Cystic Fibrosis on the African Continent

Running Title: CF in Africa

Cheryl Stewart (PhD, BSc) and Michael S. Pepper (MD, PhD)

Department of Immunology and Institute for Cellular and Molecular Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Author for correspondence

Prof. Michael S. Pepper

Dept. of Immunology

Faculty of Health Sciences

University of Pretoria

P.O. Box 2034

0001 Pretoria

South Africa

Tel: +27 12 319 2179

Fax: +27 12 319 2946

E-mail: michael.pepper@up.ac.za

ABSTRACT

Cystic fibrosis (CF; OMIM #219700) is a life-shortening and costly autosomal recessive disease which has been most extensively studied in individuals of Caucasian descent. There is ample evidence however that it also affects other ethnicities. In Africa there are several reports of CF, but there has been no concerted effort towards establishing the molecular epidemiology of this disease on the continent which is the first step towards outlining a public health strategy to effectively address the needs of these patients. A literature search has revealed reports from only 12 of the 54 African states on the molecular analysis of the mutations present in suspected CF patients, resulting in the identification of 79 mutations. Based on previous functional investigations, 38 of these cause CF, 10 are of varying clinical consequence, five have no associated evidence as to whether or not they cause CF, four are synonymous, five were novel, and 21 are unique to Africa. We propose that CF be more thoroughly investigated on the continent to ensure that the public health needs of African CF patients, both in Africa and those of African descent, are met.

Key words: cystic fibrosis, Africa, molecular epidemiology

INTRODUCTION

Overview

Our awareness of cystic fibrosis (CF) significantly pre-dates our ability to comprehend the molecular factors which underpin its cause and affect prognosis. The dire warning, "Woe to that child which when kissed on the forehead tastes salty. He is bewitched and soon must die," has been circulating since the 1800s[1]. CF patients taste salty when kissed as an elevated level of sweat chloride is a hallmark feature of the disease[2] and, if untreated, they rarely make it past their first birthday[3]. Typical presentation also includes failure to thrive caused by pancreatic insufficiency and chronic recurrent chest infections[4].

The molecular cause of CF was only elucidated in 1989 with the identification of *CFTR* as the responsible gene[5]. Almost 2 000 mutations have been identified since[3], which have been grouped into classes. Classes I-III tend to be associated with more severe forms of CF while patients carrying mutations from classes IV-VI tend to have milder symptoms. The possible effects range from the production of functional CFTR with a reduced ability to traffic chloride ions (class VI) to the production of an immature CFTR which is destroyed before it reaches the cell membrane (class I[3]). There are drugs available that target specific mutation classes while other candidates are the subject of research or have progressed to clinical trials[6].

Diagnosing CF in Africa

The highest reported prevalence of CF is among individuals of Caucasian descent[7] in whom it is the leading cause of death among autosomal recessive diseases[7,8]. It was assumed that CF only affected Caucasians which skewed research efforts. A report from the World Health Organisation (WHO)

published in 2002 revealed that only a handful of countries, all European, had a mutation detection rate above 95% (Denmark, the Czech Republic, the Netherlands, Belgium, Switzerland and the Republic of Ireland[9]). North America, along with Australia, New Zealand, Sweden, Portugal, Luxembourg, the United Kingdom, Italy, Malta and Greece, had a mutation detection rate of 80-89%. In contrast, only two African countries had data at the time of the publication of the report – Algeria (60-69% detection rate) and South Africa (70-79%[9]).

CF was first reported in Africa in a Black South African baby who died within half an hour of birth[10]. Several years later, twin Black African boys became the first confirmed cases of CF at Baragwanath Hospital in Johannesburg, leading the authors to urge clinicians to consider CF as a possible diagnosis in this ethnic group[11]. This recommendation has not always been heeded to the detriment of non-Caucasian CF patients. The early diagnosis of CF patients is critical, since diagnosis after six weeks doubles their risk of developing severe pulmonary disease which is the leading cause of death among CF patients[12,37].

There are several factors on the African continent which prevent the identification of CF patients within the critical six week window. First, the assumption among clinicians that CF predominantly affects Caucasians has not been completely dispelled, resulting in under diagnosis among non-Caucasians. Second, CF diagnosis in Africa may be overshadowed by more rampant phenocopic illnesses such as protein energy malnutrition, chronic pulmonary infections and HIV[13]. Third, CF patients don't always display the classic triad of symptoms; in one study, only 4.6% of CF patients had all three symptoms[4] increasing the likelihood of misdiagnosis. Fourth, the gold standard diagnostic for CF (the sweat test[14]) can return false negatives particularly in mild cases[15]. The sweat

test requires technical skill and rigour and can only be reliably performed at larger health centres and hospitals, reducing its availability to rural populations[16]. In Africa, there may still be countries where this test is not available at all. In Sudan, for example, the sweat test was only introduced in 2008 and was available in one hospital in Khartoum[17]. Fifth, almost 2 000 mutations have been identified in CF patients[3] yet the full complement of causative mutations, particularly in non-Caucasian populations, remains unknown. This eliminates the possibility of relying solely on an existing genetic test for the unambiguous diagnosis of CF.

Determining the mutations which cause CF is a complex issue. While the ΔF508 mutation has been estimated to account for 70-90% of all the affected CF chromosomes globally[3,18], common variants are rare. Of the almost 2 000 variants identified only 20 of them have an individual global frequency in excess of 0.1%[3]. This means that 10-30% of the chromosomes present in CF patients can be expected to carry rare mutations. CF mutations are also population specific, varying by race and country[19]. Both the population-specific nature of the causative mutations and the fact that many of these variants are rare complicate the development of comprehensive molecular diagnostic tools. A diagnostic test can only be effective after members of the sub-populations in a country with symptoms suggestive of CF have had their entire *CFTR* sequenced and the identity and prevalence of each variant has been determined. This data can be used to define a diagnostic instrument with a local mutation detection rate above 90%, which is necessary for the test to be used with confidence. Since there are no mutation hotspots, it would be best to adhere to WHO's recommendation to use this population sequencing strategy in 50-100 putative CF patients in a given population in order to establish a relevant diagnostic tool[9].

CF Patient Survival

The life expectancy for a European CF patient is approaching 40 years, which is a marked improvement over the life expectancy in 1938 of six months[3]. European patients benefit from the fact that research has been conducted on their populations for a longer period resulting in a higher mutation detection rate. This is frequently paired with a new-born screening programme (NBSP) which enables early diagnosis. CF babies can then commence treatment, which restricts CF-related organ damage. European CF patients also tend to benefit from referral to multi-disciplinary CF centres, which have a range of CF professional staff including pulmonologists, pharmacists, psychologists, social workers and microbiologists[20,21]. This specialised holistic approach leads to fewer hospitalisations and allows these children to lead healthier lives. European CF patient registries are invaluable for accurately determining disease epidemiology. This data can be used to inform each country's specific CF health policy. It also allows clinicians and researchers to track the impact of various interventions on patient quality of life and life expectancy. It has likewise been useful in identifying which centres are the most successful in ameliorating the effects of CF[22-24]. All of this data can be used to make the necessary improvements to ensure a high standard of treatment for European CF patients.

Data from North America and Australia reflect similar trends in survival and treatment of CF patients. Americans have access to a network of accredited CF care centres[25]. By 2009, all 50 American states had implemented a CF NBSP which led to 57.5% of Americans with CF being diagnosed at birth in 2010. The age at diagnosis decreased from six months to one month[26] which positively impacts prognosis[12,37]. If mortality rates among the American CF population remain as they were in 2010 then American CF patients born in 2010 will live to at least 39 years. If mortality

rates continue to decrease as they did in the first decade of the millennium then American CF babies born in 2010 can be expected to live to 45 years[26]. In Canada, which also has a network of specialised CF centres, median survival for a CF patient based on 2007 data was 49.7 years. Canada is in the process of introducing a NBSP which is available in about 60% of its provinces[27]. A national CF NBSP commenced in 2003 in Australia[28], where the median age at death is 30 years[29].

