The Journal of Molecular Diagnostics, Vol. 18, No. 1, January 2016

the Journal of Molecular Diagnostics
jmd.amjpathol.org

The Spectrum of *CFTR* Variants in Nonwhite Cystic Fibrosis Patients

Implications for Molecular Diagnostic Testing

Iris Schrijver,*† Lynn Pique,* Steve Graham,‡ Michelle Pearl,§ Athena Cherry,*† and Martin Kharrazi¶

From the Departments of Pathology and Pediatrics,† Stanford University Medical Center, Stanford; the Genetic Disease Screening Program,‡ and the Environmental Health Investigations Branch,¶ California Department of Public Health, Richmond; and the Sequoia Foundation,§ La Jolla, California*

Accepted for publication July 7, 2015.

Address correspondence to Iris Schrijver, M.D., Department of Pathology, Stanford University Medical Center, L235, 300 Pasteur Dr., Stanford, CA 94305. E-mail: ischrijver@stanfordhealthcare.org.

Despite the implementation of cystic fibrosis (CF) newborn screening programs across the United States, the identification of *CFTR* gene variants in nonwhite populations compared with whites remains suboptimal. Our objective was to establish the spectrum of *CFTR* variants and their frequencies in CF patients in the United States with African, Native American, Asian, East Indian, or Middle Eastern backgrounds. By using direct DNA sequencing, we identified two *CFTR* variants in 89 of 140 affected nonwhite individuals with uncharacterized genotypes. Seven variants were novel. Multiplex ligation-dependent probe amplification detected 14 rearrangements in the remaining 51 patients, 6 of which were novel. Deletions and duplications accounted for 17% of unidentified alleles. A cross-sectional analysis of genotyping data from the CF Foundation Patient Registry was performed, comparing 3496 nonwhite patients with 22,206 white CF patients. Patients of Hispanic, black, or Asian ancestry were less likely to have two identified *CFTR* variants ($P < 0.0001$ for Hispanics and blacks, $P = 0.003$ for Asians), and more likely to carry no mutations on the commonly used 23 mutation carrier screening panel ($P < 0.0001$). We analyzed the mutations recorded for each ancestry and summarized the most frequent ones. This research could facilitate equity in mutation detection between white and nonwhite or mixed-ethnicity CF patients, enabling an earlier diagnosis improving their quality of life. *(J Mol Diagn 2016, 18: 39–50; http://dx.doi.org/10.1016/j.jmoldx.2015.07.005)*

Cystic fibrosis (CF; Online Mendelian Inheritance in Man no. 219700, *http://www.ncbi.nlm.nih.gov/omim*) is one of the most frequent autosomal-recessive conditions. CF has an overall birth prevalence of 1:3500 individuals in the United States.[1] It is most common among non-Hispanic whites (approximately 1:2500) and Ashkenazi Jews (approximately 1:2270) and, consequently, these two populations have been studied the most extensively.[2–4] Among nonwhite populations, CF is less frequent. In the United States, CF occurs in approximately 1:15,000 blacks, 1:35,000 individuals of Asian descent, and 1:10,900 Native Americans.[4,5] CF affects the exocrine epithelial cells of multiple tissues and organs, including the respiratory tract, the pancreas, the intestine, the male genital tract, the hepatobiliary system, and the sweat glands.[6] Morbidity and mortality in CF are attributed most commonly to pulmonary disease, characterized by chronic lung infections and airway inflammation. Other common clinical manifestations are failure to thrive, pancreatic insufficiency, meconium ileus, and infertility resulting from a congenital bilateral absence of the vas deferens.

CF is caused by mutations in the 27-exon *CFTR* gene (Online Mendelian Inheritance in Man no. *602421, *http://www.ncbi.nlm.nih.gov/omim*) that encodes the cystic fibrosis transmembrane conductance regulator (CFTR), a 1480−amino acid protein that forms chloride ion channels in the apical epithelial

Supported by clinical research grants from Cystic Fibrosis Foundation Therapeutics, Inc. (SCHRIJ11A0-NOA47100), and by the CF Foundation, which provided access to data in the CF Foundation Registry.

