

## **Why does cystic fibrosis display the prevalence and distribution observed in human populations?**

**Andrew Mowat**

Department of Paediatric Surgery, Hull Royal Infirmary, Hull, UK.

### **Abstract**

**Cystic fibrosis is a recessive monogenetic condition which is devastating for the evolutionary fitness of the sufferer. The disorder has not been eradicated by human evolution and allele frequencies as high as (1:20) are seen in Caucasian populations. The potential explanations for the persistence of the disease are discussed including a high rate of genetic mutation, the founder effect and heterozygote advantage. Three firm conclusions are reached; the first is that the most influential factor in causing the high CF prevalence amongst peoples in European ancestry is heterozygote advantage to Tuberculosis. Secondly, the founder effect accounts for the additional amplification of allele frequency in emigrant communities. Finally, new mutations help to subsidies the mutation pool at a linear rate across all ethnicities.**

**Keywords:** Fibrosis, Tuberculosis, Heterozygote advantage.

*Accepted January 30, 2017*

### **Introduction**

Cystic fibrosis (OMIM 219700) is a recessive monogenetic condition which is devastating for the evolutionary fitness of the sufferer. The disorder has not been eradicated by human evolution and allele frequencies as high as (1:20) are seen in Caucasian populations. The potential explanations for the persistence of the disease are discussed including a high rate of genetic mutation, the founder effect, and heterozygote advantage. Three firm conclusions are reached; the first is that the most influential factor in causing the high CF prevalence amongst peoples in European ancestry is heterozygote advantage to Tuberculosis. Secondly, the founder effect accounts for the additional amplification of allele frequency in emigrant communities. Finally, new mutations help to subsidies the mutation pool at a linear rate across all ethnicities.

There is good understanding as to why humans suffer from bacterial and viral infection. We are in an 'evolutionary arms race' between their virulence factors and our immune systems. Even artificial interventions such as antibiotics are now encountering resistant strains of bacteria, notably MRSA. It is, however, a more open question as to why genetic disorders, which are devastating for the evolutionary fitness of the sufferer, persist. As this review first evaluates arguments to explain the high prevalence of the inherited genetic condition cystic fibrosis (CF) and second considers why CF shows the global geographic distribution observed in the table below (Table 1).

CF is a true recessive condition. Sufferers always possess two mutated alleles and such homozygotes always show complete penetrance [1]. However, clinical presentation is not uniform. 'Mild' mutations have been identified which confer a less severe phenotype with pancreatic exocrine sufficiency providing a useful differentiator [2]. Modifier genes such as alpha-1-antitrypsin [3] and environmental factors such as air pollution also influence disease progression. Nevertheless, the question of CF disease prevalence must lead to a discussion of the factors which maintain CF alleles within the gene pool. Many of the arguments put forward to explain this allele frequency stand up to little scrutiny and are consequently given minimal attention. Other ideas possess a substantial body of supporting evidence although no one solution has been able to establish itself as scientific consensus.

The first factor considered is genetic mutation. CF alleles are continuously being brought into the population by mutation. If the rate of mutation produces disease alleles at the same rate at which they are lost through natural selection then the disease prevalence will remain constant, as is the case in CF. The causation factor of high mutation rate is sufficient to explain other genetic conditions such as achondroplasia [4] and Duchenne muscular dystrophy. Assuming that CF is in equilibrium in the UK population the required mutation rate can be calculated (Figure 1).

**Table 1.** Estimated CF gene (carrier) frequency in various populations. Adapted from Klinger [32] Carrier frequencies estimated using the Hardy-Weinberg formula

Geographic Location	Estimated Frequency of CF in New-borns	Estimated Heterozygote Frequency
Brittany	1:400	1:9.7
Ohio (US Amish)	1:625	1:12.5
Southwest Africa (Dutch)	1:676	1:13
United Kingdom	1:2,000	1:20-25
Australia	1:2,500	1:23-25
USA	1:2,500	1:23-28
France	1:2,500	1:23-27
Czechoslovakia	1:2,700	1:26
Germany	1:3,400	1:29
Pakistani (in England)	1:10,000	1:50
Blacks (residing in US)	1:29,000	1:82
Oriental (residing in US)	1:90,000	1:150

At equilibrium in recessive conditions  $sq^2 = \mu(1-q^2)$

Where:

(s) = the coefficient of selection which is the chance of reproductive failure in the disease sufferer. In CF sufferers are infertile and for the vast majority of human evolution have perished prior to reaching reproductive age therefore  $s=1$ .

