

REVIEW ARTICLE

THE CRADLE OF THE $\Delta F508$ MUTATIONDanish Saleheen^{1,2}, Philippe M. Frossard¹¹ Department of Biological and Biomedical Sciences, Aga Khan University, Stadium Road, Karachi, Pakistan.² Molecular Epidemiology Unit, Strangeways Research Labs, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

Cystic fibrosis (CF) is the most common autosomal recessive disorder caused due to mutation/s in the *CFTR* gene. The most common mutation in *CFTR* worldwide is $\Delta F508$ and cystic fibrosis genetic analysis consortium revealed that this mutation is responsible for approximately 66% of all CF chromosomes in the world. Studies looking at the DNA polymorphic haplotypes created by CF linked markers suggest that $\Delta F508$ has a single origin as this mutation has been found associated exclusively with one marker haplotype. Despite a high prevalence of this mutation in CF patients in northern parts of Europe, findings suggest that this mutation was not spread by Europeans but by a group that is speculated to have originated in the Middle East or a more eastern region in Asia (most likely subcontinent). Over here we have given a brief introduction to cystic fibrosis and classification of *CFTR* mutations and have further elaborated on the crucial issue about the spread of the $\Delta F508$ mutation. We have reviewed findings that give clues about the origin of this mutation from the Baluch ethnicity residing in Pakistan.

Keywords: CFTR, Delta F508, Mutation

Cystic fibrosis (CF) is the most common life-limiting autosomal recessive disorder; it affects nearly 1/2500 live births in Caucasians.¹ CF is a multi-system disease and can involve secretory cells, sinuses, lungs, pancreas, liver and the reproductive tract.¹ The disease has a "classical" appearance in 90% of cases and patients present with a triad of chronic obstructive pulmonary disease, exocrine pancreatic insufficiency and elevation of sodium and chloride concentrations in the sweat.¹⁻³ After extensive research in the field, in western countries, 60% of all CF patients have a diagnosis of the disease before their first birthday and 90% of all CF patients are diagnosed by 10 years of age. This has greatly affected the life expectancy of CF patients as compared to the past.¹

The first evidence of CF being a monogenic disorder came from the reviews of CF pedigrees in Italian church records in 1985.² Later on identification and cloning of the *CFTR* gene on chromosome 7 served as the basis for understanding the genetic aetiology of the disease.⁴⁻⁵ The gene causing CF spans 190 Kb of DNA and encodes a 1480 amino acid polypeptide named cystic fibrosis transmembrane regulator (*CFTR*).⁶ To date 1529 mutations have been identified in the gene.⁷ These mutations have a variety of effects at the cellular level. However they can be grouped into 6 categories⁸, in which they:

- i. block the synthesis of *CFTR*,
- ii. lead to defective processing
- iii. cause abnormal regulation of *CFTR* Cl channel
- iv. disrupt the normal conductive pathway of the channel
- v. lead to complete or partial production of *CFTR* protein
- vi. lead to defective regulation of other channels.

The most common mutation in *CFTR* worldwide is $\Delta F508$.¹ It is a class II mutation caused by deletion of three bases encoding a phenylalanine residue at position 508 within the nucleotide binding domain (NBD).⁹ Heterologous expression system analyses have shown that this mutation results in misfolding of the *CFTR* and hence mislocalization and defective processing of the mature protein.⁹ It was postulated that the mutation is severe enough to result in absence of any mature protein in the affected cells.¹⁰ Recent immuno-cytochemical studies of intestinal, respiratory and hepatobiliary epithelia, however, show that in patients homozygous for $\Delta F508$, a proportion of *CFTR* protein was shown to be targeted to the apical membranes.¹¹ The fact was lent support by *in vitro* studies done on $\Delta F508$ homozygous CF patients suggesting that at least some *CFTR* molecules reach the plasma membrane in these patients.¹² Genotype-phenotype correlations clearly indicate that this mutation is severe as far as pancreatic status is concerned but conclusions on its effect on lung disease vary between studies.⁸