Copyright © 2016 American Society for Investigative Pathology and the Association for Molecular Pathology.
Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmoldx.2015.07.005

cell membrane.[7–9] When CFTR production is disrupted, the resulting abnormal electrolyte transport contributes to CF pathogenesis. Defective ion transport in the respiratory tract, in particular, leads to depleted airway surface liquid and increased mucosal obstruction.[10,11] Currently, there are more than 2000 described *CFTR* sequence variants (*http://www.genet.sickkids.on.ca/cftr/StatisticsPage.html*; last accessed June 19, 2015) and these are distributed throughout the gene, but to date only a subset have been firmly established as pathogenic by empiric analyses.[12] The majority of *CFTR* variants are point mutations or other small sequence changes, however, up to 2% of CF alleles likely are gene rearrangements, including large deletions, insertions, and duplications.[13] The most common *CFTR* mutation, p.Phe508del (delF508 by legacy nomenclature), accounts for approximately 66% of identified mutant alleles worldwide.[14] The spectrum and frequency of individual *CFTR* variants, however, vary relative to specific ethnic groups and geographic locations.[4,14,15] For instance, the c.3120+1G>A variant, although uncommon in non-Hispanic whites, is the second most frequent CF allele among black individuals, occurring at a frequency of 10% to 12%.[16,17] Ignoring this single, overall relatively rare allele in, for example, a newborn CF program in a state with black constituents would lower the detection rate for this population and could result in delayed diagnoses.

A clinical diagnosis of CF usually is based on two criteria: the presence of at least one distinctive clinical feature and laboratory evidence of CFTR dysfunction, typically an increased sweat Cl^- concentration.[18] A diagnosis can be made much more rapidly by incorporation of molecular testing and the identification of two *CFTR* mutations.[1] Thus, the diagnosis increasingly is expedited by molecular analysis that can be applied to both symptomatic and presymptomatic patients. Symptomatic identification of CF patients (excluding newborns with meconium ileus or a pre-existing family history) on average delays the diagnosis until 14.5 months of age, resulting in postponed treatment and significant compromises to clinical status.[19,20]

To date, knowledge of the spectrum of *CFTR* variants in nonwhite patients has remained limited. As a result, CF has been diagnosed at a later age among several nonwhite groups compared with whites[19–22] and there likely remains an inequitable identification of CF variants, despite the implementation of newborn screening in all states. Nonwhite patients may be diagnosed more frequently based on symptoms, rather than through a newborn screening algorithm or molecular diagnostic testing because the testing panels in use do not sufficiently include the variants that are prevalent in nonwhite populations. To determine which *CFTR* alleles are prevalent among nonwhite CF patients, we recruited and comprehensively genotyped 140 CF probands of African, Asian, Native American, East Indian, or Middle Eastern ancestry whose molecular etiology had not been fully characterized. Combining our results with genotype data from the CF Foundation Patient Registry, we identified the individual CF variants present in affected nonwhites and estimated their frequencies. Finally, we assessed the proportion of nonwhite patients who would not be identified by commonly used mutation analyses.

Materials and Methods

Study Subjects

Eligible participants included nonwhite CF patients with 0 or 1 identified *CFTR* variants(s) (including patients without genotype testing information) who were enrolled in the CF Foundation Patient Registry via 161 CF Centers in the United States. We excluded variants known to be benign but did not aim to predict or assign pathogenicity to the identified sequence changes. Rather, our focus was to characterize the spectrum of *CFTR* sequence variants. For the purposes of this study, nonwhite patients were those who self-identified as black, Asian, Native American, East Indian, or Middle Eastern; white and Hispanic CF patients were not eligible for the study. CF patients reporting more than one race/ethnicity were not excluded from participation unless they were of mixed white and Hispanic descent. Patient recruitment began in 2009 and, based on nonidentifiable data from the CF Foundation Patient Registry and our criteria, 528 living nonwhite CF patients were eligible for participation. Patients were recruited through the CF Centers providing care to these 528 patients. The study coordinator used mail, e-mail, and telephone calls to Center Directors to invite eligible patients identified by Center staff to obtain free *CFTR* sequencing at the clinical Stanford Molecular Pathology laboratory. Over the 4-year study period, 140 patients were enrolled. Results from clinical testing were reported back to the Center and patient. Center staff were expected and reminded by study staff to add testing results into the CF Foundation Patient Registry for later analysis.

Genomic DNA Amplification and CFTR Sequencing

Whole blood was collected via venipuncture at the CF Center and sent to the clinical diagnostic Stanford Molecular Pathology laboratory, where DNA extraction, amplification, and *CFTR* sequencing were performed. Genomic DNA was isolated from peripheral blood using standard procedures. All 27 exons of the *CFTR* gene (*http://www.ncbi.nlm.nih.gov/genbank*, GenBank accession number NG_016465.1) and noncoding regions in which mutations are known to exist [5′ untranslated region, at least 20 bp on each side of an exon, and intervening sequence (IVS)12 and 22 (IVS11 and 19 by legacy numbering)] were PCR-amplified using primer pairs from flanking intronic sequences as originally described[7,23,24] or as modified where needed[25] (and unpublished data). Amplified products were purified using either the Qiaquick PCR Purification Kit or the Qiaquick Gel Extraction Kit (Qiagen, Valencia, CA) according to the

ARTICLE IN PRESS

manufacturer's instructions. Products then were sequenced with fluorescent di-deoxy terminators (Life Technologies, Grand Island, NY) and electrophoresed on an ABI 3100 or 3130xl genetic analyzer (Life Technologies). Mutation Surveyor DNA Variant Analysis software (SoftGenetics, LLC, State College, PA) assisted in the detection of sequence changes and in the assessment of overall sequence quality. Identified variants were confirmed by sequencing in the opposite direction.