In stark contrast, the life expectancy for a South African CF patient is 20.5 years[30]. There is no South African NBSP nor a population-specific genetic test. Additionally, among the North African countries, only Egypt has an NBSP but this only targets congenital hypothyroidism[31]. An internet search did not identify any African CF registries. This negatively impact life expectancy, hospitalisation frequency and the quality of life of Africans with CF.

The maximum annual treatment cost in South Africa per patient has been estimated to be R360 000[32], approximately €26,905. This is triple the average annual wage[33]. Based on European and American data, the cost of CF (including indirect costs) can vary from €16 307-€394 518 per patient[34], depending on disease severity. Using Australian CF registry data, it was determined that the annual cost per patient with mild CF was €8 736, whereas moderate and severe CF cost €22 071 and €28 994 respectively[35]. The estimated South African cost may reflect the impact of delayed diagnosis on treatment costs - CF patients are on average diagnosed at 13 months[36]. Delayed diagnosis has been shown to have a negative impact on nutrition, growth and lung function. CF patients who receive a delayed diagnosis also tend to be hospitalised more frequently with more severe pulmonary exacerbations, increasing treatment costs[37].

While it has been established that CF does indeed affect Africans, no comprehensive public health policy exists to increase the longevity of African CF patients. It may be assumed that the incidence of CF is significantly lower than the major challenges faced by African health care such as HIV and tuberculosis. However, without national CF registries and population screening for CF, there is no way to accurately determine the scale of the problem. Additionally, CF is often misdiagnosed due in part to the prevalence of phenocopic illnesses but also because the mutations which cause CF are population-specific, necessitating the development of local genetic tests. Available commercial genetic tests have largely been designed based on European data and thus are not able to detect mutations specific to a given African nation. In the absence of an appropriate public health response, African CF patients are frequently misdiagnosed allowing organ damage to proceed unchecked. Misdiagnosis therefore drives up the cost of the illness while simultaneously decreasing life expectancy. This is in contrast to the European example in which there is a comprehensive continental public health framework for treating CF with a concomitant improvement in the quality of life and longevity. It was therefore our aim to survey the literature to assess Africa's state of readiness to implement the kind of public health programme that would make a significant difference in the quality of life of African CF patients.

MATERIALS AND METHODS

A survey was conducted of the molecular epidemiology of CF in Africa until January 22, 2015. Google Scholar was used to search for the term "cystic fibrosis" and the name of each of the 49 states that make up continental Africa was used together with the Republic of Cabo Verde, São Tomé and Príncipe, the Republic of Seychelles, the Union of Comoros and the Republic of Madagascar. To be included, the paper either had to divulge the results of molecular screening done

in or outside Africa on African citizens (namely individuals born in Africa who had migrated or the children born to these persons). We also included the results of Jewish populations of African origin in whom we could expect minimal admixture with the people of their current country of residence. We included patients with congenital bilateral absence of the vas deferens (CBAVD) which is a CF-related disease. Carrier screening data were excluded. When potential CF patients were screened for only one mutation and a second mutation remained unidentified, this was not counted in the category of "unknown mutation" since it is possible that the second mutation could have been uncovered if a more thorough investigation had been performed. If the report detailed the search for more than one mutation, any chromosome that remained unidentified was counted in the category of "unknown mutation", ("U"). Investigations which searched for only one mutation and where none of the chromosomes were positive for that mutation were excluded. The data were recorded per chromosome (and not per patient) and only individuals in whom at least one mutation was identified were included in the collated data. We chose to record the data per chromosome for consistency given that many CF patients are heterozygotes.

RESULTS

Twelve African countries had published a total of 26 reports of molecular investigations into the cause of CF – Morocco[38-40], Algeria[41-43], Tunisia[40,41, 44-46], Libya[47], Egypt[48-50], Sudan[17], Rwanda[51], Senegal[52], Cameroon[52-53], Namibia[54], Zimbabwe[55], and South Africa[2,36,55-60] – with a total of 2 344 chromosomes having been screened. Only two publications met our inclusion criteria while also screening non-CF patients[39,58] which means that the data cannot be used to determine carrier rates in the general population. Since the publications span the period 1990-2014, methods as diverse as single strand conformation polymorphism and sequencing were employed. There was no standard

list of mutations used for screening resulting in differences between studies (Table S1). The data therefore report the total number of chromosomes tested for each mutation as well as the number of chromosomes found to be positive for a given mutation. The largest cohorts were investigated in North Africa (1 334 chromosomes) and Southern Africa (760 chromosomes; Table S2). Sudan, Rwanda, Cameroon and Zimbabwe were the only countries where ΔF508 was either not included in the genetic screen or was not detected. In the remaining eight countries, 2 052 chromosomes were screened for ΔF508 which was detected at a frequency of 48%; this is lower than the global average of 70-90%[3,18].

A total of 79 variants were detected in the included reports of which 38 have been shown empirically to cause CF, 10 are of varying clinical consequence[61-62], five have no associated evidence of pathogenicity, four are synonymous, five were novel, and 21 are unique to Africa (four of the 21 were novel). The most frequently detected alleles were ΔF508 (992 chromosomes), 3120+1G>A (83 chromosomes), G542X (58 chromosomes), N1303K (51 chromosomes), W1282X (48 chromosomes), E1104X (41 chromosomes), 711+1G>T (36 chromosomes), 3272-26A>G (17 chromosomes), and 394delTT (15 chromosomes). It should also be noted that for 224 chromosomes, the methods used did not result in the resolution of the patient's molecular CF status. These instances are represented by U in Table 1.

Only 14 of the 26 papers did not use kits or mutation-specific screening methods and could therefore identify novel or Africa-specific variants (table S1). These studies generated data for 1 120 or 49% of the chromosomes screened and led to the identification of 21 mutations that are unique to Africa. The rest of the chromosomes were tested using commercially available kits or mutation-

Table 1. Summary of mutations reported in African CF patients.

Mutation	# of alleles tested	+ alleles	Allele Frequency/%	Nationality
ΔF508	2 052	992	48.34	Egyptian (CBAVD), Algerian, Tunisian, Libyan, Moroccan, Namibian, South African
U	1 448	224	15.47	Egyptian, Algerian, Tunisian Jews, Tunisian, Libyan, Moroccan, Senegalese, South African, Rwandan
G542X	1 192	58	4.87	Tunisian, South African
N1303K	1 134	51	4.50	Egyptian, Algerian, Tunisian, Libyan, South African
W1282X	1 118	48	4.29	Tunisian, South African
3120+1G>A	724	83	11.46	Rwandan, South African, Zimbabwean
E1104X	712	41	5.76	Algerian, Tunisian, Libyan
711+1G>T	710	36	5.07	Algerian, Tunisian, Moroccan
V201M	568	5	0.88	Tunisian
D1270N	560	5	0.89	Tunisian, Moroccan, South African
T665S	556	2	0.36	Egyptian, Tunisian
R74W	542	3	0.55	Tunisian, Moroccan
2766del8	540	10	1.85	Tunisian
F1166C	540	1	0.19	Tunisian
G85E	540	6	1.11	Tunisian
L1043R	540	1	0.19	Tunisian
R1066C	540	1	0.19	Tunisian

