Available Online at http://www.recentscientific.com

ISSN: 0976-3031

International Journal of Recent Scientific Research
Vol. 5, Issue, 11, pp.2136-2140, November, 2014

International Journal of Recent Scientific Research

RESEARCH ARTICLE

ALLELE FREQUENCY AND GENETIC DISTANCE ANALYSIS OF BARAK VALLEY POPULATION WITH GLOBAL POPULATIONS FOR SICKLE CELL ANAEMIA

Arup Kumar Malakar, *Supriyo Chakraborty and Prosenjit Paul

Department of Biotechnology, Assam University, Silchar, Assam, India

ARTICLE INFO

Article History:

Received 5th, October, 2014
Received in revised form 12th, October, 2014
Accepted 6th, November, 2014
Published online 28th, November, 2014

Key words:

sickle cell anaemia, allele frequency, heterozygosity, genotype frequency

ABSTRACT

The inherited disorders of hemoglobin are the commonest monogenic disorders in India. Of these, the sickle cell disease is genetically transmitted as hemo-globino pathy. Its prevalence has ranged from 9.4-22.2% in different communities in endemic areas with a varying frequency of sickle cell gene. The present investigation attempts to study the population genetics of sickle cell disease for heterozygosity, genetic distance and genetic identity and its prevalence in Barak valley region of Northeastern India with respect to 23 global populations. Our results revealed that the overall population of India (frequency of 0.109) was much more affected by sickle cell anaemia than the Barak valley population (frequency of 0.089). However, from the comparative analysis of the genetic distance and genetic identity between Barak valley population with other global populations, we observed that the Barak valley population showed the highest genetic distance (0.00076) and hence the lowest genetic identity (0.99924) with DR Congo for sickle cell anaemia. But the Barak valley population showed the lowest genetic distance (0.00003) and the highest genetic identity (0.99997) with the US population for sickle cell anaemia.

INTRODUCTION

Genetic disorders are heritable and are passed down from the parents to their offspring. In India, the occurrence of congenital abnormalities is up to 37 per 1000 births (Talukder and Sharma 1979). An allele is one of the alternative forms of a gene or genetic locus. Allele frequency is the proportion of a particular allele of a gene. In population genetics, allele frequencies are used to depict the amount of genetic diversity at the individual, population and species level. The study of single locus or a few loci at a time in human beings assumes tremendous significance in population genetics in view of the difficulty in analyzingthe alleles of nearly 25,000 genes simultaneously.

Sickle-cell anaemia is an autosomal recessive genetic hematologic disorder characterized by the irregular sickle shape of the erythrocytes and is caused by a point mutation in the β-globin chain of haemoglobin. In this disease, the hydrophilic amino acid glutamic acid is replaced with the hydrophobic amino acid valine at the 6th position. The β-globin gene is found on chromosome 11 (Schmaier and Lazarus 2011). The association of two wild-type α-globin subunits with two mutant β-globin subunits forms haemoglobin S (HbS). The most common types of sickle cell disease (SCD) are SS, SC and S beta-thalassemia. Other more rare forms include SD-Punjab, SO Arab, SLepore and SE disease. The sickle haemoglobin (HbS) mutation confers a genetic advantage against malaria, so the carrier frequency in the population is the highest in areas where malaria is highly endemic, including Africa, Mediterranean Europe, the Middle East, the Caribbean, South and Central America and some regions of India (Weatherall 2011). In the US, the SCD is the most common single gene disorder among the African-Americans population and affects more than 70,000 people with an estimated frequency of 1 in every 375 newborns (Steiner and Miller 2006). In the UK, it affects 1 in 2400 live births across ethnic groups, and more than 12000 individuals are now living with SCD. This makes SCD the most common and fastest-growing genetic disorder in the UK. Sickle cell variants namely haemoglobin S alpha and beta-thalassemia are commonly found in some European countries like Turkey, Italy and Greece. The incidence of SCD in other parts of the world is increasing due to population migration (LIVEX 2009). Although sickle cell anemia is believed to have originated among the Venddoits in the Middle East, still there are some contradictory reports on the origin of the disease in Africa. The disease was first described as a clinical entity by Herrick from Chicago in 1910 (Savitt and Goldberg 1989). However, Dresbach (1904) reported a case of peculiar red cells occurring in a Negro medical student. According to Desai and Hiren (2004), it is believed that the first presentation of SCD was in 1670 in one of the Ghanian families in Africa (Desai and Dhanani 2004). At first sickle cell anemia was believed to be familial, later on it was reported that sickle cell anemia was an autosomal recessive inherited disease associated with the sickling of the red blood cell as a result of oxygen depletion (Serjeant 2001).

