BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 11(7)

Allele frequency for Cystic fibrosis in Indians *vis-a-vis* global populations

Karnajit Kumar Bepari, Arup Kumar Malakar, Prosenjit Paul, Binata Halder & Supriyo Chakraborty*

Department of Biotechnology, Assam University, Silchar 788011, Assam, India; Supriyo Chakraborty – E mail: supriyoch_2008@rediffmail.com; *Corresponding author

Received June 02, 2015; Revised July 23, 2015; Accepted July 24, 2015; Published July 31, 2015

Abstract:
Cystic fibrosis (CF) is an autosomal recessive disease caused by mutations in the cystic fibrosis transmembrane conductance regulator gene. This gene encodes a protein involved in epithelial anion channel. Cystic fibrosis is the most common life-limiting genetic disorder in Caucasians; it also affects other ethnic groups like the Blacks and the Native Americans. Cystic fibrosis is considered to be rare among individuals from the Indian subcontinent. We analyzed a total of 29 world's populations for cystic fibrosis on the basis of gene frequency and heterozygosity. Among 29 countries Switzerland revealed the highest gene frequency and heterozygosity for CF (0.022, 0.043) whereas Japan recorded the lowest values (0.002, 0.004) followed by India (0.004, 0.008). Our analysis suggests that the prevalence of cystic fibrosis is very low in India.

Key words: Allele frequency, cystic fibrosis, India, other populations.

Background:
Genetic disease or disorder is the result of changes, or mutations, in an individual's DNA. Any mutation in the coding region of a gene may lead to the production of a protein which is no longer functional. In India, the occurrence of congenital abnormalities is up to 37.0 per 1000 births and for live births alone from 9.8 to 21.0 per 1000 **[1]**. Genetic diseases can be inherited if the mutation occurs in the germ cell of the body, as germ cells are involved in passing genetic information from one generation to next generation or from parents to offspring. Cystic fibrosis, an autosomal recessive disease, occurs when both the parents have mutated genes. Cystic fibrosis (CF) is the most common potentially lethal and life threatening genetic disorder among the white Caucasian population of Europe, North America and Australia. It is caused by mutation of the cystic fibrosis transmembrane conductance regulator (CFTR) gene **[2].** In UK approximately 1 in 2500 children are born with CF incidence **[3],** less common in African Americans (1 in 15 000) and Asian Americans (1 in 31 000) **[4].**

The CFTR gene translates a protein which transports sodium and chloride across cell membranes. Mutation of *CFTR* gene leads to too much salt deposit and not enough water within the cells. As a result, the mucus becomes thick, sticky and salty instead of sweaty. The thick, sticky mucus builds up in the respiratory and digestive passage ways, which gives rise to symptoms of the disease condition. The clinical features include pancreatic insufficiency, male infertility, meconium ileus in the newborn, and chronic lung infection with excessive inflammation, leading to progressive deterioration of lung function **[5].** The loss of lung function is the main cause of death in CF patients. Most current therapies treat the symptoms of the disease and have increased the median life expectancy for individuals with CF to ~39 years **[6].**

In 1989 the CFTR gene was identified, and located on the seventh human chromosome at the position 7ql3 **[7]**. More than 1,900 mutations have been identified in the CFTR gene (http://www.genet.sickkids.on.ca/cftr). The most common mutation is the ΔF508 that is seen in two thirds of all

cystic fibrosis cases all over the world **[8].** It is a three-nucleotide deletion at the 508th codon causing the deletion of a phenylalanine residue and subsequent defective intracellular processing of the *CFTR* protein which is an important chloride channel **[9].** In 1968 it was reported that cystic fibrosis was rare in India **[10]**. Recent studies on CF cases indicate that the disease is far more common in people of Indian origin than previously thought. CF in the migrant population from the Indian subcontinent in the USA is 1 in 40 000, and in the UK is 1 in 10 000–12 000 **[11].** A study on 955 cord blood samples suggested a carrier rate of the common mutation DF 508 (DF 508) as 0.4%, and incidence of CF in India is 1 in 40 000 **[12],** but its diagnosis often escapes in the majority of cases. In view of the above facts the present study was undertaken to estimate the allele frequency of cystic fibrosis gene in Indian *vis-a-vis* 28 other global populations.