Mutation	# of alleles tested	+ alleles	Allele Frequency/%	Nationality
Y122X	540	1	0.19	Tunisian
3272-26A>G	488	17	3.48	South African
G551D	488	5	1.02	South African
5T	420	27	6.43	Egyptian (CBAVD), Algerian, Tunisian, Moroccan
1717-1G>A	402	1	0.25	South African
2789+5G>A	402	1	0.25	South African
3659delC	402	1	0.25	South African
394delTT	402	15	3.73	South African
621+1G>T	402	1	0.25	South African
Q493X	402	1	0.25	South African
R1162X	402	1	0.25	South African
R117H	402	1	0.25	South African
R553X	402	4	1.00	South African
S549N	402	1	0.25	South African
11TG	300	4	1.33	Moroccan
12TG	300	4	1.33	Moroccan
1811+5A>G	136	1	0.74	Tunisian
4016insT	136	1	0.74	Tunisian
4268+2T>G	136	2	1.47	Tunisian
I1203V	136	2	1.47	Tunisian
R1158X	136	2	1.47	Tunisian
R785X	136	1	0.74	Tunisian

Mutation	# of alleles tested	+ alleles	Allele Frequency/%	Nationality
405+4A>G	122	1	0.82	Cameroonian
A204T	120	1	0.83	Rwandan
c.1001+11C>T	120	5	4.17	Rwandan
c.1898+152T>A	120	9	7.50	Rwandan
c.2752-15C>G	120	3	2.50	Rwandan
c.3041-71A>G	120	2	1.67	Rwandan
c.3272-32T>C	120	1	0.83	Rwandan
c.4575+2G>A	120	1	0.83	Rwandan
E527E	120	3	2.50	Rwandan
F693L	120	2	1.67	Rwandan
M470V	120	13	10.83	Rwandan
P1290P	120	6	5.00	Rwandan
Q1463Q	120	14	11.67	Rwandan
T854T	120	52	43.33	Rwandan
7T	80	57	71.25	Egyptian (CBAVD)
9T	80	9	11.25	Egyptian (CBAVD)
c.1418delG	74	?		Egyptian
c.2620-15C>G	74	?		Egyptian
c.3718-24G>A	74	?		Egyptian
c.3877G>A	74	?		Egyptian
3849+10kbC>T	34	1	2.94	Moroccan Jews
G1244E	34	1	2.94	Moroccan Jews
G1249E	34	7	20.59	South African

Mutation	# of alleles tested	+ alleles	Allele Frequency/%	Nationality
S549R	34	4	11.76	Moroccan Jews
-94G>T	28	1	3.57	South African
2183delAA	28	1	3.57	South African
3196del54	34	2	5.88	South African
1812-1G>A	20	1	5.00	Algerian
c.1670delC	20	2	10.00	Libyan
1898+3A>C	16	1	6.25	Egyptian
405+1G>A	12	8	66.67	Tunisian Jews
D579G	6	1	16.67	Sudanese
R1102K	6	1	16.67	Sudanese
3729delAinsTCT	4	2	50.00	Tunisian
c.54-1161_c.164+1603del2875	4	2	50.00	South African, Zimbabwean
N	4	1	25.00	Tunisian
1609delCA	2	2	100.00	Algerian
c.4136+1G>A	2	1	50.00	Senegalese
EX17a-EX18del	2	2	100.00	Senegalese
R1070W	2	1	50.00	Moroccan
Y1109X	2	2	100.00	Cameroonian

Mutations in red have been empirically shown to cause CF, mutations in purple were reported as novel in the original publication, mutations in blue have varying clinical consequence, mutations highlighted in yellow are unique to Africa. ? = number of alleles was not reported, U = unknown mutations, N = normal CFTR allele.

specific methods. Only twelve variants were found in more than one country, underscoring the population-specific nature of CF mutations. Nine countries have published data on the *CFTR* of fewer than 100 patients (as per WHO's 2002 directive[9]). It should be noted that in nine of the twelve countries, variants were discovered that have never been reported outside of Africa. **Tunisia** heads the list with seven variants (2766del8, T665S, F1166C, 4268+2T>G, L1043R, 3279delAinsTCT and 1811+5A>G), followed by **Rwanda** (A204T, 3041-71A>G, 4575+2G>A, 3272-32T>C), **Egypt** (1898+3A>C, T665S), **South Africa** (-94G>T, c. 54-1161_c.164+1603del2875), **Senegal** (c.4136+1G>A, EX17a-EX18del), **Cameroon** (Y1109X, 405+4A>G), **Sudan** (R1102K) and **Zimbabwe** (c. 54-1161_c.164+1603del2875). Only two of these Africa-specific variants (T665S and c. 54-1161_c.164+1603del2875) occurred in more than one nation. As of this writing, the majority of African countries (42) have no published account of a search for *CFTR* mutations in their population. As depicted in Figure 1, most of the molecular work has been done in the Northern African populations of Morocco, Algeria, Tunisia, Libya and Egypt and in South Africa. In East Africa, only Rwanda has any published work, while West Africa is represented by Senegal and Cameroon. In southern Africa, published molecular reports are from Zimbabwe, Namibia and South Africa.

It should also be noted that none of the papers (save two[48,49]) that could identify novel mutations relied solely on sequencing. Most investigations utilised less sensitive methods (PCR-RFLP for instance; Table S1) and only progressed to sequencing if there was an abnormal pattern. In some instances, the authors chose to screen for a specific mutation (such as ΔF508) and if no mutations were detected proceeded to a different screening method whose result would again determine if sequencing would be done. In others, if the mutation specific test yielded no results, the authors

proceeded directly to sequencing. While this was one way to handle the cost and complexity of sequencing a large gene (*CFTR* is 189 kb long) it does raise the question of what mutations may have been missed via this strategy.