The present study was undertaken to estimate the magnitude of genetic distance and genetic identity between the Barak valley population and some global populations on the basis of the allele frequency for sickle cell anaemia. This might help us understand the pattern of genetic variation and the extent of genetic differentiation of Barak valley population from global populations from the point of quantitative genetics.

* Corresponding author : **Supriyo Chakraborty**
Department of Biotechnology, Assam University, Silchar, Assam, India

MATERIALS AND METHODOLOGY

The study comprised of a total of 23 different populations for sickle cell anaemia along with Barak valley populations. The population data of SCA were obtained from the published literature and websites given in **Table 2**. The allele frequency, homozygosity and heterozygosity, genetic distance and genetic identity for sickle cell anaemia were estimated using the formulae given by Hardy Weinberg (1908) and Masatoshi Nei (1972). The Hardy-Weinberg Equation is very useful in predicting the percentage of population that may be heterozygous carriers of recessive alleles for certain genetic diseases. This law further provides information how gene frequencies change from generation to generation.

The allele frequency of sickle cell anaemia was estimated at Hardy Weinberg (HW) equilibrium *i.e.*

$p^2+2pq+q^2 = 1$,

where,p = frequency of the dominant allele
q = frequency of the recessive allele
p^2= genotype frequency of dominant homozygotes
2pq = genotype frequency of heterozygote
q^2= genotype frequency of recessive homozygotes

Using the allele frequency estimates, the standard genetic distance (*D*) between two populations was calculated. For this, at first the genetic identity (*I*) was calculated as:

$$I = \frac{J_{xy}}{\left(J_x J_y\right)^{\frac{1}{2}}}$$

Where,

$$J_{xy} = \sum_{i=1}^{n} P_{i.x} P_{i.y}$$

$$J_x = \sum_{i=1}^{n} P^2_{i.x}$$

$$J_y = \sum_{i=1}^{n} P^2_{i.y}$$

and, $P_{i.x}$ and $P_{i.y}$ are the frequencies of the i^{th} allele in populations x and y.

D (genetic distance) was then estimated as
D = -ln(I)
$= -\ln J_{xy}+1/2 \ln J_y+1/2 \ln J_x$
The Hardy-Weinberg heterozygosity (*HE*) of a population for a particular locus with *n* alleles can be calculated as:

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

Where, P_i^2 = genotype frequency of observed homozygotes

Results

In this study, we analyzed the allele frequency of the gene for sickle cell anaemia, heterozygosity, genetic distance and genetic identity of the Barak valley (BV) population with a total of 23 other global populations for sickle cell anaemia. Our analysis revealed that the frequency of sicklers in India (0.109) is significant when compared to other world populations and followed by the populations of Oman, DR Congo, Uganda, Nigeria and Bahrain (figure 1B). The genotype frequency and heterozygosity in India were found to be 0.014 and 0.191 respectively (figure 1A). Two populations i.e. Saudi Arabia and Kuwait showed the SCA allele frequency lying between that of India and the Barak Valley population. The estimates of HbS alleles in twenty-three world populations along with Barak valley are presented in table 1. The population size and disease incidence of sickle cell anaemia in Barak Valley population and 23 other populations are tabulated in table 2, showing the distribution of the frequency of diseased population world-wide in which Nigeria (1887600) preceded USA (587310) and India (68680), respectively.