Cystic fibrosis genetic analysis consortium revealed that $\Delta F508$ is responsible for approximately 66% of all CF chromosomes in the world.¹³ However its frequency varies from region to region.¹³ The highest frequency has been found in patients from Northern Europe, where it accounts for 75%-88% of all CF alleles.¹⁴ Indeed, the frequency reaches to a maximum of 100% in the isolated Faroe Islands of Denmark.¹⁵ Importantly, evaluation of studies that have been done on European populations reveal a clear-cut north-west to south-east gradient of $\Delta F508$ prevalence in CF patients showing a maximum prevalence of 88% in Denmark (excluding Faroe

islands) to a minimum of 24.5% in the Turkish population.¹³⁻¹⁵

Although studies investigating the origin of $\Delta F508$ have failed to identify any founder population but have established strong evidence in ruling out various potential regions (for eg. Faroe islands) where the mutation could have possibly originated.^{13,16-19} These DNA based studies normally utilize single nucleotide polymorphisms (SNP's) and several short tandem repeat polymorphisms which are also known as 'microsatellites' within a gene of interest to find and trace the origin of a mutation. By utilizing such a methodology, evolutionary origin of some CF causing mutations have been identified; for instance *CFTR* mutation G542X had a single origin in the ancient Phoenicians.²⁰ Similarly *CFTR* mutation G551D has been associated with Celtic tribes.²¹

SNP's or microsatellite markers by virtue of linkage disequilibrium can be used to create haplotypic maps. A haplotype containing a *CFTR* mutation or a SNP linked with a *CFTR* mutation displays a geographical gradient starting from a higher frequency in the founder population residing in a specific region to a lower frequency in a region where migration from the prior region has occurred lately. Based on this principle ancestry of a *CFTR* mutation G551D was revealed, as one particular haplotype containing this mutation remained unaltered from more than 170 generations and displayed variable frequencies in different regions.^{20,22}

Studies looking at the DNA polymorphic haplotypes created by CF linked markers suggest that $\Delta F508$ has a single origin.¹⁶ This mutation has been found associated exclusively with one marker haplotype. Morral *et al.* have suggested that the original time of this mutation ranges from 3000 years to 53000 years ago.¹⁶ Despite a high prevalence in northern parts of Europe, findings suggest that the mutation is more ancient and did not originate in Europe.¹⁶⁻¹⁹ Moreover, analyses of wild type chromosome containing *CFTR* gene of various European populations suggest that the general genetic background of these populations is different than the diseased *CFTR* chromosome and thus the origin of most frequent CF mutations should be non-European.¹⁷⁻¹⁸

Based on these findings, Mateu *et al.* investigated the origins of five common *CFTR* mutations including $\Delta F508$ in 949 unrelated autochthonous healthy individuals of various geographical origins and concluded that the haplotype containing $\Delta F508$ was present in very low frequencies in the European populations.¹⁹ However, the highest frequency was detected in one Middle Eastern population.¹⁹ Although Mateu *et al.* did not report the data on the different Middle Eastern ethnicities, a

comprehensive pattern of CF causing mutations in two Middle Eastern regions has been reported by Frossard *et al.*^{23,24} The data from Middle East show high frequencies of $\Delta F508$ in particular ethnicities. Despite the fact that data reported by Frossard *et al.* does not show the haplotype frequencies, it is now widely believed that $\Delta F508$ is more ancient and that it was not spread by Europeans but by a group that preceded them.^{16-19,21,23} This group is speculated to have originated in the "Middle East" or a more eastern region in Asia (most likely subcontinent).^{23,25} Thus it is important to study this mutation along with specific haplotypes in the CF patients of Middle East, Pakistani and Indian origin to discover more about its origins.

Studies on CF in indigenous subcontinental population are severely lacking. Two studies have been done in this regard looking at the prevalence of $\Delta F508$ prevalence in CF patients. Both studies were done on patients who were admitted to tertiary care hospitals. In both the studies the patients were diagnosed of CF by pilocarpine iontophoresis test initially. Study done by Bhutta ZA *et al* on Pakistani population was limited to very small sample size.²⁶ They evaluated 15 patients belonging to different ethnicities in Pakistan and found a 60% prevalence of this mutation in the Pakistani CF population. The other study done by Kabra SK *et al* looked at a patient population coming from different regions of India and Pakistan. $\Delta F508$ was identified in 19% chromosomes out of 290 tested.²⁷ It is surprising to note that $\Delta F508$ was highest (56%) in patients of Pakistani origin as compared to a 12% prevalence of this mutation in Indian CF patients.²⁷ Findings of both studies could be challenged on the fact that $\Delta F508$ is a severe mutation as far as pancreatic status is concerned and patients carrying this mutation are hence more likely to be hospitalized compared to those with milder mutations, which would confer a spuriously high prevalence of $\Delta F508$ in a hospital-based sample population. A 42% difference in prevalence between Pakistani and Indian CF populations, however, cannot be ignored.