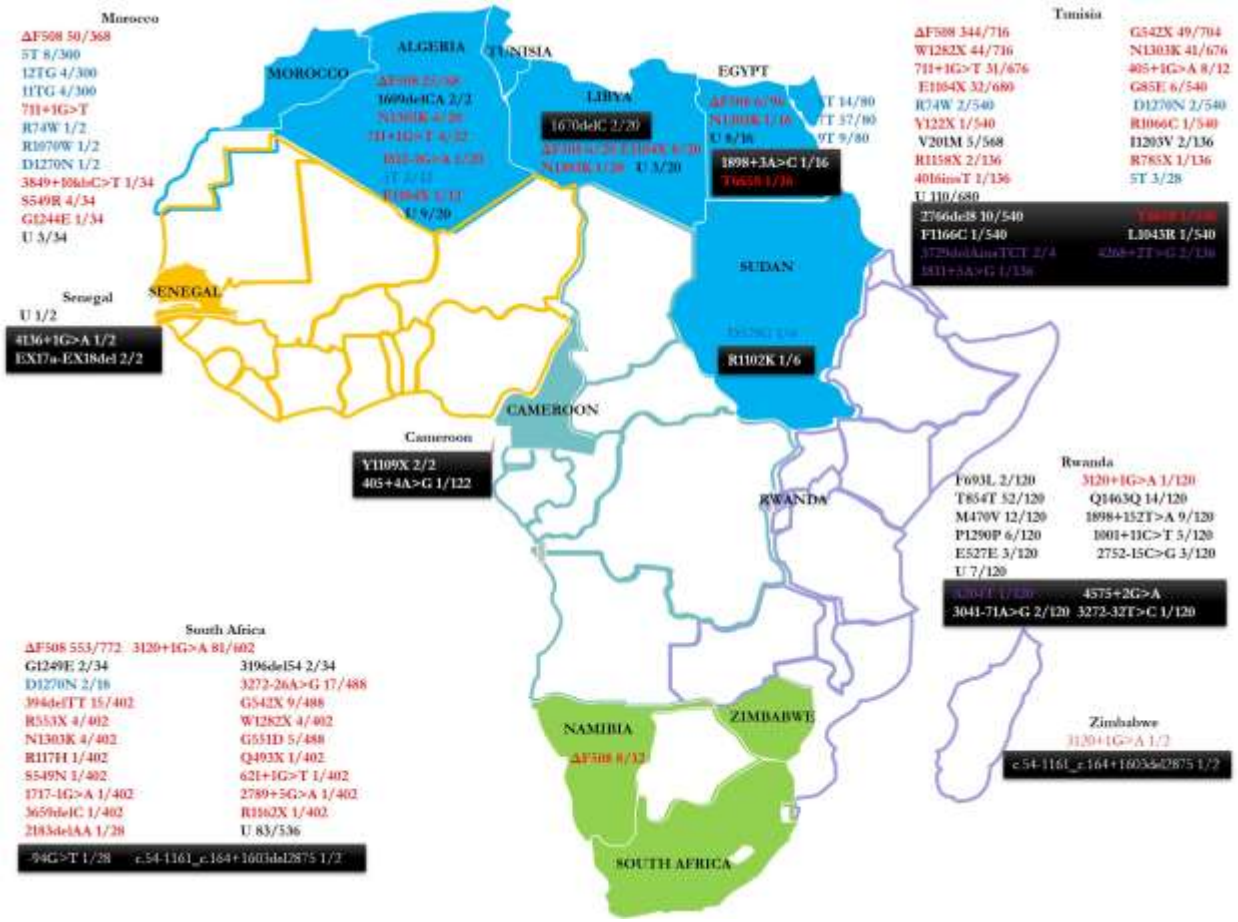

Figure 1: Mutations identified in suspected African CF patients. Mutations in red have been empirically shown to cause CF, mutations in purple were reported as novel in the original publication, mutations in blue have varying clinical consequence, mutations in black boxes are unique to Africa, countries in white have no molecular data. The fractions represent the number of each allele identified in a population. The continent was divided into North, West, Central, East and Southern Africa and colour coded accordingly.

DISCUSSION

Cystic fibrosis patients on the African continent have a life expectancy of 20.5 years[30], whereas in developed countries CF patients have a life expectancy upwards of 40 years[3, 25-27]. A public health intervention could be useful in raising African CF patients' life expectancy with a concomitant improvement in their quality of life. As an example, in 2002 South African life expectancy, including the impact of HIV, was 54.68 years; in 2014 the life expectancy had risen to 62.5. This marked improvement is due in large part to a concerted public health thrust to ensure the availability of anti-retroviral drugs and to reduce mother to child transmission of the virus[63]. Similarly, a public health intervention could radically improve the quality of life of Africans with CF. However, our survey has revealed that there is much to be done before such a strategy could be implemented.

One of the issues that needs to be addressed is the lack of epidemiological data. We recommend that each African nation establishes a CF registry. This will facilitate the accurate determination of the prevalence and incidence of CF in Africa which would define its impact and inform each nation's health policy. The registries would also be an invaluable tool for tracking the impact of various interventions on the overall health of CF patients and will allow governments to realise when they have met or exceeded targets set for improving the life expectancy and quality of life of CF patients[64].

Diagnosis of CF relies on two standard biochemical analyses: the sweat test and the faecal elastase-1 assay. Both tests should be made available in at least the major medical centres and hospitals. Molecular diagnosis would be an unambiguous way of identifying a CF patient if the full complement of causative mutations was known. However, since the causative mutations tend to be population specific, each country will need to determine the mutations present among its peoples

and tailor the molecular diagnostic tools accordingly. Africa is lagging in this regard, with 78% of countries having no published record of any molecular investigation of CF. Three-quarters of the countries that have published molecular screening work appear to have screened less than 100 patients, indicating that there are likely to be additional unidentified mutations.

Of the 79 variants identified, 10 are of varying clinical consequence, five are synonymous, and 38 are empirically pathogenic. Most of this data was generated as part of CFTR2, a collaborative initiative which has defined the pathogenicity of variants identified in North American and European CF patients[61-62]. 60% of the variants identified in Africans have been assessed by CFTR2 and thus have also been described in patients of mainly Caucasian descent. This high percentage may firstly be explained by the fact that, in several studies, investigators chose to screen for known mutations which had previously been identified in European patients as opposed to taking an unbiased sequencing approach to identifying *CFTR* mutations. Second, it has been shown that there is a certain amount of admixture in North American CF patients whose population substructures reveal African, Indian and Mexican influences[65]. Third, there is evidence that Europeans returned to Africa after the initial "out of Africa" migration leading to European admixture among the peoples of the Mother Continent[66].

There are 27 mutations identified in African CF patients whose impact on the CFTR still needs to be investigated. The discovery of *CFTR* mutations is insufficient evidence of causation as the identified mutations could be benign. Functional studies, similar to those described recently[61], therefore need to be conducted on 34% of the mutations listed in Table 1 to ascertain how many are deleterious. It is generally accepted that pathogenic mutations reduce CFTR's ion trafficking ability

to less than 10%[67]. Once the pathogenicity of these variants has been empirically established, this data can be used to develop molecular CF diagnostic tests specific to the sub-populations on the African continent.

It should be noted that for 51% of the chromosomes studied, commercially available kits or mutation-specific screening methods were used. Since these kits were largely developed without African data, they could not identify Africa-specific alleles nor novel mutations. To underscore the point, 21 of the 27 mutations that require functional studies have been identified only in Africa. Of these 21, 19 are unique to the particular African nation in which they were identified, lending further credence to the idea that CF mutations are likely to be population specific. Only 3.8% of the chromosomes were subjected to sequencing without first using some less sensitive molecular screen, like PCR-RFLP. While screening is one way to reduce the cost of sequencing the gene, it also introduces the possibility that some mutations may not have been detected, which is less than ideal.

The abundance of private mutations among African CF patient populations is not unexpected since Africans are known to possess the highest level of genomic variation on the planet and the largest number of population-specific alleles[68]. There are several reasons for this including the fact that Africa is the birthplace of modern humans, the diversity of the African landscape (which includes deserts and the second largest tropical rainforest in the world) and the selection pressure applied by the high burden of various infectious diseases[69]. Genetic diversity is also indicated by the fact that Africa is home to about 2 000 ethnolinguistic groups, which may be taken as a proxy for unique genetic sub-populations since individuals are more likely to reproduce with those who speak their language(s)[70]. The participants in the out of Africa migration experienced a population bottleneck,

which resulted in the loss of the considerable levels of genomic diversity resident in their African forebears[68,70].

Despite the unparalleled genomic diversity known to exist among Africans, these populations remain understudied, as has been revealed here with African CF patients. In order to begin to address this genomic knowledge deficit, the African Genome Variation Project (AGVP) genotyped or performed whole genome sequencing on 16 populations living in sub-Saharan Africa. The Project identified 16-24% novel variants and 11-23% private variants among these people groups[66]. Based on our survey of variants present in African CF patients, 27% of the mutations identified were private, or unique to the continent. Given the known genetic diversity among the peoples of Africa, it is therefore imperative that African nations develop African screening tools in order to effectively diagnose and manage African CF patients.

Initially, a complete molecular diagnosis was only important to unambiguously determine a patient's CF status. With recent advances in pharmacotherapy, a molecular diagnosis has also become essential in determining what treatment avenues a patient may benefit from[6]. The approval of Kayledeco by the Food and Drug Administration (FDA), indicated for the treatment of patients carrying at least one G551D mutation, heralded the start of a new era[7]. Kayledeco is a potentiator that addresses gating defects and may thus be useful in the treatment of other class III mutations. It has been shown to improve conductance in cell lines expressing missense and splicing mutations associated with residual CFTR function[7]. Determining what mutations a CF patient has also means knowing what class(es) those mutations fall into and thus what drug therapies, experimental or approved, may be beneficial. More class-specific drugs are being developed, which means that

clinicians would also know which clinical trials a patient may be eligible for. African patients can only benefit from both of these scenarios if investment is made into diagnosing CF at the molecular level.