Detection and Characterization of CFTR Rearrangements

Residual, de-identified DNA samples with only one or no variant(s) identified by sequencing were analyzed, as part of an Institutional Review Board–approved research protocol, for exon rearrangements by multiplex ligation-dependent probe amplification (MLPA), using the SALSA MLPA probe mix P091-D1 CFTR (MRC-Holland, Amsterdam, The Netherlands) and following the manufacturer's protocol. Samples with potential homozygous changes and a homozygous pattern of single-nucleotide polymorphisms also were analyzed by MLPA for the presence of large deletions involving one or more exons. Capillary electrophoresis of amplified MLPA products was performed using an ABI 3500 Genetic Analyzer (Life Technologies), and the resulting data were analyzed using GeneMarker 1.51 software (SoftGenetics, LLC).

For samples whose MLPA results indicated either a deletion of exons 2 to 3 or a deletion of exons 17a to 18, previously published primers were used to confirm the presence of the c.54-5940_273+10250del21kb deletion and the c.2988+1173_c.3468+2111del8898 deletion, respectively.[26,27] An alternative primer pair (forward, 5′-GATGGAGTCTCACTCTGTTG-3′; reverse, 5′-GACACTGTCTTTCTTTTCTGTG-3′) was developed to confirm the exon 17a to 17b deletion (c.2988+1615_3367+357del3796ins62bp[28]). A short 417-bp fragment flanking the deletion breakpoint was amplified using MyTaq DNA Polymerase (Bioline, London, UK) and the following cycling conditions: 95°C for 3 minutes, 95°C for 15 seconds, 57°C for 15 seconds, and 72°C for 10 seconds for 35 cycles. These conditions do not amplify the corresponding 4151-bp region of the unaffected allele.

Novel rearrangements were characterized by long-range PCR using the Expand Long Template PCR System 2 for 9- to 12-kb fragments (Roche Applied Systems, Indianapolis, IN), 100 ng of genomic DNA, and a series of walking primer pairs with a walking distance of approximately 1 kb. For putative deletions, primers were designed upstream and downstream of the missing exons; primers for suspected duplications, however, were designed to amplify across the duplication junction. Amplified products were purified using either the Qiaquick PCR Purification Kit or the Qiaquick Gel Extraction Kit (Qiagen), and then sequenced with walking primers and electrophoresed on an ABI 3730xl sequencing instrument (Life Technologies). For each novel deletion and duplication, a PCR-based detection method was developed. Short products (<500 bp) were amplified using AmpliTaq Gold Polymerase (Life Technologies) and primer pairs flanking the junction points (Table 1). Amplicons were sequenced in both the forward and reverse directions to confirm rearrangement breakpoints.

Array comparative genomic hybridization was performed to examine further the extent of the rearrangements in probands NCCF-28 and NCCF-84 and to aid in the design of walking primers for subsequent long-range PCR analysis. Briefly, 0.5 μg of patient- and sex-matched control DNA were digested for 2 hours with Alu I and Rsa I at 37°C. The digested DNA was purified and random-prime labeled with Cy3 or Cy5 dyes (Enzo Life Sciences, Farmingdale, NY). The labeled products were purified and mixed with Cot-1 DNA, blocking agent, and hybridization buffer, denatured at 95°C, and hybridized to the Human Genome 180 k whole-genome oligonucleotide arrays (GRCh37/hg19 assembly; Agilent Technologies, Santa Clara, CA; Illumina, San Diego, CA) at 65°C for 24 hours. The slides were washed and then scanned on an Agilent Technologies scanner and the raw signals were processed and analyzed by BlueFuse Multi software (Illumina).