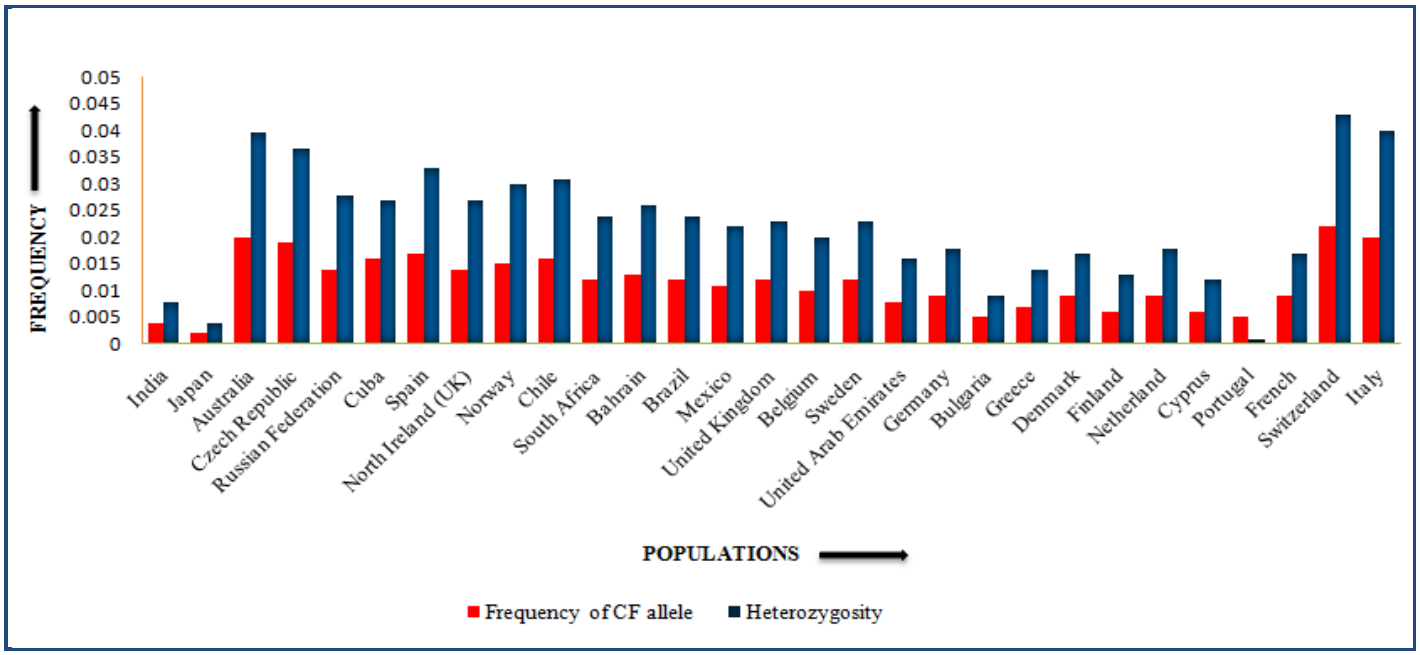

Figure 1: Comparison of gene frequency and heterozygosity for cystic fibrosis in India and 28 global populations

Methodology:

Data collection

The data were retrieved from the molecular genetics epidemiology of cystic fibrosis (MGECF) that were published by WHO since 1983. Most documents report joint meetings and workshops organized by WHO in association with various organizations devoted to CF, including the International Cystic Fibrosis (Mucoviscidosis) Association (ICFMA) now Cystic Fibrosis Worldwide, the Cystic Fibrosis Foundation (CFF), the European Cystic Fibrosis Thematic Network (ECFTN) and the European Cystic Fibrosis Society (ECFS). Some data are collected from the countries of European Union (EU) during 2004.