It must be kept in mind that the 'Pakistani population' does not have a homogenous distribution and is composed of people belonging to different ethnicities. It is highly essential to look at the individual ethnic data while studying the genetic mutation spectrum of a disorder like CF. Studies looking at the $\Delta F508$ prevalence in CF patients of Pakistani origin living overseas have been variable and have looked at heterogeneous Pakistani populations rather than individual ethnicities. The most striking evidence of a high prevalence of $\Delta F508$ in a Pakistani ethnic group comes from the findings of Frossard *et al.*, who studied CF genetic mutations in the United Arab Emirates.²³ They investigated 17 unrelated families and found a 86% prevalence (14/16 alleles) of $\Delta F508$ in Baluchis

residing in UAE. Out of this Baluch population, all Pakistani Baluch CF patients carried $\Delta F508$ mutation (100% prevalence). These findings were subsequently confirmed in the Omani population, in which 100% of CF patients of Baluch origin were also $\Delta F508$ homozygotes.²⁴ Keeping in view the findings discussed earlier about a non-European origin of $\Delta F508$ and a high prevalence of $\Delta F508$ found in Pakistani Baluchis, it has been proposed that this mutation might have originally arisen in Baluchis. This hypothesis is analogous to the findings of Quaipe *et al.*²⁸ who studied the spectrum of beta-thalassemia in the UAE population and concluded that some specific mutations were introduced into UAE population by Baluchi immigrants.

The Baluch ethnicity currently resides in Iran, Pakistan (Baluchistan) and Afghanistan, and its people are known to have immigrated through successive waves to the Punjab, India and Gulf countries (Oman and the UAE). However, the actual geographical location of Baluchis refers only to the western province of Pakistan, Baluchistan. To prove a founding effect of $\Delta F508$ mutation in the Pakistani Baluch population, it is highly essential to study Baluchis of Pakistani province Baluchistan using appropriate haplotype analyses. This will answer the crucial question in the field of CF to know whether indeed 'Baluchistan' was the cradle of the $\Delta F508$ mutation.



Figure-1: Starred line represents approximate boundaries of the Baluch ethnicity. The arrow indicates the migrations that occurred historically from Baluchistan to Oman and the UAE.

REFERENCES

1. Cutting GR. Cystic fibrosis. In: Rimoin DL, Connor JM, Pyeritz RE, Korf BR. Principles and Practice of Medical Genetics. 4th ed. Harcourt publishers, 2002; p. 1561–606.
2. Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal screening for cystic fibrosis: addition of molecular

