distribution of ABO, Rh (D) and PTC tasting ability to understand the genetic structure and micro differentiation among different endogamous Muslim populations namely Syed, Sheikh, Pathan, Ansari and Shia of Aligarh. For ABO system, only Pathan and Ansari showed significant differences in allele frequencies while other combinations showed non-significant values. The Pathan and Ansari populations separated earlier than the Sheikh, as well as Syed and Shia cluster which might have been the migrants to Indian population from outside quite later.

Simmons *et al.* (2007) carried out the admixture estimates based on ABO, Rh and nine STRs in two Venezuelan regions. The European contribution was high in both groups. Aarzo and Mohammad (2005) carried out

a study on the gene frequencies of ABO, Rh, PTC taste ability, sickling and G6PD systems for different endogamous groups: Sheikh, Syed, Pathan, Ansari, Saifi and Hindu Bania. All the groups at most loci showed statistically non-significant differences, except for ABO and PTC traits, for which interpopulational differences were seen.

Materials and methods

The present study comprised of 25 human populations including the Barak Valley Hindus and other 24 populations in the route of journey of mankind from Africa and was conducted in the Department of Biotechnology, Assam University (A Central University), Silchar 788 011, Assam, India from July, 2008 to till September, 2010 (ongoing work). The frequencies of *O*, *A* and *B* alleles of *ABO* blood group gene in each population were calculated from the ABO blood group distribution/phenotypic data obtained from diverse sources (Tab. 2) using the following formulae (Hedrick, 2005):

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

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

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

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

To calculate the genetic distance between any two populations for *ABO* gene, at first the genetic identity (I) was calculated (Hedrick, 2005) as:

$$I = \frac{J_{xy}}{\sqrt{J_x J_y}} \text{ where } J_{xy} = \sum_{i=1}^n p_{i,x} p_{i,y}$$

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

And $p_{i,x}$ and $p_{i,y}$ are the frequencies of the 'i'th allele in two populations. The value of I (identity) ranges from 0 to 1. It was multiplied by 100 to express genetic identity in percentage for convenience.

Nei's standard genetic distance (D) between two populations for *ABO* gene was calculated (Nei, 1972) as:

$$D = \ln(I) \text{ where } I = \text{genetic identity (in proportion)}$$

$$= \ln J_{xy} + \frac{1}{2} \ln J_y + \frac{1}{2} \ln J_x$$

D value was multiplied by 100 to express it in percentage. D value in percent can also be calculated as 100- I (%).

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

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

Results and discussion

Barak Valley 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.

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. 2). In general, the frequency of *O* allele was the highest in all the populations. *B* allele was not reported in Australians.

Genetic distance and genetic identity

The estimates of Nei's genetic distance between Barak Valley Hindus (BVH) with other 24 nations for *ABO* gene (Tab. 3 and Fig. 1) revealed that BVH showed relatively low genetic distance with nations like Saudi Arabia (0.07%), India (0.13%), Borneo (0.40%), Russia (0.59%), Central Asia (0.60%), Siberia (0.60%), South China (0.71%) and Sri Lanka (0.93%). But the BVH showed highest genetic distance with Australia (5.18%) followed by Norway (4.13%), Sudan (3.89%) and Sweden (3.60%). The estimates of genetic distance ranged from 0.07 to 5.18%.

Genetic identity is the reverse of genetic distance. The BVH showed high genetic identity with nations showing low genetic distance *e.g.* nations like Saudi Arabia, India, Borneo, Russia, Central Asia and Siberia. The genetic distance of BVH with nations in the route of journey of mankind for *ABO* gene did not reflect a gradual increase in genetic distance with the increase in geographical distance. It suggests that migration was not the key determining factor in changing the *ABO* gene frequency in human populations. In other words, genetic factors like mutation, selection and genetic drift in addition to migration might have played major role in changing the *ABO* gene frequency in human populations.

Several studies have been carried out on genetic distance measurements across different populations. Roy *et al.*, (1990) analyzed the genetic distance and the gene diversity among 10 endogamous groups in Chattisgarh,

India using the gene frequency data of three genetic loci and observed that the gene differentiation among these population groups is only about 2 per cent.