( $q^2$ ) = the birth rate of sufferers, in the UK  $q^2=1/2000$ .

( $\mu$ ) = the mutation rate per gene per generation.

Solving for  $\mu$  gives the value of  $5 \times 10^{-4}$

**Figure 1.** The calculation of required mutation rate in order to explain CF allele frequency by mutation rate alone. Adapted from Human Molecular Genetics 3 Chapter 4

This value is exceptionally high. This in itself does not render the mutation rate theory impossible but when combined with haplotype data this factor can be considered implausible. If the mutation rate was as high as suggested by this formula, then one would expect to see new mutations constantly appearing on a wide range of genetic backgrounds (haplotypes). Instead a European wide survey was that 92% of CF chromosomes are associated with haplotype B, which was defined by the presence of two extragenic polymorphisms (XV2C and HM19) [5]. The theory is further contradicted by the nature of mutations leading to CF. Although 1400 mutations have been catalogued 70% of all CF chromosomes carry the same triple base loss leading to the deletion of phenylalanine ( $\Delta F508$ ) in the protein product. If mutations rates were responsible for the high frequency of alleles a more heterogeneous picture would be expected. The possibility of mutation rate would be more plausible if it were accompanied by genetic heterogeneity; the concept that mutations at multiple different loci all produce the same phenotypic end point. Until 1987 this idea was the widely accepted explanation of CF prevalence.

However, it has since been discredited by two separate lines of inquiry. The first of these comes from the study of consanguineous marriages. Romeo looked at CF and consanguineous marriage rates in Italy and found that the birth-rates of CF were consistent with those predicted by Dahlberg’s formula, using the assumption of genetic homogeneity [6]. The second line of evidence comes from data showing strong linkage disequilibrium between the CF locus and restriction fragment length polymorphisms (RFLP) [7]. KM.19 the RFLP nearest the CF gene pointed to the existence of a single CF gene located at 7q31. This has subsequently been mapped through the use of complementary DNA [8]. To summarise, although mutations were initially necessary for the creation of CF alleles mutations of new alleles plays little role in maintenance of CF alleles at such a high frequency. As for genetic heterogeneity, there has never been any need to invoke the presence of a second locus contributing to CF therefore this theory can be dismissed.

The next hypothesis to be considered is the ‘founder effect’. The idea that when human population numbers

go through narrow bottlenecks there is a resultant loss of genetic variation. When the population re-establishes itself any 'alleles' present in high numbers amongst the 'founders' will be present at much higher frequencies than prior to the bottleneck. The theory relies on the premise that the genetic makeup of the founders will be significantly different to the population from which they came. Accordingly, it is more pronounced with smaller founder groups. As the number of founders increases the likelihood of great disparity between the gene pools becomes reduced. The effect has been used to explain phenomena such as the total colour blindness on the island of Pingelap [9]. One advantage this theory has over mutation rate and genetic heterogeneity is that it provides a potential explanation of the non-uniform geographic distribution of CF observed in Table 1. For example, the high CF prevalence in Caucasian populations could be explained by a high proportion of the first European settlers being heterozygous for the CF allele. However, known socio-geographic factors argue against its influence. Europe was populated by constant waves of immigration, and at no point did the ancestors of the Caucasian race go through a bottleneck tight enough to explain the disparity in allele frequency observed. This does not render the founder effect totally redundant. The effect can be used to explain the extremely high prevalence of the allele seen in the Ohio Amish populations (1:12.5) [10]. The Amish have a common Swiss-German ancestry who emigrated in the seventeenth century, leaving too short a period of time for different selective pressures to be of any significance. In this case the population bottleneck is comfortably small enough to explain the difference between the original and founder populations. Though, in the global context of CF distribution it would seem that the 'founder effect' is no more than a partial solution.

The final factor to be considered is that of heterozygote advantage. That is the idea that CF heterozygotes have or had an advantage over homozygotes for the 'normal'

allele. This theory becomes more appealing when some mathematical analysis is used to reveal how marginal such an advantage would have to be (Figure 2).