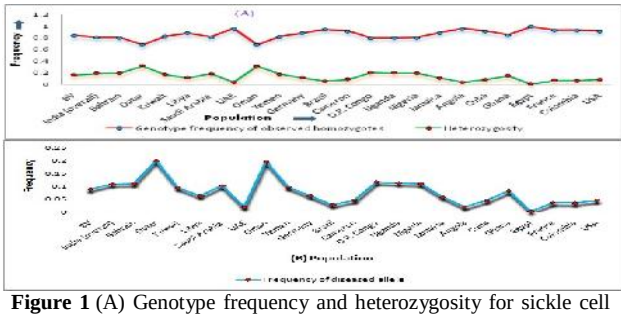

Figure 1 (A) Genotype frequency and heterozygosity for sickle cell anaemia (SCA)
(B) Allele frequency for SCA including BV and world populations

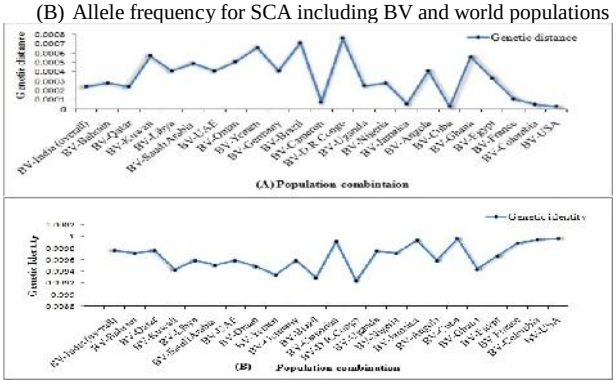

Figure 2 (A) Genetic distance of BV population with 23 global populations for sickle cell anaemia; (B) Genetic identity between BV and 23 other world populations for sickle cell anaemia

Heterozygosity of a gene is widely used to measure the extent of genetic variation in natural population. Heterozygosity values range from 0 to 1, where 0 indicates no heterozygosity/genetic variation whereas 1 means the highest genetic variation. From our analyses it was evident that the Qatar population (0.320) displayed the highest genetic variation and Egypt (0.006) showed the lowest genetic variation for the gene as indicated by heterozygosity. Both the Barak valley population and the Indian population (as a whole) represented moderate values of heterozygosity indicating moderate genetic variation for the concerned gene. The comparison of Nei's genetic distance between Barak valley population with other twenty-three populations for SCA (table 3) revealed that Barak valley population recorded relatively low genetic distance with USA (0.00003), Cuba (0.00003), Colombia (0.00005), Jamaica (0.00006) and Cameron (0.00008) for the gene. But Barak valley population revealed the high genetic distance with DR Congo (0.00076), Brazil (0.00071), Yemen (0.00066), and Kuwait (0.00057). The values of genetic distance ranged from 0.00003 to 0.00076 (figure 2A). Genetic identity is the reverse of the genetic distance. In the present analysis, Barak valley population showed utmost genetic identity with the populations of the USA (99.997%), Cuba (99.997%), Colombia (99.995%),

Table 1 Estimates of allele frequency, genotype frequency and heterozygosity of sickle cell anaemia in BV population and 23 world populations

Population	Frequency of diseased allele	Frequency of normal allele	Genotype frequency of diseased condition	Genotype frequency of normal condition	Genotype frequency of observed homozygotes	Heterozygosity
BV	0.089	0.911	0.008	0.830	0.838	0.162
India (overall)	0.109	0.891	0.014	0.795	0.809	0.191
Bahrain	0.110	0.890	0.012	0.792	0.804	0.196
Qatar	0.2	0.8	0.04	0.64	0.68	0.32
Kuwait	0.095	0.905	0.009	0.819	0.828	0.172
Libya	0.063	0.937	0.004	0.878	0.882	0.118
Saudi Arabia	0.103	0.897	0.011	0.805	0.816	0.184
UAE	0.02	0.98	0.0004	0.960	0.9604	0.040
Oman	0.195	0.805	0.038	0.648	0.686	0.314
Yemen	0.097	0.903	0.009	0.815	0.824	0.176
Germany	0.063	0.937	0.004	0.878	0.882	0.118
Brazil	0.028	0.972	0.0008	0.945	0.946	0.054
Cameron	0.046	0.954	0.002	0.910	0.912	0.088
DR Congo	0.117	0.883	0.014	0.780	0.794	0.206
Uganda	0.113	0.887	0.013	0.787	0.8	0.2
Nigeria	0.11	0.89	0.012	0.792	0.804	0.196
Jamaica	0.059	0.941	0.003	0.885	0.888	0.112
Angola	0.02	0.98	0.0004	0.960	0.9604	0.040
Cuba	0.045	0.955	0.002	0.912	0.914	0.086
Ghana	0.082	0.918	0.007	0.843	0.85	0.15
Egypt	0.003	0.997	0.000009	0.994	0.994	0.006
France	0.036	0.964	0.001	0.929	0.93	0.07
Colombia	0.035	0.965	0.001	0.931	0.932	0.068
USA	0.045	0.955	0.002	0.912	0.914	0.086