It is unfortunately the case that genetics services are seen as luxuries in countries where malnutrition and infection still form a large part of the disease burden. However, there may be a higher prevalence of genetic disorders in developing countries, which tend to underreport these diseases, than in developed nations, making genetic services more relevant in the former. The burden of genetic diseases tends to exact a higher cost there as well, due to limited resources[71]. Having a child with an inherited disease also poses the risk of impoverishing the family and putting their siblings at higher risk of poor health and mortality in countries with a high prevalence of malnutrition and infection[72]. The WHO therefore recommends that "Public health authorities must acknowledge the reality that these conditions are indeed major causes of disease, disability, suffering and death in their countries, and recognise that there are approaches for their management and prevention that can significantly reduce their burden in a cost efficient manner." It should also be noted that "The societal costs of inaction in genetics, measured in terms of avoidable human suffering and burden to public health, are very high." The WHO also recognises that CF belongs to a group of the more common severe genetic diseases and that it could "…contribute significantly to chronic morbidity in childhood in many developing countries." It therefore recommends that "Programs for the prevention and care of affected children with these conditions may significantly reduce the overall burden due to chronic disease at the community level."[71]

The World Health Assembly (WHA) has recognised that a lack of epidemiological data may prevent effective management of birth defects, defined as any structural or functional abnormality present from birth, such as CF[73]. Based on our survey of CF on the continent, there is no published data concerning the search for the molecular cause of CF in most of the African states. There are no registries for CF either, which means that there is no reliable baseline epidemiological data (such as incidence and prevalence). Additionally, in the absence of population-specific genetic tests, the true incidence of CF will be underreported since there will be an elevated false negative rate.

The WHA has made several recommendations to improve the treatment of those with birth defects: a) resources should be devoted both to programmes that prevent birth defects and that provide care for those carrying these defects; b) awareness should be raised concerning the importance of new-born screening programmes; c) data should be collected in order to accurately determine the scale of the problem; and d) appropriate community genetic services should be dispensed within the purview of the primary health care system[73]. All of these recommendations are relevant to Africa. It is clear from our survey that even with the underreporting, there are many CF patients on the continent. A new-born screening programme would assist in identifying CF patients soon after birth who could then benefit from starting the appropriate course of treatment sooner, thereby improving their quality of life and longevity. A community genetics approach would allow for both the establishment of national registries and genetic counselling. The former would assist governments in accurately characterising their CF population and proposing an appropriate public health strategy to improve their overall health. The latter would assist parents of CF patients to understand their child's condition and to make an autonomous decision concerning their future reproductive choices.

By virtue of being born in Africa, CF patients are at a disadvantage and can expect to live only half as long as their European counterparts[3,30]. This disparity is due in large part to the fact that an effective and comprehensive public health strategy to deal with CF does not exist in Africa. This situation must be viewed in the light of Article 24 of the United Nations' Convention on the Rights of the Child (CRC) which states that every child should enjoy "…the highest attainable standard of health." It calls upon States to "…strive to ensure that no child is deprived of his or her right of access to such health care services…" via the "application of readily available technology…"[74]. Cystic fibrosis has long been a challenge to health care systems, particularly among European populations where the disease prevalence is apparently highest. It is undeniable that CF is also an issue among the peoples of Africa. However, more epidemiological data needs to be gathered in national registries before we can accurately measure the scale of the problem on the Mother Continent. This includes the need to sequence the *CFTR* of suspected CF patients in all the countries that constitute Africa in order to identify the responsible mutations. Given the diversity of Africans, we can reasonably expect that a number of private mutations will be uncovered. Using this molecular data to diagnose patients and commencing treatment as soon as possible would allow for the fulfilment of the basic health rights afforded to children by the CRC.

ACKNOWLEDGEMENTS

The authors would like to thank the University of Pretoria Vice-Chancellor's Post-doctoral Research Programme, the Institute for Cellular and Molecular Medicine at the University of Pretoria and the Genomics Research Institute (a University of Pretoria Institutional Research Theme) for funding this study.

References

1. Hodson M, Bush A, Geddes D, editors. Cystic Fibrosis, 3rd ed. Boca Raton, Florida: CRC Press, 2007.
2. Goldman A, Graf C, Ramsay M, Leisegang F, Westwood ATR. Molecular diagnosis of cystic fibrosis in South African population. *S Afr Med J* 2003;93:518-519.
3. Ikpa PT, Bijvelds MJC, de Jonger HR. Cystic fibrosis: toward personalised therapies. *Int J Biochem Cell Biol* 2014;52:192-200.
4. Westwood T, Henderson B, Ramsay M. Diagnosing cystic fibrosis in South Africa. *S Afr Med J* 2006;96:304-306.
5. Riordan JR, Rommens JM, Bat-Sheva Kerem A, et al. Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. *Science* 1989;245:1066-1073.
6. Amin R, Ratjen F. Emerging drugs for cystic fibrosis. *Expert Opin Emerg Drugs* 2014;19:143-155.
7. Bell SC, De Boeck K, Amaral MD. New pharmacological approaches for cystic fibrosis: Promises, progress, pitfalls. *Pharmacol Ther* 2015;145:19-34.
8. Snouwaert JN, Brigman KK, Latour AM, et al. An animal model for cystic fibrosis made by gene targeting. *Science* 1992;257:1083-1088.
9. *The molecular genetic epidemiology of cystic fibrosis*. Genoa: World Health Organisation; 2002.
10. Grové SS. Fibrocystic disease of the pancreas in the Bantu. *S Afr J Lab Clin Med* 1959;5:113-118.
11. Levin SE, Blumberg H, Zamit R, Schmaman A, Wagstaff L. Mucoviscidosis (cystic fibrosis of the pancreas) in Bantu twin neonates. *S Afr Med J* 1967;41:482-485.
12. Wang SS, O'Leary LA, FitzSimmons SC, Khoury MJ. The impact of early cystic fibrosis diagnosis on pulmonary function in children. *J Pediatr* 2002;141:804-810.
13. Mutesa L, Bours V. Diagnostic challenges of cystic fibrosis in patients of African origin. *J Trop Pediatr* 2009;55:281-286.
14. Gibson LE, Cooke RE. Test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilising pilocarpine by iontophoresis. *Pediatr* 1959;23:545-549.

15. Highsmith WE, Burch LH, Zhou Z, et al. A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. *N Engl J Med* 1994;331:974-980.
16. Goldman A, Labrum R, Claustres M, et al. The molecular basis of cystic fibrosis in South Africa. *Clin Genet* 2001;59:37-41.
17. Ibrahim SA, Elmola MAF, Karrar ZA, et al. 2014. Cystic fibrosis in Sudanese children: first report of 35 cases. *Sudan J Pediatr* 2014;14:39-44.
18. Cystic Fibrosis Genetic Analysis Consortium. Population variation of common cystic fibrosis mutations. *Hum Mutat* 1994;4:167-177.
19. Bobadilla JL, Macek Jr. M, Fine JP, Farrell, PM. Cystic fibrosis: a worldwide analysis of *CFTR* mutations – correlation with incidence data and application to screening. *Hum Mutat* 2002;19:575-606.
20. Castellani C, Conway S, Smyth AR, Stern M, Elborn JS. Standards of care for cystic fibrosis ten years later. *J Cyst Fibros* 2014;13:S1-S2.
21. Kerem E, Conway S, Elborn S, Heijerman H. Standards of care for patients with cystic fibrosis: a European consensus. *J Cyst Fibros* 2005;4:7-26.
22. Kerem E, Webb KA. European cystic fibrosis society standards of care: A road map to improve CF outcome. *J Cyst Fibros* 2014;13:357-358.
23. Stern M, Bertrand, DP, Bignamini E, et al. European Cystic Fibrosis Society Standards of Care: Quality Management in cystic fibrosis. *J Cyst Fibros* 2014;13:S43-S59.
24. Sheppard DN. The European cystic fibrosis patient registry: The power of sharing data. *J Cyst Fibros* 2010;9:S1-S2.
25. Cystic Fibrosis Foundation - Care Center Network. 2015. Available at: http://www.cff.org/LivingWithCF/CareCenterNetwork/. Accessed February 2, 2015.
26. MacKenzie T, Gifford AH, Sabadosa KA, et al. Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: Survival analysis of the cystic fibrosis foundation patient registry. *Ann Intern Med* 2014;161:233-241.