This means that even without the creation of any new CF allele by mutation that the number of alleles will remain at equilibrium if heterozygotes have, at the mean, 2.3% more surviving children than 'normal' homozygotes. Such an advantage could be conferred in two ways. The first is if the heterozygote has a reproductive advantage. The second is if possessing a single mutated allele produces total or partial immunity to an infectious disease.

Two groups have published findings reporting that CF heterozygotes parents produced more children than controls. Danks et al. [11] first published data in 1965 reporting the phenomena in Australia. This was reinforced by Knudson two years later who released extremely significant data ( $P < 0.01$ ) showing that relationships between CF heterozygotes produced 4.34 live offspring whilst controls produced 3.43 [12]. If the numbers had proved reflective of the population, they would have explained the maintenance of CF allele frequency. However, it was later shown that the results of the two groups were influenced by 'bias of ascertain'. Specifically, larger families were more likely to provide the cystic fibrosis cases whose parents were used as the heterozygous cohort. A further study was conducted in which 143 grandparent couples of CF cases in Utah were contrasted with 20 control couples randomly selected from the Utah Genealogical Database [13]. In this case ascertainment correction was applied to the data which subsequently showed no significant difference. Jorder and Lathrop applied the same ascertainment correction retrospectively to the results of Danks et al. [11] and Hallet et al. [12] eradicating any significant difference between the data. More recent data has come from the Hutteries of South Dakota, a genetic isolate with a relatively high CF carrier frequency. A population wide screen has been conducted on the only two mutations present in the population which

At equilibrium in a population  $p/q = s_2/s_1$

Where:

(q) = the proportion of CF alleles within the population as  $q^2$  (the disease frequency) is **1/2000** in newborns in the UK working back gives  **$q = 0.022$** .

(p) = the proportion of non-CF alleles which is  **$1 - q = 0.978$**

(s<sub>1</sub>) and (s<sub>2</sub>) are the coefficients of selection against 'normal' homozygotes and CF sufferers respectively therefore (s<sub>2</sub>) = **1** (see Fig 2 for clarification)

Solving for (s<sub>1</sub>) gives **0.023**.

**Figure 2.** Calculation of selective advantage required by the heterozygote in order to maintain allele frequency. Adapted from *Human Molecular Genetics 3 Chapter 4*

sought to establish an association between carrier status and fertility. The work provided no positive evidence of non-random transmission of mutations or skewed sex ratios in children of carrier parents [14]. It would seem especially unlikely that heterozygotes achieve a reproductive advantage since CF sufferers are rendered infertile.

The possibility of ‘meiotic drive’ has been invoked [15]. This mechanism is widespread in *Drosophila*. It allows a heterozygote allele to force itself into the next generation without producing a direct selective advantage by ‘cheating’ meiotic segregation and being present in greater than 50% of gametes. However, results from Kitzis et al. [16] using data collected from Europe and North America show no preferential transmission through the paternal line as was originally claimed [17]. To close, there is no strong evidence to suggest that reproductive advantage plays any role in CF allele maintenance. This leaves the hypothesis that CF heterozygotes have at least partial immunity to an infectious agent.

It has been established that CF carriers, display phenotypic differences non-carriers. A recent search of the medical literature between 1966 and 2015, aggregated data from 15 separate studies. The meta-analysis found that rates of asthma were 60% higher in carriers than non-carriers [18].

An example of such a relationship is known to exist between a recessive genetic condition and an infectious condition; namely sickle cell anaemia and malaria. This association has been firmly established as scientific consensus since its proposal by a group at Oxford University [19].

In order for any infectious agent to achieve such a status with respect to CF it must conform to three criteria. First, there must be a theoretical mechanism by which the advantage is conferred. Second, the infectious agent must have provided a strong selective pressure in the Caucasian population in the time frame in which the allele frequency increased. Furthermore, the infectious agent should be significantly less prevalent in populations in which the CF allele is much lower, for example Orientals (Table 1). Finally, there should be clinical evidence of the difference in response to the agent between ‘normal’ and

heterozygotes. Three candidate agents have emerged and will be considered here:

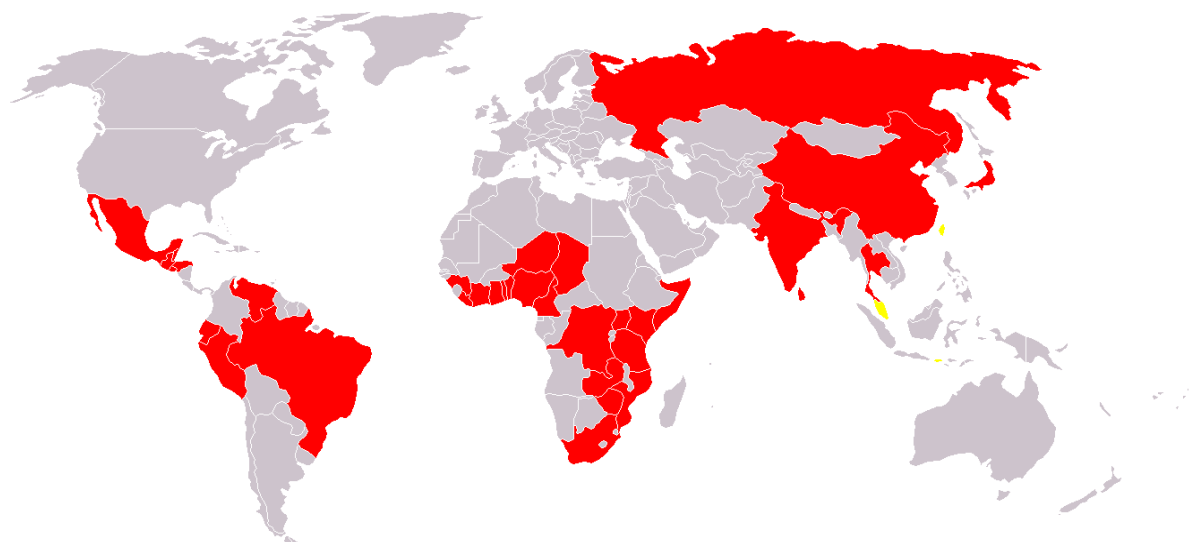
- 1.) Cholera combined with other diarrheal bacteria such as *E. coli*
- 2.) Typhoid
- 3.) Tuberculosis

Cholera has been a strong candidate since its proposal by a variety of sources. The strength of the hypothesis derived from the mechanism of pathogenicity of *Vibrio cholerae*. The bacteria release a potent exotoxin which irreversibly activates the G-protein subunits in the intestinal epithelium. This induces a series of downstream events culminating in the increase in cAMP<sub>o</sub> and chloride channel opening by protein kinase A. The clinical sequelae are watery diarrhoea and a 50% chance of mortality if left untreated. As most CFTR mutations, notably ΔF508, produce a protein product which never reaches the apical membrane [20] it has been suggested that presumably have half the number of chloride channels would have a greater chance of survival in acute cholera infection. This hypothesis relies on the number of chloride channels being the limiting factor for chloride secretion. Scientific inquiry into this issue has been highly inconsistent as can be seen in Table 2 [20-35].

Hogenauer et al. [21] make three alternative suggestions as to the nature of the limiting factor in chloride secretion. The most convincing of these is that the level of chloride secretion is restricted by the activity of the basolateral transporter critical in establishing the electrochemical gradient necessary for chloride secretion. If any of these theories prove to be valid it will negate if not totally dismiss the CF cholera heterozygote hypothesis. Another problem with the cholera hypothesis is that global distribution of the disease does not match CF distribution in the same way that malaria shadows sickle cell distribution. The original cholera reservoir existed in the Ganges River in India and was only spread to Europe via trade routes in the nineteenth century. Of the seven major outbreaks thereafter only five reached Europe and most occurred on the Indian subcontinent an area in which cholera is still endemic (Figure 3).