Table 2 Population size and disease incidence of sickle cell anaemia in BV population and 23 other populations

Population	Population size	Disease incidence	Source
BV	250	2	Present study
India (overall)	9315474	68680	(Feroze and Aravindan 2000; Patra, Chauhan et al. 2011; Arjunan 2013), www.rmrct.org, www. Cips.org, www.icmr.nic.in
Bahrain	5685	68	(Al-Arrayed, Hafadh *et al*. 2003)
Qatar	1702	68	(Fawzi, Al Hilali *et al*. 2003)
Kuwait	2386	21	(Marouf, D'souza *et al*. 2002)
Libya	1350	5	(Jain 1985)
Saudi Arabia	30055	318	(Jain 1985)
UAE	22200	9	(Al Hosani, Salah *et al*. 2005)
Oman	5000	190	(White, Byrne *et al*. 1986)
Yemen	5000	47	(White, Byrne *et al*. 1986)
Germany	34084	14	(Lobitz, Frömmel *et al*. 2014)
Brazil	1217833	974	(de Castro Lobo, Ballas *et al*. 2014)
Cameron	467	1	(Badens, di Montemuros *et al*. 2000)
DR Congo	31304	428	(Tshilolo, Aissi *et al*. 2009)
Uganda	857	11	(Okwi, Byarugaba *et al*. 2010)
Nigeria	156000000	1887600	(Luzzatto 2012)
Jamaica	8000	28	(Serjeant, Forbes *et al*. 1974)
Angola	758000	11996	(Piel, Patil *et al*. 2013)
Cuba	108000	185	(Piel, Patil *et al*. 2013)
Ghana	737000	4924	(Piel, Patil *et al*. 2013)
Egypt	72000000	757	www.who.int/genomics
France	253466	341	www.api.ning.com
Colombia	1729	2	(Alvear, Barboza *et al*. 2012)
USA	293655405	587310	www.cureresearch.com

Jamaica (99.994%), Cameron (99.992%) for the gene whereas it showed less genetic identity with DR Congo (99.924%), Brazil (99.929%), Yemen (99.934%), and Kuwait (99.943%) for sickle cell anaemia (figure 2B).

DISCUSSION

The present study was carried out to analyze the allele frequency, heterozygosity, genetic distance and genetic identity of the gene governing sickle cell anaemia in global populations as well as in the population of Barak valley. Barak valley is geographically located in southern part of Assam state in North East India. It 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 an average rainfall of 318 cm and 146 rainy days per annum. Nearly 70% of the population depends on agriculture for livelihood. Several studies have been carried out on the genetic distance across different populations (Roy, Datta *et al.* 1990; Chakraborty 2010). The estimates of HbS alleles in twenty-three world populations along with Barak valley are presented in table 1. The frequency of HbS was found within the range of 0.200 to 0.003. The average frequency of HbS and HbA alleles in those populations were 0.078458 and 0.921542, respectively. Surprisingly the frequency of HbS in Barak valley was much higher than the average value of HbS for all the countries. According to Hardy- Weinberg equilibrium, the sum of frequencies of both alleles i.e. HbS and HbA should be unity. For SCA in the existing population the HbS allele on an average occurred more than HbA indicating that heterosis or outbreeding enhancement might have occurred in majority of those populations. In population genetics, allele frequency simply refers to gene frequency.