- diagnostics to increase specificity. *Biochem Med Metab Biol* 1993;49:200–11.
3. Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry Annual Report 2000. Bethesda: Cystic Fibrosis Foundation, 2001.
4. Riordan JR, Rommens JM, Kerem BS, Alon N, Rozmahel R, Grzelczak Z, *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989;245:1066–73.
5. Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, *et al.* Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* 1989;245: 1059–65.
6. Ellsworth RE, Jamison DC, Touchman JW, Chissole SL, Braden Maduro VV, Bouffard GG, *et al.* Comparative genomic sequence analysis of the human and mouse cystic fibrosis transmembrane conductance regulator genes. *Proc Natl Acad Sci U S A*. 2000;97(3):1172–7.
7. The Cystic Fibrosis Genetic Analysis Consortium. Cystic Fibrosis Mutation Data Base. <http://www.genet.sickkids.on.ca/cftr/> (accessed Feb 3, 2007)
8. Vankeerberghen A, Cuppens H, Cassiman JJ. The cystic fibrosis transmembrane conductance regulator: an intriguing protein with pleiotropic functions. *J Cystic Fibrosis* 2002;1:13–29.
9. Rowntree RK, Harris A. The phenotypic consequences of CFTR mutations. *Ann Hum Genet* 2003; 67:471–85.
10. Cheng, SH, Gregory RJ, Marshall J, Paul S, Souza, DW, White GA, O'Riordan CR and Smith AE. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* 1990; 63:827–34
11. Kartner, N, Augustinas O, Jensen TJ, Naismith AL and Riordan JR. Mislocalization of delta F508 CFTR in cystic fibrosis sweat gland. *Nat Genet* 1992;1:321–7.
12. Kalin N, Claass A, Sommer M, Puchelle E and Tummeler B. DeltaF508 CFTR protein expression in tissues from patients with cystic fibrosis. *J Clin Invest* 103:1379–89.
13. Bobadilla JL, Macek M, Fine JP and Farrell PM. Cystic Fibrosis: A worldwide analysis of CFTR mutations-correlation with incidence data and application to screening. *Human Mutation* 2002;19:575–606.
14. Estivill X, Bancells C, Ramos C, Biomed CF Mutation Analysis Consortium. 1997. Geographic distribution and regional origin of 272 cystic fibrosis mutations in European populations. *Hum Mutat* 10:135–54
15. Schwartz M, Sørensen N, Brandt NJ, Høgdall E, Holm T. High incidence of cystic fibrosis on the Faroe Islands: a molecular and genealogical study. *Hum Genet* 1995;95:703–6.
16. Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Gimenez J, Reis A, Varon-Mateeva R, Macek M Jr, Kalaydjieva L, *et al.* The origin of the major cystic fibrosis mutation (delta F508) in European populations. *Nat Genet* 1994;7(2):169–75.
17. Mateu E, Calafell F, Ramos MD, Casals T, Bertranpetit J. Can a place of origin of the main cystic fibrosis mutations be identified? *Am J Hum Genet* 2002;70:257–64.
18. Morral N, Dork T, Llevadot R, Dziadek V, Mercier B, Ferec C, Costes B, Girodon E, Zielenski J, Tsui LC, Tummeler B, Estivill X. Haplotype analysis of 94 cystic fibrosis mutations with seven polymorphic CFTR DNA markers. *Hum Mutat* 1996;8:149–59.
19. Mateu E, Calafell F, Lao O, Bonne-Tamir B, Kidd JR, Pakstis A, Kidd KK, Bertranpetit J. Worldwide genetic analysis of the CFTR region. *Am J Hum Genet* 2001;68:103–17.
20. Loirat F, Hazout S, Lucotte G. G542X as a probable Phoenician cystic fibrosis mutation. *Hum Biol.* 1997;69:419–25.
21. Macek M Jr, Macek M, Serre JL, Vavrova V, Burger J, Reis A, *et al.* Population study of CFTR gene mutations in Bohemia and Moravia: hypothesis on the historical spread of G551D and deltaF508 in Europe. *Am J Hum Genet* 1991;49:A474,A2703
22. Cashman SM, Patino A, Martinez A, Garcia-Delgado M, Miedzybrodzka Z, Schwarz M, *et al.* Identical intragenic microsatellite haplotype found in cystic fibrosis chromosomes

- bearing mutation G551D in Irish, English, Scottish, Breton and Czech patients. *Hum Hered.* 1995;45:6–12.
23. Frossard PM, Girodon E, Dawson KP, Ghanem N, Plassa F, Lestringant GG, *et al.* Identification of cystic fibrosis mutations in the United Arab Emirates. *Mutations in brief* no. 133. *Online Hum Mutat* 1998;1:412–3
 24. Frossard PM, Dawson KP, Das SJ, Alexander PC, Girodon E, Goossens M. Identification of cystic fibrosis mutations in Oman. *Clin Genet.* 2000;57:235–6.
 25. Dawson KP, Frossard PM. A hypothesis regarding the origin and spread of the cystic fibrosis mutation deltaF508. *QJM* 2000;93:313–5.
 26. Bhutta ZA, Moattar T, Shah U. Genetic analysis of cystic fibrosis in Pakistan: a preliminary report. *J Pak Med Assoc* 2000;50:217–9
 27. Kabra SK, Kabra M, Lodha R, Shastri S, Ghosh M, Pandey RM, *et al.* Clinical profile and frequency of delta F508 mutation in Indian children with cystic fibrosis. *Indian Pediatr* 2003;40:612–9.
 28. Quaiife R, al-Gazali L, Abbas S, Fitzgerald P, Fitches A, Valler D, *et al.* The spectrum of beta thalassaemia mutations in the UAE national population. *J Med Genet* 1994;31:59–61.
-

Address for Correspondence:

Dr. Danish Saleheen, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Worts Causeway, Cambridge CB2 8RN. Tel: 01223 741302, Fax: 01223 741339.

Email: ds436@medschl.cam.ac.uk