Papiha *et al.* (1982) observed that the genetic differentiation in Indian populations was low (0.26-1.70%). In Assam, Das (1979) analyzed the variation among three caste populations namely Brahmin, Kalita and Kaibarta on the

Tab. 2. Estimates of allele frequency of *ABO* gene in the nations en route the journey of mankind

Sl. No	Population	Allele Frequency				Reference*
		<i>O</i>	<i>A</i>	<i>B</i>	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.com
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 <i>et al.</i> (1986)
7	Borneo (Malaysia)	0.62	0.22	0.16	1.00	Kamil <i>et al.</i> (2010)
8	South China	0.53	0.23	0.24	1.00	www.bloodbook.com
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 <i>et al.</i> (1983)
15	Eastern Europe (Poland)	0.57	0.28	0.15	1.00	www.bloodbook.com
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 Hindus	0.60	0.19	0.21	1.00	Chakraborty (2010)

*Detailed reference in text

Tab. 3. Genetic distance and genetic identity between Barak Valley Hindus (BVH) and other nations *en route* the journey of mankind for *ABO* gene

Sl. No	Population combination	Genetic Distance (<i>D</i>)	<i>D</i> (%)	Genetic Identity (<i>I</i>)	<i>I</i> (%)
1	BVH-Kenya	0.0101	0.01	0.9899	98.99
2	BVH-Sudan	0.0389	3.89	0.9611	96.11
3	BVH-Saudi Arabia	0.0007	0.07	0.9993	99.93
4	BVH-India (Overall)	0.0013	0.13	0.9987	99.87
5	BVH-Sri Lanka	0.0093	0.93	0.9907	99.07
6	BVH-West Indonesia	0.0124	1.24	0.9876	98.76
7	BVH-Borneo (Malayasia)	0.0040	0.40	0.9960	99.60
8	BVH-South China	0.0071	0.71	0.9929	99.29
9	BVH-Australia	0.0518	5.18	0.9482	94.82
10	BVH-Bulgaria	0.0267	2.67	0.9733	97.33
11	BVH-Hungary	0.0143	1.43	0.9857	98.57
12	BVH-Austria	0.0266	2.66	0.9734	97.34
13	BVH-Pakistan	0.0193	1.93	0.9807	98.07
14	BVH-Central Asia	0.0060	0.60	0.9940	99.40
15	BVH-Eastern Europe (Poland)	0.0144	1.44	0.9856	98.56
16	BVH-Siberia	0.0060	0.60	0.9940	99.40
17	BVH-Russia	0.0059	0.59	0.9941	99.41
18	BVH-Alaska	0.0262	2.62	0.9738	97.38
19	BVH-USA (Whites)	0.0233	2.33	0.9767	97.67
20	BVH-Britain	0.0335	3.35	0.9665	96.65
21	BVH-Norway	0.0413	4.13	0.9587	95.87
22	BVH-Sweden	0.0360	3.60	0.9640	96.40
23	BVH-Iceland	0.0279	2.79	0.9721	97.21
24	BVH-Denmark	0.0226	2.26	0.9774	97.74

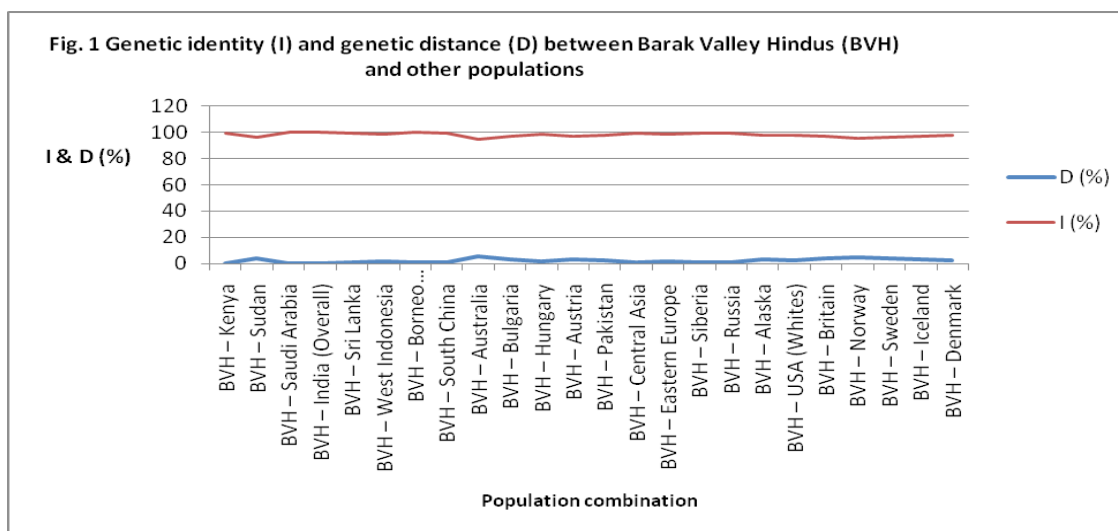


Fig. 1. Genetic identity (I) and genetic distance (D) between Barak Valley Hindus (BVH) and other populations

basis of ABO blood groups and other anthropometric characters. He asserted that the Kaibarta stand apart from the Brahmin and the Kalita, who are similar to each other. Danker-Hopfe *et al.* (1988) examined thirteen Assamese populations including two Muslim groups for the distribution of anthropometric, anthroposcopic and dermatoglyphic traits and observed 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).