Table 3 Genetic distance and genetic identity of Barak Valley (BV) population with 23 other populations for sickle cell anaemia

Sl. No.	Population combination	Genetic distance (D)	Genetic identity (I)	I (%)
1	BV-India (overall)	0.00024	0.99976	99.976
2	BV-Bahrain	0.00028	0.99972	99.972
3	BV-Qatar	0.00024	0.99976	99.976
4	BV-Kuwait	0.00057	0.99943	99.943
5	BV-Libya	0.00041	0.99959	99.959
6	BV-Saudi Arabia	0.00049	0.99951	99.951
7	BV-UAE	0.00041	0.99959	99.959
8	BV-Oman	0.00051	0.00049	99.949
9	BV-Yemen	0.00066	0.99934	99.934
10	BV-Germany	0.00041	0.99959	99.959
11	BV-Brazil	0.00071	0.99929	99.929
12	BV-Cameron	0.00008	0.99992	99.992
13	BV-D R Congo	0.00076	0.99924	99.924
14	BV-Uganda	0.00025	0.99975	99.975
15	BV-Nigeria	0.00028	0.99972	99.972
16	BV-Jamaica	0.00006	0.99994	99.994
17	BV-Angola	0.00041	0.99959	99.959
18	BV-Cuba	0.00003	0.99997	99.997
19	BV-Ghana	0.00056	0.99944	99.944
20	BV-Egypt	0.00033	0.99967	99.967
21	BV-France	0.00011	0.99989	99.989
22	BV-Colombia	0.00005	0.99995	99.995
23	BV-USA	0.00003	0.99997	99.997

The estimated values of Nei's genetic distance between Barak valley population and other world populations for SCA are given in table 3. In case of SCA, genetic distance of Barak valley population was found the lowest with the population of USA (0.00003), Cuba (0.00003), Colombia (0.00005), Jamaica (0.00006) and Cameron (0.00008) indicating highest genetic identity of Barak valley population with these countries. On the other hand, the populations of DR Congo (0.00076), Brazil (0.00071), Yemen (0.00066) and Kuwait (0.00057) showed the highest genetic distance with Barak valley population and revealed the lowest genetic identity of these populations with Barak valley population. These results suggest that migration was not the key determining factor in changing the HbS gene frequency in human populations. Several studies were carried out on genetic distance across different populations (Chakraborty 2010). In contradiction to Roy (Roy, Datta *et al.* 1990), our analysis revealed that for sickle cell anaemia the genetic differentiation among different populations is nearly 7%.

CONCLUSION

Very few studies have been carried out to describe the clinical and genetic profile of sickle cell anaemia in the population of Barak valley. The present study was taken up to check the prevalence of sickle cell anaemia and compare its status world-wide. Our results revealed that the population in India, as a whole, was much more affected by sickle cell anaemia than the population of Barak valley as a part.

The frequency of HbS in Barak valley was much higher than the average values of HbS for all the countries. However on comparing the genetic distance and the genetic identity for sickle cell anaemia, we found that Barak valley population showed the highest genetic distance (0.00076) with DR Congo and hence the lowest genetic identity (0.99924). But Barak valley population showed the lowest genetic distance (0.00003) with USA and thus the highest genetic identity (0.99997) with USA. Genetic distance and genetic identity are, in fact, indirectly proportional to each other. From the above results, it can be concluded that sickle cell anaemia is less prevalent in Barak valley region as compared to India as a whole. Hence it is imperative to take up further research initiative on this field to elucidate the actual cause for lesser prevalence of sickle cell anaemia in Barak valley population.

Acknowledgement

We are thankful to Assam University for providing the necessary facilities to undertake the research work.

Conflict of interest

There is no conflict of interest in this research work.

References

Al-Arrayed, S., N. Hafadh, *et al.* (2003). "Student screening for inherited blood disorders in Bahrain." Eastern Mediterranean health journal= La revue de sante de la Mediterranee orientale= al-Majallah al-sihhiyah li-sharq al-mutawassit 9(3): 344-352.

Al Hosani, H., M. Salah, *et al.* (2005). "Incidence of haemoglobinopathies detected through neonatal screening in the United Arab Emirates." East Mediterr Health J 11(3): 300-307.

Alvear, C. C., M. Barboza, *et al.* (2012). "Pilot study of hemoglobinopathies in newborns of the Rafael Calvo maternity clinic of Cartagena, Colombia." Colombia Médica: CM 43(3): 196.

Arjunan, A. (2013). Prevalence of the Beta-S Gene and Sickle Cell Disease in India, University of Pittsburgh.