For the present study we have selected a total of 29 populations of different countries (including India) across different continents *i.e,* Europe, Australia, Latin America, Africa and Asia. The population data of cystic fibrosis were obtained from the published literature given in **Table 1 (see supplementary material) [13-18].**

Estimation of allele frequency

Assuming dominance model with two alleles for cystic fibrosis locus, the recessive allele frequency for different populations was estimated following Hedrick (2005) **[19]** as

$$q = \left(\frac{N_{22}}{N} \right)^{1/2}$$

Where, N_{22} = No. of recessive homozygotes
N = Total individuals

The dominant allele frequency (p) was estimated as

$$p = 1 - q$$

If p and q are the frequencies of a biallelic gene, then at Hardy-Weinberg equilibrium the frequencies of genotypes are given by

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

Where, p = the frequency of the dominant allele (represented here by A)
q = the frequency of the recessive allele (represented here by a)
p^2= frequency of AA (homozygous dominant) genotype
2pq= frequency of Aa (heterozygous) genotype
q^2= frequency of aa (homozygous recessive) genotype

Hardy-Weinberg heterozygosity of a population for a particular locus with *n* alleles can be estimated as

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

Where, p_i^2 = genotype frequency of observed homozygotes.

Heterozygosity is used as an index of genetic variation in a locus within a population and provides information about the occurrence of an allele upon segregation.

Results & Discussion:

Genetic study of the allele frequency of loci involved in genetic disorders plays a very crucial role in formulating research projects for follow-up detailed study and in identifying the thrust areas for therapeutic interventions by the governmental agencies. Based on allele frequency analysis it can be predicted how frequently the diseased condition will occur in common population of a country. The population size and incidence of cystic fibrosis in 29 countries are presented in **Table 1 (see supplementary material)**.

Assuming a digenic model with complete dominance and a recessive allele governing the cystic fibrosis condition, we analyzed 29 human populations for estimating the frequencies of normal/dominant (healthy condition) and recessive (diseased condition) alleles and the extent of heterozygosity for cystic fibrosis locus. Heterozygosity is used as an index for understanding the degree of genetic variation occurring in a diseased locus. Greater the heterozygosity, greater the chance of occurrence of the disease as recessive homozygotes upon Mendelian segregation and recombination in a population.

The gene frequency and heterozygosity for cystic fibrosis in 29 countries is presented in **Figure 1**. Earlier study on cystic fibrosis reported it be rare in India **[10]**. But subsequent studies reported that cystic fibrosis increases every year in the population. Keeping this point in mind, we estimated the frequency of the recessive allele for cystic fibrosis in 29 global populations including India from the genotypic frequencies. Our analysis revealed that Japan recorded the lowest allele frequency and heterozygosity for CF (0.002, 0.004) followed by India (0.004, 0.008).

The predicted incidence for CF among Asians in the United Kingdom (Indians/ Pakistani) is 1 in 10000 **[20]** and 1 in 40000 in the USA **[11].** But in India, the CF incidence is reported to be 1 in every 40000 to 100000 live births **[21].** Switzerland (0.022, 0.043), Italy (0.020, 0.040) and Australia (0.020, 0.039) record relatively higher frequency of the recessive allele and heterozygosity for CF **Table 2 (see supplementary material)**. It is evident from our analysis that Asian countries showed lower frequencies of the CF recessive allele compared to European countries like Switzerland and Italy.

Literature on the study of CF in Middle East countries is scanty. Our analysis revealed that the Middle East countries namely UAE (0.008, 0.016) and Bahrain (0.013, 0.026) had intermediate CF recessive allele frequency and heterozygosity. The global distribution pattern of the CF recessive allele frequency and heterozygosity (**Figure 1**) provides a clue that during the process of migration of human population from Africa to other continents, the frequency of CF recessive allele might have increased in the colder regions of the world due to natural selection. Till date there is no detailed study on the genetic and clinical profile of cystic fibrosis among Indian children **[22].** Thus, an in-depth study on CF across the global populations at molecular level could provide further insights in understanding the occurrence of CF cases across different populations all over the globe.