Roychoudhury (1982) analyzed the genetic distance between Jews and Non-Jews using gene frequency data of nine blood groups and protein loci. He 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. They asserted that the Algerians were closer to Negroids while the other Mediterraneans were closer to Caucasoids.

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

Conclusions

Analysis of the genetic distance and genetic identity between Barak Valley Hindus and other twenty four nations along the route of historic journey of mankind from Africa as proposed by Stephen Oppenheimer based on ABO blood group gene frequency revealed insights on the evolutionary relationship and genetic closeness of the Hindus with other nations. Genetic distance estimate ranged from 0.07 to 5.18%. Barak Valley Hindus (BVH) showed relatively low genetic distance for *ABO* gene with the populations of Saudi Arabia (0.07%), India (0.13%), Borneo (0.40%), Russia (0.59%), Central Asia (0.60%), Siberia (0.60%), South China (0.71%) and Sri Lanka (0.93%) suggesting high genetic identity and possible evolutionary relationship of BVH during migration with these nations. But the BVH showed highest genetic distance with Australia (5.18%) followed by Norway (4.13%), Sudan (3.89%) and Sweden (3.60%) indicating low genetic identity of BVH with these nations. Possibly migration was not the key determining factor in changing the *ABO* gene frequency in human populations.

Acknowledgements

The authors are thankful to Assam University, Silchar, Assam for providing the necessary facilities to take up this research work. The authors are also thankful to Dr. Z.B. Zeng, W. N. Reynolds Professor of Genetics and Statistics, North Carolina State University, Raleigh, USA and to Dr. G.N. Hazarika, Assam Agricultural University, Jorhat, Assam, India for their valuable suggestions.

References

- Aarzo SS, Afzal M (2005). Gene diversity in some Muslim populations of North India. *Human Biol* 77(3):343-353
- 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.
- Ara G, Siddique YH, Beg T, Afzal M (2008). Gene diversity among some Muslim populations of Western Uttar Pradesh, India. *Anthropol* 10(1):5-9.
- Breguet G, Ney R, Gerber H, Garner MF (1986). Treponemal serology and blood groups on Bali Island, Indonesia. *Genitourin Med* 62:298-301.
- 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-Hopf 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.
- Hedrick PW (2005). *Genetics of Populations* (3rd Ed.). Jones and Bartlett Publishers, Sudbury, USA, p. 63-106.
- Kamil M, Han Al-Jamal, Yusoff NM (2010). Association of ABO blood groups with diabetes mellitus. *Lib J Med*, 5 (<http://journals.sfu.ca/coactions/index.php/ljm/article/view/article/4847/5365>).
- Landsteiner K (1900). Note the antifermantative, lytic and agglutinating activity of blood serum and lymph. *Centralblatt f. Bacteriol Infect Dis Parasit Cust* 27:357-362.
- Lewontin RC (1974). *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York, p. 19-94.
- Nei M (1972). Genetic distance between populations. *Am Nat* 106:283-292.
- 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.
- Revavov AA, Asanov A, Lunga IN, Bakhranov SM (1983). Frequencies of ABO system blood groups and haptoglobins in Uzbekistan-The problems of sampling studies. *Genetika* 19(7):1193-1197.
- 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.
- Simmons AD, Rodriguez-Arroyo G, Rodriguez-Larralde A (2007). Admixture estimates based on ABO, Rh and nine STRs in two Venezuelan regions. *Ann Hum Biol* 34(10):56-67.
- Sokal RR (1988). Genetic, geographic and linguistic distances in Europe. *Proc Natl Acad Soc, USA* 85:1722-1726.
- Spitsyn VA, Batsevich VA, El-Chinova GI, Kobiliansky ED (2009). Genetic position of Chuvashes in the system of Finno-Ugric and Turkic speaking peoples. *Russian J Genet* 45(9):1117-1123.
- Triantaphyllidis CD, Kouvatzi 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.
- Zhu F, Tao S, Xu X, Ying Y, Hong X, Zhu H, Yan L (2010). Distribution of ABO blood group allele and identification of three novel alleles in the Chinese Han population. *Vox Sanguinis* 98(4):554-559.