27. Cystic Fibrosis Canada - Annual Report of The Canadian Cystic Fibrosis Registry. 2012. Available at: http://www.cysticfibrosis.ca/wp-content/uploads/2014/03/Canadian-CF-Registry-English-FINAL-FOR-WEB1.pdf. Accessed February 3, 2015.
28. Martin B, Schecter MS, Jaffe A, et al. Comparison of the US and Australian Cystic Fibrosis Registries: The impact of newborn screening. *Pediatr* 2012;129:e348-e355.
29. Australian Cystic Fibrosis Data Registry - 15th Annual Report from the Australian Cystic Fibrosis Data Registry. 2012. Available at: http://www.cysticfibrosis.org.au/media/wysiwyg/CF-Australia/medical-documents/ACFDR_2012/ACFDR_Annual_Report_2012r.pdf. Accessed February 3, 2015.
30. Westwood ATR. The prognosis of cystic fibrosis in South Africa: a 33 year study. *J Cyst Fibros* 2008;7(S2):458.
31. Saadallah AA, Rasheed MS. Newborn screening: Experiences in the Middle East and North Africa. *J Inherit Metab Dis* 2007;30: 482-489.
32. South African Cystic Fibrosis Association - Medication. 2015. Available at: http://www.sacfa.org.za/about.asp?r=&spage=9. Accessed July 25, 2014.
33. Statistics South Africa - Census 2011. Available at: http://www.statssa.gov.za/publications/P03014/P030142011.pdf. Accessed July 25, 2014.
34. Angelis A, Tordrup D, Kanavos P. 2014. Socio-economic burden of rare diseases: A systematic review of cost of illness evidence. *Health Policy* 2014. doi: http://dx.doi.org/10.1016/j.healthpol.2014.12.016
35. van Gool K, Norman R, Delatycki MB, Hall J, Massie J. Understanding the costs of care for cystic fibrosis: an analysis by age and health state. *Value Health* 2013;16:345-355.
36. Masekela R, Zampoli M, Westwood AT, et al. Phenotypic expression of the 3120+ 1G> A mutation in non-Caucasian children with cystic fibrosis in South Africa. *J Cyst Fibros* 2013;12:363-366.

37. Grosse SD, Boyle CA, Botkin JR, et al. Newborn screening for cystic fibrosis: evaluation of benefits and risks and recommendations for state newborn screening programs. *MMWR Recomm Rep* 2004; 53:1-36.

38. de Prada Merino A, Bütschi, FN, Bouchardy I, et al. [R74W;R1070W;D1270N]: a new complex allele responsible for cystic fibrosis. *J Cyst Fibros* 2010;9:447-449.

39. Ratbi I, Génin E, Lengendre M, et al. Cystic fibrosis carrier frequency and estimated prevalence of the disease in Morocco. *J Cyst Fibros* 2008;7:440-443.

40. Quint A, Lerer I, Sagi M, Abeliovich D. Mutation spectrum in Jewish cystic fibrosis patients in Israel: Implication to carrier screening. *Am J Med Genet* 2005;136A:246-248.

41. Boudaya M, Fredj SH, Haj RB, et al. Cystic fibrosis transmembrane conductance regulator mutations and polymorphisms associated with congenital bilateral absence of vas deferens in a restricted group of patients from North Africa. *Ann Hum Biol* 2012;39:76-79.

42. Loumi O, Baghriche M, Delpech M, Kaplan JC, Bienvenu T. Analysis of the complete coding region of the CFTR gene in ten Algerian cystic fibrosis families. *Hum Hered* 1999;49:81-84.

43. Lucotte G, Barrè E, Berriche S. Frequency of the cystic fibrosis mutation ΔF508 in Algeria. *Hum Genet* 1991;87:759.

44. Fredj SH, Boudaya M, Oueslati S, et al. New frameshift CF mutation 3729delAinsTCT in a Tunisian cystic fibrosis patient. *J Genet* 2013;92:81.

45. Fredj SH, Messaoud T, Templin C, des Georges M, Fattoum S, Claustres M. Cystic fibrosis transmembrane conductance regulator mutation spectrum in patients with cystic fibrosis in Tunisia. *Genet Test Mol Biomarkers* 2009;13:577-581.

46. Messaoud T, Fredj SBH, Bibi A, Elion J, Férec C, Fattoum S. Épidémiologie moléculaire de la mucoviscidose en Tunisie. *Ann Biol Clin* 2005;63:s69-70.

47. Fredj SH, Fattoum S, Chabchoub A, Messaoud T. First report of cystic fibrosis mutations in Libyan cystic fibrosis patients. *Ann Hum Biol* 2011;38:561-563.

48. El-Seedy A, Pasquet M-C, Shafiek H, El-Komy M, Kitzis A, Ladevèze V. Cystic fibrosis in Egypt: new mutational detection of the *CFTR* gene in patients from Alexandria, northern Egypt. *J Cyst Fibros* 2012;12:S53.
49. Naguib ML, Schrijver I, Gardner P, et al. Cystic fibrosis detection in high-risk Egyptian children and CFTR mutation analysis. *J Cyst Fibros* 2007;6: 111-116.
50. Lissens W, Mahmoud KZ, El-Gindi E, et al. Molecular analysis of the cystic fibrosis gene reveals a high frequency of the intron 8 splice variant 5T in Egyptian males with congenital bilateral absence of the vas deferens. *Mol Hum Reprod* 1999;5:10-13.
51. Mutesa L, Azad AK, Verhaeghe C, et al. Genetic analysis of Rwandan patients with cystic fibrosis-like symptoms: identification of novel Cystic Fibrosis Transmembrane Conductance Regulator and Epithelial Sodium Channel gene variants. *Chest* 2009;135:1233-1242.
52. Feuillet-Fieux MN, Ferrec M, Gigarel N, et al. Novel *CFTR* mutations in black cystic fibrosis patients. *Clin Genet* 2004;65:284-287.
53. Ghanem N, Costes B, Girodon E, Martin J, Fanen P, Goossens M. Identification of eight mutations and their sequence variations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics* 1994;21:434-436.
54. Cystic Fibrosis Genetic Analysis Consortium. Worldwide survey of the ΔF508 mutation – report from the Cystic Fibrosis Genetic Analysis Consortium. *Am J Hum Genet* 1990;47: 354-359.
55. des Georges M, Guittard C, Templin C, et al. WGA allows the molecular characterisation of a novel large *CFTR* rearrangement in a Black South African cystic fibrosis patient. *J Mol Diagn* 2008;10:544-548.
56. de Carvalho C, Ramsay M. CFTR structural rearrangements are not a major mutational mechanism in black and coloured southern African patients with cystic fibrosis. *S Afr Med J* 2009;99:724.
57. Westwood T, Brown R. Cystic fibrosis in black patients: Western Cape experiences. *S Afr Med J* 2006;96:288-289.

58. Padoa C, Goldman A, Jenkins T, Ramsay M. Cystic fibrosis carrier frequencies in populations of African origin. *J Med Genet* 1999;36:41-44.

59. Carles S, Desgeorges M, Goldman A, et al. First report of CFTR mutations in Black cystic fibrosis patients of Southern African origin. *J Med Genet* 1996;33:802-804.