**Table 2.** Table documenting the evidence for and against the non-uniform secretion between heterozygotes and ‘normal’

	<b>Murine Models</b>	<b>Human Models</b>
<b>Supportive</b>	Gabriel et al. [26] produced supportive evidence showing a linear relationship between functioning allele number and secretion levels, i.e., CF mice did not produce any secretion. Heterozygotes produced 50% that of normal when exposed to cholera toxin.	Behm et al. [34] produced indirect evidence. They showed that heterozygotes have a 35% reduction in sweat volume when perspiration was induced by stimulation with cholinergic and beta-adrenergic agonists. As sweating requires secretion through the small chloride channels this is supportive.
<b>Unsupportive</b>	Cuthbert et al. [35] found no significant difference in secretion levels between normal and heterozygotes when stimulated with a range of secretagogues such as cholera toxin and isoprenaline.	Hogenauer et al. [21] found no difference in jejunum secretion levels between heterozygotes and controls when a prostaglandin analogue was used as a chloride secretagogue.



**Figure 3.** Map showing the prevalence of cholera in the modern world. Countries with endemic cholera levels are shown in red whilst those with sporadic outbreaks are yellow in colour. The image was adapted from <http://geo.arc.nava.gov/sge/health/diseases>

However, CF allele frequency is low in this area 1:50 (Table 1). Proponents of the cholera hypothesis have countered by suggesting that the decreased sweat levels in heterozygotes negate the cholera advantage in hot climates, a highly speculative suggestion. A final obstacle is the recently published work of Poolman and Galvani [22] who analysed British mortality records from 1861 to 1870 to determine the mortality rates from the three major candidates. The conclusion was that cholera, even working in combination with other diarrheal inducers such as *E.coli* could not produce the allele frequency observed. Even if being a heterozygote for CF conferred total resistance, an equilibrium incidence of 1:10,000 would be achieved, significantly lower than that observed (1:2000). This evidence is given particular strength by the fact that the data reviewed related the period of the fourth cholera outbreak in which death rates from cholera were anomalously high. In summary, cholera has a neat theoretical mechanism by which CF heterozygosity could impart immunity. However, evidence for the mechanism working *in vivo* is weak. This combined with its failure to offer a satisfactory explanation of CF distribution and new evidence suggesting it did not kill enough people mean that the cholera heterozygote hypothesis is frail. It has been given an easy ride in the scientific community perhaps through lack of a suitable alternative.

The second candidate disease is typhoid. The theoretical advantage is based on the work of Pier et al. [23] who showed that the normal CFTR was used by the disease causing bacteria *Salmonella typhi* for entry into epithelial cells. The group found that human epithelial cells expressing wild-type CFTR ingested significantly greater amounts of *S. typhi* when compared to heterozygotes expressing the most common CF mutation,  $\Delta F508$ . The work of van de Vosse et al. [24] was also supportive showing that in typhoid endemic Indonesia the  $\Delta F508$

mutation was not present in any of the 775 typhoid patients sampled. However, the typhoid hypothesis suffers from the same difficulties as cholera. The geographic dispersion of typhoid presence in the modern day does not show any correlation with CF frequency. Furthermore the work of Poolman and Galvani [22] showed that death rates in the envelope analysed could only produce a CF disease prevalence of 1:7600 live births, even if resistance to the bacteria was total amongst heterozygotes.

The final candidate disease is Tuberculosis (TB). Here, the weakness in this hypothesis is that the theoretical advantage of carrying a CF allele to tuberculosis is unclear. Meindl suggested that fibroblasts of heterozygotes produce an excess of hyaluronic acid, a substance believed to have a role in defence against pathogens such as *Mycobacteria tuberculosis* [25]. However, this is highly speculative and based on little experimental observation. The strength of the TB hypothesis comes from its compatibility with the historical-geographic perspective. TB was seen as ‘the white plague’ and offers an explanation of observed CF distribution in a way that cholera and typhoid do not. As humans moved north hunter-gatherers became farmers and cattle were domesticated. This provided two separate advantages for TB to spread. The first is the intimate human contact produced by close knit communities [26]. The second is that many scientists suspect that modern TB originated from bovine species around the time of domestication [27]. This produces a causative link between Caucasian people and CF advantage:

The work of Poolman and Galvani [22] is supportive of the TB hypothesis. They show that TB easily killed enough people to produce the observed allele frequency, indeed the death rates from TB were so pronounced that a resistance rate of only 13% amongst heterozygotes would explain the current allele frequency. This is the beauty of the TB hypothesis; TB was so prevalent that only a mild

selective advantage would produce rapid allele frequency increases. Such a marginal physiological difference may not necessarily have been documented by the scientific community especially as little investigation into this area has been carried out. Poolman additionally modelled the CF incidence in India showing that if tuberculosis had been 70% of that in the UK it would have produced the observed Indian CF frequency, a figure which is plausible historically. Evidence also comes from Ashkenazi-Jewish populations. This group, of eastern European origin, show a high rate of the fat storage condition Tay-Sachs, believed to be due to heterozygote advantage to TB. The Ashkenazi-Jews were concentrated into towns in a more intense manner than the majority of Europeans and show a higher CF prevalence (1:3000) than other Jewish populations, for example Iranian Jews (1:39000). Furthermore it is clear that as European emigrants colonised much of the globe and a major factor in their success was their immunity to disease which had their origin in livestock. For example, in the 1880s white settlers built the Canadian Pacific Railroad through the Saskatchewan province bringing them into contact with certain Native Americans tribes for the first time. In the subsequent years the Natives proceeded to die of TB at a rate of 9% per annum. In conclusion, TB superficially appears a weak candidate for heterozygote advantage as no neat theoretical advantage has been suggested. However, further investigation reveals a strong body of evidence dwarfing that of typhoid and cholera.

More recently, there have been different suggestions about how heterozygote advantage could manifest itself. Groups have suggested that rather than giving an advantage against an infective agent, being a heterozygote could be advantageous against environmental stressors. Two hypotheses have been suggested and will be discussed. The first is that heterozygotes have an increased resilience to lactose induced diarrhoea. This idea utilises the same theoretical hypothesis as cholera. Heterozygotes have only half the number CFTR channels in their gastrointestinal tract therefore would have reduced diarrhoea, if presented with an identical stressor. In the lactose intolerant, consumption of lactose induces a constellation of gastrointestinal symptoms including abdominal pain and diarrhoea. These typically start between 30 mins and 2 h, after the consumption of milk [28]. Before the domestication of cattle, lactose intolerance, was ubiquitous amongst human populations. As cow's milk became a staple component of human's diet diarrhoea would have been an inevitable side effect. Dehydration would have been a common cause of a common of death, and an increased ability to tolerate dehydration, an evolutionary advantage. Modiano et al. [29] show a clear correlation between populations with a high CF allele frequency and high levels of the dominant P allele of the lactase, LCT, gene, which results in lactose tolerance. Like the TB hypothesis, the theory, neatly fits the geographical distribution of cystic fibrosis. Northern European populations were the first domestic cattle, and have had the longest period of exposure to

lactose. Cystic fibrosis remains rare in populations, who have high levels of lactose intolerance, such as China [30]. However, the theory has similar weaknesses to cholera induced diarrhoea, namely, that it is unclear that being a heterozygote for mutant CFTR, would have any protective effect on diarrhoea, induced fluid loss. As the number of CFTR channels is unlikely to be the limiting factor in fluid loss into the gastrointestinal tract.

The second environmental factor which has been proposed is that carriers have had a respiratory advantage during evolution. Airway fluid is lost through evaporation, particularly when breathing cold, dry air. When the CF allele frequency was increasing in frequency, in Europe, the climate was cold and dust laden. Vladimir et al. suggests that individuals with one CF mutation due to slower fluid reabsorption, and therefore may have had better clearance of inhaled dust [31]. However, the theory is reliant on any advantage to a dust laden environment being exclusively felt in Europe. This is unlikely, and there is no therefore no clear mechanism why such an advantage would lead to rise of allele numbers exclusively in northern European populations (Figure 4).

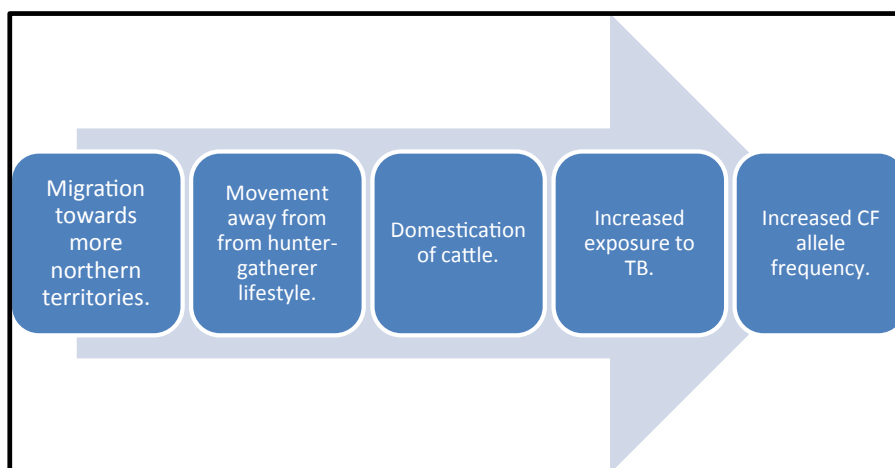
As there are no established examples of other genetic conditions being caused by a heterozygote advantage to similar environmental conditions, such hypotheses, represent, scientific speculation and are difficult to verify.