Conclusion:

The present study revealed that the recessive allele frequency and heterozygosity for cystic fibrosis are low in India *i.e.*, 0.004 and 0.008, respectively. But Japan recorded the lowest values for allele frequency and heterozygosity (*i.e.* 0.002, 0.004) whereas Switzerland had the highest recessive allele frequency and heterozygosity (*i.e.* 0.022, 0.043) for cystic fibrosis. India runs a relatively low risk for cystic fibrosis among twenty-nine global populations. Further research at molecular level may be initiated to understand the factors responsible for increased frequency of cystic fibrosis in European countries and decreased frequency in Asian countries.

Acknowledgement:

We are thankful to Assam University for providing the necessary facilities to undertake the research work. The work was not funded by any external funding agency.

References:

[1] Malakar AK *et al. IJSR* 2014: 2136
[2] Kraemer R e*t al. Respir Res.* 2006 **7**: 138 [PMID: 17137500]
[3] Dodge JA *et al. Arch Dis Child.* 1997 **77:** 493 [PMID: 9496181]
[4] Hamosh A *et al. J Pediatr. 1*998 **132**: 255 [PMID: 9506637]
[5] Zielenski J, *Respiration* 2000 **67:** 117 [PMID: 10773783]
[6] Ashlock MA & Olson ER, *Annu Rev Med*. 2011 **62**: 107 [PMID: 21226613]
[7] Kerem, B *et al. Science* 1989 **245:** 1073 [PMID: 2570460]
[8] Flume PA & Stenbit A, A*m J Med Sci.* 2008 **335**: 51 [PMID: 18195584]
[9] Morales MM *et al. Braz J Med Biol Res*. 1999 **32**: 1021 [PMID: 10454765]
[10] Mehta S *et al. Indian Pediatr.* 1968 **5**: 185 [PMID: 5681613]
[11] Powers CA *et al. Arch Pediatr Adolesc Med*. 1996 **150**: 554 [PMID: 8620244]
[12] Kapoor V *et al. J Cyst Fibros.* 2006 **5:** 43 [PMID: 16311077]
[13] Dodge JA *et al. Eur Respir J.* 2007 **29:** 522 [PMID: 17182652]
[14] Chung F *et al. Respir Med.* 2002 **96:** 963 [PMID: 12477209]
[15] Lucotte G *et al. Hum Biol. 1*995 **67**: 797 [PMID: 8543293]-
[16] Slieker MG *et al. Chest.* 2005 **128**: 2309 [PMID: 16236889]
[17] Yiallouros PK *et al. Clin Genet.* 2007 **71**: 290 [PMID: 17309655]
[18] Southern KW *et al. J Cyst Fibros.* 2007 **6**: 57 [PMID: 16870510]
[19] Hedrick PW & Goodnight C, *Evolution* 2005 **59**: 1633 [PMID: 16329237]
[20] Goodchild MC *et al. Arch Dis Child*. 1974 **49**: 739 [PMID: 4420998]
[21] Kabra SK *et al. Indian Pediatr*. 2003 **40**: 612 [PMID: 12881616]
[22] Kabra M *et al. Am J Med Genet*. 2000 **93:** 161 [PMID: 10869121]

Edited by P Kangueane
Citation: Bepari *et al.* Bioinformation 11(7): 348-352 (2015)

BIOINFORMATION

open access

Supplementary material:

Table 1: Population size and disease incidence of cystic fibrosis in Indian and 28 other populations