60. Herbert JS, Retief AE. The frequency of the delta F508 mutation in the cystic fibrosis genes of 71 unrelated South African cystic fibrosis patients. *S Afr Med J* 1992;82:13-15.

61. Sosnay PR, Siklosi KR, Van Goor F, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. *Nat Genet* 2013;45:1160-1167.

62. Aznarez I, Chan EM, Zielenski J, Blencowe BJ, Tsui L-C. Characterisation of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 14 of the cystic fibrosis transmembrane conductance regulator gene. *Hum Mol Genet* 2003;12:2031-2040.

63. Mid-year population estimates, 2015. Available at: http://www.statssa.gov.za/publications/P0302/P03022015.pdf. Accessed July 28, 2015.

64. *Community genetics services*. Geneva: World Health Organisation; 2010.

65. Li W, Sun L, Corey M, et al. Understanding the population structure of North American patients with cystic fibrosis. *Clin Genet* 2011;79:136-146.

66. Gurdarsani D, Carstensen T, Tekola-Ayele F, et al. The African Genome Variation Project shapes medical genetics in Africa. *Nat* 2015;517:327-332.

67. Ramalho AS, Beck S, Meyer M, Penque D, Cutting GR, Amaral MD. Five percent of normal cystic fibrosis transmembrane conductance regulator mRNA ameliorates the severity of pulmonary disease in cystic fibrosis. *Am J Respir Cell Mol Biol* 2002;27:619-627.

68. Tishkoff SA, Kidd KK. Implications of biogeography of human populations for "race" and medicine. *Nat Genet* 2004;36:S21-S27.

69. Reed FA, Tishkoff SA. African human diversity, origins and migrations. *Curr Opin Genet Dev* 2006;16:597-605.

70. Campbell MC, Tishkoff SA. African Genetic Diversity: implications for human demographic history, modern human origins and complex disease mapping. *Annu Rev Genomics Hum Genet* 2008;9:403-433.
71. *Services for the prevention and management of genetic disorders and birth defects in developing countries*. Geneva: World Health Organisation; 1999
72. *Community approaches to the control of hereditary diseases: Report of a WHO Expert Advisory Group*. Geneva: World Health Organisation; 1985.
73. *Sixty-Third World Health Assembly WHA63.17: Birth defects*. Geneva: World Health Organisation; 2010.
74. United Nations Convention on the Rights of the Child. Available at: http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx. Accessed August 6, 2015.

Supplementary material

Table S1. Cystic fibrosis mutations identified in Africa by nationality.

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
c.1418delG	Sequencing	74	?		Egyptian
c.3877G>A		74	?		Egyptian
c.2620-15C>G		74	?		Egyptian
c.3718-24G>A		74	?		Egyptian
T665S		16	1	6.25	Egyptian
ΔF508		16	5	31.25	Egyptian
1898+3A>C		16	1	6.25	Egyptian
N1303K		16	1	6.25	Egyptian
U		16	8	50.00	Egyptian
ΔF508	Kit, PCR-	80	1	1.25	Egyptian CBAVD
5T	RFLP, exon	80	14	17.50	Egyptian CBAVD
7T	11 and poly-T	80	57	71.25	Egyptian CBAVD
9T	tract	80	9	11.25	Egyptian CBAVD
	sequencing				
1609delCA	Exon 10 heteroduplex formation and sequencing	2	2	100.00	Algerian

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
ΔF508	ΔF508 screen,	68	25	36.76	Algerian
N1303K	DGGE &	20	4	20.00	Algerian
711+1G>T	PCR-RFLP	32	4	12.50	Algerian
1812-1G>A		20	1	5.00	Algerian
5T	DGGE &	12	2	16.67	Algerian
E1104X	DHPLC	12	1	8.33	Algerian
U		20	9	45.00	Algerian
ΔF508	Kit, DHPLC	716	344	48.04	Tunisian
G542X	& DGGE	704	49	6.96	Tunisian
W1282X		716	44	6.15	Tunisian
N1303K		676	41	6.07	Tunisian
711+1G>T		676	31	4.59	Tunisian
405+1G>A	Mutation screen	12	8	66.67	Tunisian Jews
2766del8	PAGE	540	10	1.85	Tunisian
E1104X	Kit, DHPLC & DGGE	680	32	4.71	Tunisian
G85E	Kit & DGGE	540	6	1.11	Tunisian
R74W		540	2	0.37	Tunisian
D1270N		540	2	0.37	Tunisian

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
Y122X	Kit & DGGE	540	1	0.19	Tunisian
T665S		540	1	0.19	Tunisian
R1066C		540	1	0.19	Tunisian
F1166C		540	1	0.19	Tunisian
L1043R		540	1	0. 19	Tunisian
V201M		540	1	0.19	Tunisian
U		680	110	16.18	Tunisian
3729delAinsTCT	DGGE &	4	2	50.00	Tunisian
N	DHPLC	4	1	25.00	Tunisian
I1203V	PCR-RFLP,	136	2	1.47	Tunisian
R1158X	DGGE,	136	2	1.47	Tunisian
4268+2T>G	DHPLC &	136	2	1.47	Tunisian
R785X	ΔF508 screen	136	1	0.74	Tunisian
4016insT		136	1	0.74	Tunisian
1811+5A>G		136	1	0.74	Tunisian
5T	11 exon	28	3	10.71	Tunisian
V201M	screen,	28	4	14.29	Tunisian
	DGGE,				
	DHPLC				
D579G	?	6	1	16.67	Sudanese

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
R1102K	?	6	1	16.67	Sudanese
ΔF508	ΔF508 screen,	20	6	30.00	Libyan
E1104X	DGGE,	20	8	40.00	Libyan
c.1670delC	DHPLC	20	2	10.00	Libyan
N1303K		20	1	5.00	Libyan
U		20	3	15.00	Libyan
ΔF508	32 mutation	334	26	7.78	Moroccan
5T	kit	300	8	2.67	Moroccan
12TG		300	4	1.33	Moroccan
11TG		300	4	1.33	Moroccan
711+1G>T	Kit & exon	2	1	50.00	Moroccan
R74W	sequencing	2	1	50.00	Moroccan
R1070W		2	1	50.00	Moroccan
D1270N		2	1	50.00	Moroccan
ΔF508	Kit, SSCP,	34	24	70.59	Moroccan Jews
3849+10kbC>T	exon	34	1	2.94	Moroccan Jews
S549R	sequencing	34	4	11.76	Moroccan Jews
G1244E		34	1	2.94	Moroccan Jews
U		34	3	8.82	Moroccan Jews
F693L	DHPLC,	120	2	1.67	Rwandan

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
	sequencing				
3120+1G>A	DHPLC,	120	1	0.83	Rwandan
A204T	sequencing	120	1	0.83	Rwandan
c.4575+2G>A		120	1	0.83	Rwandan
T854T		120	52	43.33	Rwandan
Q1463Q		120	14	11.67	Rwandan
M470V		120	13	10.83	Rwandan
c.1898+152T>A		120	9	7.50	Rwandan
P1290P		120	6	5.00	Rwandan
c.1001+11C>T		120	5	4.17	Rwandan
E527E		120	3	2.50	Rwandan
c.2752-15C>G		120	3	2.50	Rwandan
c.3041-71A>G		120	2	1.67	Rwandan
c.3272-32T>C		120	1	0.83	Rwandan
U		120	7	5.83	Rwandan
405+4A>G	DGGE, SSCP, ASO	122	1	0.82	Cameroonian
Y1109X	DGGE,	2	2	100.00	Cameroonian
EX17a-EX18del	sequencing &	2	2	100.00	Senegalese
c.4136+1G>A	PCR-RFLP	2	1	50.00	Senegalese