In final contemplation, three conclusions can be drawn:

1. The most influential factor in causing the high CF prevalence amongst peoples of European ancestry is the heterozygote advantage to TB.
2. The 'founder effect' accounts for the additional amplification of allele frequency in emigrant communities such as the Ohio Amish.
3. New mutations help to subsidise the mutation pool at a linear rate across all ethnicities.

It is important to note that none of the factors considered are mutually exclusive. The observed prevalence and dispersion could be explained by any combination working in conjunction with one another.

In terms of the future of this debate, academic questions such as this one are often given low priority in modern science as there is little to be gained clinically through their solution. However, if significant scientific resources are directed at the question then the first conclusion may be confirmed by culturing the bacteria *Mycobacterium tuberculosis* in conditions mimicking those of the CF heterozygote pulmonary epithelium. If such experiments show that the bacteria displays inferior growth when compared to growth in 'normal' pulmonary epithelium then TB will possess all three aspects of the triad required for any infectious agent to explain the high prevalence of a genetic condition through heterozygote advantage and may become scientific fact.



**Figure 4.** Flow chart displaying the causation pattern between Caucasian populations and increased CF allele frequency

## References

- Sosnay PR, Raraigh KS, Gibson RL. Molecular genetics of cystic fibrosis transmembrane conductance regulator: Genotype and phenotype. *Pediatr Clin North Am* 2016; 63: 585-598.
- Sheppard DN, Rich DS, Gregory OR, et al. Mutations in CFTR associated with mild disease-form Cl<sup>-</sup>channels with altered pore properties. *Nature* 1993; 362, 160-164.
- Meyer P, Braun A, Roscher AA. Analysis of the two common alpha-1-antitrypsin deficiency alleles PiMS and PiMZ as modifiers of *Pseudomonas aeruginosa* in cystic fibrosis. *Clin Genet* 2002; 62: 325-327.
- Arnheim N, Calabrese P. Germline stem cell competition, mutation hot spots, genetic disorders and older fathers. *Annu Rev Genomics Hum Genet* 2016; 17: 219-243.
- Sereth H, Shoshani T, Bashan, et al. Extended haplotype analysis of cystic fibrosis mutations and its implications for the selective advantage hypothesis. *Hum Genet* 1993; 92: 289-295.
- Dahlberg G. *Mathematical methods for population genetics*. Karger Basel 1947; 61-67.
- Estivilli X, Farall M, Scouler PJ, et al. candidate for the cystic fibrosis locus isolated by selection for methylation islands. *Nature* 1987; 326: 849-845.
- Riordan JR, Romman JM, Kerem B, et al. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* 1989; 16: 1066-1073.
- Gardner JC, Michaelides M, Hardcastle AJ. Cone opsins, colour blindness and cone dystrophy: Genotype-phenotype correlations: The new millennium. *S Afr Med J* 2016; 106: 75-78.
- Da Silva Martins R, Fonseca ACP, Acosta FES, et al. Severe phenotype in an apparent homozygosity caused by a large deletion in the CFTR gene: A case report. *BMC Res Notes* 2014; 7: 583.
- Danks DM, Allan J, Anderson CM. A genetic study of fibrocystic disease of the pancreas. *Ann Human Genet* 1965; 28: 323.
- Hallett WY, Knudson AG, Massey FJ. Absence of detrimental effect of the carrier state for the cystic fibrosis gene. *Am Rev Respir Dis* 1965; 92: 714.
- Jorder LB, Lathrop GM. A test of the heterozygote advantage hypothesis in cystic fibrosis carriers. *Hum Genet* 1988; 42: 808-815.
- Romeo IG, Ober C. CFTR mutations and reproductive outcomes in a population isolate. *Hum Genet* 2008; 122: 583-588.
- Romeo G, Devoto M, Galiotta LJV. Why is the cystic fibrosis gene so frequent? *Hum Genet* 1989; 84: 1-5.
- Kitzis A, Chomel JC, Kaplan JC, et al. Unusual segregation of cystic fibrosis alleles to males. *Nature* 1988; 333: 215.
- Vidaud M, Kitzis A, Ferec C, et al. Confirmation of linkage disequilibrium between haplotype B and cystic fibrosis allele in the French population. *Hum Genet* 1989; 81: 183-184.
- Nielsen AO, Qayum S, Bouchelouche PN, et al. Risk of asthma in heterozygous carriers for cystic fibrosis: A meta-analysis. *J Cyst Fibros* 2016; 15: 563-567.
- Allison AC. Protection afforded by sickle-cell trait against subterian induced infection. *Br Med J* 1954; 6: 290-294.
- Sheppard DN, Welsh MJ. Structure and function of the CFTR chloride channel. *Physiol Rev* 1999; 79: 23-45.
- Hogenauer C, Santance CA, Porter JL, et al. Active intestinal chloride secretion in human carriers of cystic fibrosis mutation: an evaluation of the hypotheses that heterozygotes have submucosal active intestinal chloride secretion. *Ann J Hum Genet* 2000; 67: 1422-1427.
- Poolman EM, Galvani AP. Evaluating candidate agents of selective pressure for cystic fibrosis. *J R Soc Interface* 2007; 4: 91-98.
- Pier GB, Grout M, Zaidi T, et al. *Salmonella typhi* uses CFTR to enter intertinal epithelial cells. *Nature* 1998; 393: 79-82.