No	Country	Population Size	Disease Incidence	Source
1	India	70000	1	MGECF (1983)
2	Japan	225000	1	MGECF (1983)
3	Australia	2500	1	MGECF (1983)
4	Czech Republic	2833	1	MGECF (1983)
5	Russian Federation	4900	1	MGECF (1983)
6	Cuba	3900	1	MGECF (1983)
7	Spain	3500	1	MGECF (1983)
8	North Ireland (UK)	5350	1	MGECF (1983)
9	Norway	4500	1	MGECF (1983)
10	Chili	4000	1	MGECF (1983)
11	South Africa	7056	1	MGECF (1983)
12	Bahrain	5800	1	MGECF (1983)
13	Brazil	6902	1	MGECF (1983)
14	Mexico	8500	1	MGECF (1983)
15	United Kingdom	60271000	8284	[17]
16	Belgium	10348000	1065	[18]
17	Sweden	7300	1	MGECF (1983)
18	United Arab Emirates (UAE)	15876	1	MGECF (1983)
19	Germany	82425000	6835	[19]
20	Bulgaria	7518000	170	[18]
21	Greece	10648000	555	[19]
22	Denmark	54100	412	[19]
23	Finland	25000	1	MGECF (1983)
24	Netherlands	16318000	1275	[20]
25	Cyprus	776000	26	[21]
26	Portugal	10524000	285	CF foundation US -06
27	France	60424000	4533	[22]
28	Switzerland	2000	1	MGECF (1983)
29	Italy	2438	1	MGECF (1983)

Table 2: Estimates of allele frequency, genotype frequency and heterozygosity of CF in 29 populations

Country	Frequency of diseased allele	Frequency of normal allele	Genotype frequency of diseased condition	Genotype frequency of normal condition	Genotype frequency of observed homozygotes	Heterozygosity
India	0.004	0.996	0.000016	0.992	0.99201	0.008
Japan	0.002	0.998	0.000004	0.996	0.99600	0.004
Australia	0.020	0.98	0.00040	0.96	0.96040	0.039
Czech Republic	0.019	0.981	0.00036	0.963	0.96336	0.036
Russian Federation	0.014	0.986	0.00020	0.972	0.97220	0.028
Cuba	0.016	0.984	0.00026	0.972	0.97230	0.027
Spain	0.017	0.983	0.00029	0.967	0.96729	0.033
North Ireland (UK)	0.014	0.986	0.000 19	0.973	0.97319	0.027
Norway	0.015	0.985	0.00022	0.970	0.97022	0.030
Chile	0.016	0.984	0.00025	0.969	0.96925	0.031
South Africa	0.012	0.988	0.00014	0.976	0.97614	0.024
Bahrain	0.013	0.987	0.00017	0.974	0.97417	0.026
Brazil	0.012	0.988	0.00015	0.976	0.97615	0.024
Mexico	0.011	0.989	0.00012	0.978	0.97812	0.022
United Kingdom	0.012	0.988	0.00013	0.977	0.97713	0.023
Belgium	0.010	0.990	0.00010	0.980	0.98010	0.020
Sweden	0.012	0.988	0.00014	0.977	0.97714	0.023
United Arab Emirates (UAE)	0.008	0.992	0.00006	0.984	0.98406	0.016
Germany	0.009	0.991	0.00008	0.982	0.98208	0.018
Bulgaria	0.005	0.995	0.00002	0.991	0.99102	0.009

Greece	0.007	0.993	0.00005	0.986	0.98605	0.014
Denmark	0.009	0.991	0.00007	0.983	0.98307	0.017
Finland	0.006	0.994	0.00004	0.987	0.98704	0.013
Netherlands	0.009	0.991	0.00007	0.982	0.98207	0.018
Cyprus	0.006	0.994	0.00003	0.988	0.98803	0.012
Portugal	0.005	0.995	0.00002	0.990	0.99002	0.001
France	0.009	0.991	0.00007	0.983	0.98307	0.017
Switzerland	0.022	0.978	0.00050	0.956	0.95650	0.043
Italy	0.020	0.980	0.00040	0.960	0.96040	0.040