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
U		2	1	50.00	Senegalese
ΔF508	?	12	8	66.67	Namibian
ΔF508	ΔF508 screen	772	553	71.63	South African
3120+1G>A	DGGE, PCR-RFLP & sequencing	602	81	13.46	South African
G1249E	DGGE,	34	2	5.88	South African
3196del54	sequencing	34	2	5.88	South African
D1270N	PCR-RLP, kit	18	2	11.11	South African
3272-26A>G	PCR, ASO,	488	17	3.48	South African
394delTT	ARMS, RFLP	402	15	3.73	South African
G542X		488	9	1.84	South African
R553X		402	4	1.00	South African
W1282X		402	4	1.00	South African
N1303K		402	4	1.00	South African
G551D		488	5	1.02	South African
R117H		402	1	0.25	South African
Q493X		402	1	0.25	South African
S549N		402	1	0.25	South African
621+1G>T		402	1	0.25	South African

Mutation	Method of Identification	# of alleles tested	# + alleles	Allele Frequency/ %	Nationality
1717-1G>A		402	1	0.25	South African
2789+5G>A		402	1	0.25	South African
3659delC		402	1	0.25	South African
R1162X		402	1	0.25	South African
-94G>T		28	1	3.57	South African
2183delAA		28	1	3.57	South African
c.54-1161_c.164+1603del2875	MLPA	2	1	50.00	South African
U	PCR, ASO, ARMS, RFLP	536	83	15.49	South African
3120+1G>A	MLPA	2	1	50.00	Zimbabwean
c.54-1161_c.164+1603del2875	MLPA	2	1	50.00	Zimbabwean

Mutations in red have been empirically shown to cause CF, mutations in purple were reported as novel in the original publication, mutations in blue have varying clinical consequence, mutations highlighted in yellow are unique to Africa. ? = number of alleles was not reported, U = unknown mutations, N = normal CFTR allele. Sequencing = Sanger sequencing; kit = commercial screening diagnostic kit; PCR-RFLP = polymerase chain reaction restriction fragment length polymorphism; DGGE = denaturing gradient gel electrophoresis; DHPLC = denaturing high performance liquid chromatography; PAGE = polyacrylamide gel electrophoresis; SSCP = single strand conformation polymorphism; ASO = allele specific oligonucleotide hybridisation; ARMS = amplification refractory mutation system; MLPA = multiplex ligation-dependent probe amplification.

Table S2. Cystic fibrosis mutations in Africa by region.

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
North Africa – 1 334 chromosomes tested				
c.1418delG	74	?		Egyptian
c.3877G>A	74	?		Egyptian
c.2620-15C>G	74	?		Egyptian
c.3718-24G>A	74	?		Egyptian
T665S	556	2	0.36	Egyptian, Tunisian
ΔF508	1 268	456	35.96	Egyptian, Egyptian (CBAVD), Algerian, Tunisian, Libyan, Moroccan, Moroccan Jews
1898+3A>C	16	1	6.25	Egyptian
N1303K	732	47	6.42	Egyptian, Algerian, Tunisian, Libyan
5T	420	27	6.43	Egyptian (CBAVD), Algerian, Tunisian, Moroccan
7T	80	57	71.25	Egyptian (CBAVD)
9T	80	9	11.25	Egyptian (CBAVD)

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
1609delCA	2	2	100.00	Algerian
711+1G>T	170	36	21.18	Tunisian, Algerian, Moroccan
1812-1G>A	20	1	5.00	Algerian
E1104X	712	41	5.76	Algerian, Tunisian, Libyan
G542X	704	49	6.96	Tunisian
W1282X	716	44	6.15	Tunisian
405+1G>A	12	8	66.67	Tunisian Jews
2766del8	540	10	1.85	Tunisian
G85E	540	6	1.11	Tunisian
R74W	542	3	0.55	Tunisian, Moroccan
D1270N	542	3	0.55	Tunisian, Moroccan
Y122X	540	1	0.19	Tunisian
R1066C	540	1	0.19	Tunisian
F1166C	540	1	0.19	Tunisian
L1043R	540	1	0.19	Tunisian
V201M	568	5	0.88	Tunisian
3729delAinsTCT	4	2	50.00	Tunisian

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
N	4	1	25.00	Tunisian
I1203V	136	2	1.47	Tunisian
R1158X	136	2	1.47	Tunisian
4268+2T>G	136	2	1.47	Tunisian
R785X	136	1	0.74	Tunisian
4016insT	136	1	0.74	Tunisian
1811+5A>G	136	1	0.74	Tunisian
D579G	6	1	16.67	Sudanese
R1102K	6	1	16.67	Sudanese
c.1670delC	20	2	10.00	Libyan
12TG	300	4	1.33	Moroccan
11TG	300	4	1.33	Moroccan
R1070W	2	1	50.00	Moroccan
3849+10kbC>T	34	1	2.94	Moroccan Jews
S549R	34	4	11.76	Moroccan Jews
G1244E	34	1	2.94	Moroccan Jews
U	770	133	17.27	Libyan, Tunisian, Tunisian Jews, Algerian, Egyptian, Moroccan Jews

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
East Africa – 120 chromosomes tested				
F693L	120	2	1.67	Rwandan
3120+1G>A	120	1	0.83	Rwandan
A204T	120	1	0.83	Rwandan
c.4575+2G>A	120	1	0.83	Rwandan
T854T	120	52	43.33	Rwandan
Q1463Q	120	14	11.67	Rwandan
M470V	120	13	10.83	Rwandan
c.1898+152T>A	120	9	7.50	Rwandan
P1290P	120	6	5.00	Rwandan
c.1001+11C>T	120	5	4.17	Rwandan
E527E	120	3	2.50	Rwandan
c.2752-15C>G	120	3	2.50	Rwandan
c.3041-71A>G	120	2	1.67	Rwandan
c.3272-32T>C	120	1	0.83	Rwandan
U	120	7	5.83	Rwandan
Central Africa – 124 chromosomes tested				
405+4A>G	122	1	0.82	Cameroonian
Y1109X	2	2	100.00	Cameroonian

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
West Africa – 6 chromosomes tested				
EX17a-EX18del	2	2	100.00	Senegalese
c.4136+1G>A	2	1	50.00	Senegalese
U	2	1	50.00	Senegalese
Southern Africa – 760 chromosomes tested				
ΔF508	784	561	71.56	Namibian, South African
3120+1G>A	603	82	13.60	South African, Zimbabwean
G1249E	34	2	5.88	South African
3196del54	34	2	5.88	South African
D1270N	18	2	11.11	South African
3272-26A>G	488	17	3.48	South African
394delTT	402	15	3.73	South African
G542X	488	9	1.84	South African
R553X	402	4	1.00	South African
W1282X	402	4	1.00	South African
N1303K	402	4	1.00	South African
G551D	488	5	1.02	South African
R117H	402	1	0.25	South African

Mutation	# of alleles tested	# + alleles	Allele Frequency/%	Nationality
Q493X	402	1	0.25	South African
S549N	402	1	0.25	South African
621+1G>T	402	1	0.25	South African
1717-1G>A	402	1	0.25	South African
2789+5G>A	402	1	0.25	South African
3659delC	402	1	0.25	South African
R1162X	402	1	0.25	South African
-94G>T	28	1	3.57	South African
2183delAA	28	1	3.57	South African
c.54-1161_c.164+1603del2875	4	2	50.00	South African, Zimbabwean
U	536	83	15.49	South African
Africa – 2 344 chromosomes tested				

Mutations in red have been empirically shown to cause CF, mutations in purple were reported as novel in the original publication, mutations in blue have varying clinical consequence, mutations highlighted in yellow are unique to Africa. ? = number of alleles was not reported, U = unknown mutations, N = normal CFTR allele.