Badens, C., F. M. di Montemuros, *et al.* (2000). "Molecular basis of haemoglobinopathies and G 6 PD deficiency in the Comorian population." The Hematology Journal 1(4): 264-268.

Chakraborty, S. (2010). "Comparative Study of Various Genetic Distance Measures between Populations for the ABO Gene." Notulae Scientia Biologicae 2(4): 12-17.

de Castro Lobo, C. L., S. K. Ballas, *et al.* (2014). "Newborn screening program for hemoglobinopathies in Rio de Janeiro, Brazil." Pediatric blood & cancer 61(1): 34-39.

Desai, D. and H. Dhanani (2004). "Sickle cell disease: history and origin." The internet journal of hematology 1(2): 1540-2649.

Fawzi, Z., A. Al Hilali, *et al.* (2003). "Distribution of hemoglobinopathies and thalassemias in Qatari nationals seen at Hamad hospital in Qatar." QATAR MEDICAL JOURNAL 12(1): 20-24.

Feroze, M. and K. Aravindan (2000). "Sickle cell disease in Wayanad, Kerala: gene frequencies and disease characteristics." The National medical journal of India 14(5): 267-270.

Jain, R. (1985). "Sickle cell and thalassaemic genes in Libya." Transactions of the Royal Society of Tropical Medicine and Hygiene 79(1): 132-133.

LIVEX, O. A. (2009). "The feasibility of establishing sickle cell disease screening services at health centers in Uganda."

Lobitz, S., C. Frömmel, *et al.* (2014). "Incidence of sickle cell disease in an unselected cohort of neonates born in Berlin, Germany." European Journal of Human Genetics.

Luzzatto, L. (2012). "Sickle cell anaemia and malaria." Mediterranean journal of hematology and infectious diseases 4(1).

Marouf, R., T. D'souza, *et al.* (2002). "Hemoglobin electrophoresis and hemoglobinopathies in Kuwait." Medical Principles and Practice 11(1): 38-41.

Okwi, A. L., W. Byarugaba, *et al.* (2010). "An up-date on the prevalence of sickle cell trait in Eastern and Western Uganda." BMC Hematology 10(1): 5.

Patra, P. K., V. S. Chauhan, *et al.* (2011). "Screening for the sickle cell gene in Chhattisgarh state, India: an approach to a major public health problem." Journal of community genetics 2(3): 147-151.

Piel, F. B., A. P. Patil, *et al.* (2013). "Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates." The Lancet 381(9861): 142-151.

Roy, M., U. Datta, *et al.* (1990). "Genetic distance and gene diversity among ten endogamous groups in Chhattisgarh, Central India." International Journal of Anthropology 5(2): 109-115.

Savitt, T. L. and M. F. Goldberg (1989). "Herrick's 1910 case report of sickle cell anemia: the rest of the story." Jama 261(2): 266-271.

Schmaier, A. H. and H. M. Lazarus (2011). Concise Guide to Hematology, John Wiley & Sons.

Serjeant, B. E., M. Forbes, *et al.* (1974). "Screening cord bloods for detection of sickle cell disease in Jamaica." Clinical chemistry 20(6): 666-669.

Serjeant, G. R. (2001). "The emerging understanding of sickle cell disease." British journal of haematology 112(1): 3-18.

Steiner, C. A. and J. L. Miller (2006). "Sickle Cell Disease Patients in U.S. Hospitals, 2004: Statistical Brief #21."

Talukder, G. and A. Sharma (1979). Genetic Diseases in India-Present Status. Proc. Indian Natl. Sci. Acad. B.

Tshilolo, L., L. Aissi, *et al.* (2009). "Neonatal screening for sickle cell anaemia in the Democratic Republic of the Congo: experience from a pioneer project on 31 204 newborns." Journal of clinical pathology 62(1): 35-38.

Weatherall, D. (2011). "The inherited disorders of haemoglobin: an increasingly neglected global health burden." Indian J Med Res 134: 493-497.

White, J., M. Byrne, *et al.* (1986). "Red cell genetic abnormalities in Peninsular Arabs: sickle haemoglobin, G6PD deficiency, and alpha and beta thalassaemia." Journal of medical genetics 23(3): 245-251.

****