24. van de Vosse E, Ali S, Visser AW, et al. Susceptibility to typhoid fever is associated with a polymorphism in the cystic fibrosis transmembrane conductance regulator (CFTR). *J Hum Genet* 2005; 118: 138-140.
25. Meindl RS. Hypothesis: A selective advantage for cystic fibrosis in heterozygotes? *J Med Genet* 1987; 74: 39-45
26. Porter R. *The GREATEST BENEFIT TO MANKIND*. Fontana Press. 1997.
27. Stead W, Eisenach JK, Cave M, et al. When did Mycobacteria tuberculosis infection first occur in the new world? The important question with public health implications. *Am J Res Med* 1995; 151: 1267-1268.
28. Ingram CJ, Mulcare CA, Itan Y, et al. Lactose digestion and the evolutionary genetics of lactase persistence. *Hum Genet* 2009; 124: 579-591.
29. Modiano G, Ciminelli BM, Pignatti PF. Cystic fibrosis: Cystic fibrosis and lactase persistence: A possible correlation. *Eur J Hum Genet* 2007; 15: 255.
30. Yang Y, He M, Cui H, et al. The prevalence of lactase deficiency and lactose intolerance in Chinese children of different ages. *Chinese Medical Journal* 2000; 113: 1129-1132.
31. Vladimir B, Tomašević B, Kurbel S. Hypothesis: Possible respiratory advantages for heterozygote carriers of cystic fibrosis linked mutations during dusty climate of last glaciation. *J Theor Biol* 2014; 363: 164-168.
32. Klinger KW. Genetics of cystic fibrosis. *Sem Respir Med* 1994; 6: 243.
33. Gabriel SE, Brigman KN, Koller BH, et al. Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. *Science* 1994; 266: 107-109.
34. Behm JK, Hagiarama G, Lewiston NJ, et al. Hyposecretion of  $\beta$  adrenergically induced sweating in cystic fibrosis heterozygotes. *Pediatr Res* 1987; 22: 271-276.
35. Cuthbert AW, Halstead J, Ratcliff R, et al. The genetic advantage hypothesis in cystic fibrosis heterozygotes: A murine study. *Physiol* 1995; 482: 449-454.

**Correspondence to:**

Andrew Mowat,  
Department of Paediatric Surgery,  
Hull Royal Infirmary, Hull,  
UK.  
Tel: 07903910917  
E-mail: andrewmowat@live.com