CF Foundation Patient Registry Data Analysis

Because this study was designed to improve *CFTR* genotyping of nonwhites with CF, we compared the number of current CF patients with two or more variants identified in the 2008 and 2013 CF Foundation Patient Registry. "Current CF patients" were defined as those who were alive and who were seen at a CF Center within the last year of the data set. The proportion of individuals who were fully genotyped (with two identified variants) versus those who were not (with no or one identified variant) was calculated for CF patient populations in the following racial/ethnic groups: white, Hispanic, black, Asian, and Native American. Middle Eastern individuals could not be separated within this comparison because in 2013 they were no longer distinguished as a separate category by the Patient Registry. The *CFTR* variant spectrum in nonwhites with CF was determined by comparing the genotype relative frequencies of individuals with zero, one, or two copies of the c.1521_1523delCTT, p.Phe508del mutation in the current 2013 CF patient population. The allelic frequencies for total CF patients in the 2013 CF Foundation Registry also were calculated and a list of the 50 most frequent DNA variants was compiled for each of the racial/ethnic groups listed earlier. The total CF patients included all patients in the CF Foundation Registry, living and deceased. We further assessed the number of CF patients across the ethnic ancestry groups in the current 2013 Registry population who would not be identified by the 23 mutation carrier screening panel[3,17] that often is used for diagnostic testing. In the aforementioned analyses known benign variants were excluded.

By using SAS version 12.1 statistical software (SAS Institute, Inc., Cary, NC), the CF Foundation Patient

Table 1 Primers for PCR-based Detection of Novel *CFTR* Rearrangements

Exon(s)	Type	Position	Primers	Amplicon size (bp)
1, 2	Del	c.-17,310_165-2332del43,916insAG	F: 5′-GGAAGAGAGGAAGTCAGATTG-3′ R: 5′-GTCACAAGGCAGTTATGAAATG-3′	380
4, 7	Del	c.274-2973_1116+122del12,541insCTTT	F: 5′-AACTACAAACCACTGCTCAAG-3′ R: 5′-GCTAGCTACATCAGTATTATTG-3′	229
6b-8	Dup	c.744-272_1209+3261dup9,094	F: 5′-GATTCAATTATCTCCCACTGG-3′ R: 5′-GTTTGGTAAATGCCTCCTATG-3′	228
11	Del	c.1585-6285_1679+741del7120	F: 5′-CCAACTTCAACAATAGTACATC-3′ R: 5′-CCTGGTTCAAACTGTGACTC-3′	479
19-24	Dup	chr7:117,266,922-117,760,437dup493,515	F: 5′-ACATCTTAGACTGGAGTTCTC-3′ R: 5′-AAGTCATTGGCATACATCTGC-3′	405
24	Del	c.4242+724_4443+1766del2585ins12	F: 5′-GAATTCAAGACCAGCCCAGG-3′ R: 5′-GTGGTTGCCAGGCATTAGAG-3′	301

Exons are listed according to legacy numbering.

Del, deletion; Dup, duplication; F, forward; R, reverse.

Registry database was analyzed and the frequencies were calculated, by ethnic origin, for the number of sequence variants identified per individual, for each genotype group, for sequence variants identified by the 23-mutation panel, and for the overall distribution of sequence changes. To determine whether distribution differences exist based on ethnic origin in the number of *CFTR* variants identified, in genotype groups, and in the results of the 23-mutation panel testing, χ^2 statistics and *P* values were calculated using "white only" as the reference group.

Results

CFTR Sequencing and Deletion/Duplication Analysis

Of the approximately 500 eligible nonwhite CF patients in the CF Foundation Patient Registry database with none or one *CFTR* variant(s) instead of the expected two, 140 probands were enrolled in our study. Of these probands, sequencing revealed two *CFTR* sequence variants in 89. However, in 25 probands only a single sequence variant was identified, and in 26 probands none were detected. We then specifically investigated the prevalence of deletion and duplication rearrangements, which are not detectable by Sanger sequencing. The 51 patients with incomplete genotypes were analyzed by MLPA and known deletions or duplications were confirmed by PCR. Novel deletion/duplication breakpoints were fully characterized by a walking PCR technique with subsequent sequencing and, in two cases with very large rearrangements, by array comparative genomic hybridization. MLPA testing identified 14 deletion or duplication variants in 12 probands: 2 probands (NCCF 84 and NCCF 145) each carried two deletions, but 1 of these patients (NCCF 84) already had a pathogenic mutation identified by sequencing (c.1521_1523delCTT, p.Phe508del). Thus, only 13 of 14 rearrangements accounted for unidentified alleles (13 of 77; 16.8%) (Table 2). Nine distinct variants were detected: seven represented deletions and two were duplications. Six of the gene rearrangements were novel (Table 2) and included one simple deletion, three complex deletions containing short (2 to 12 bp) insertions, one duplication, and a multigenic duplication of *CFTR* exons 19 to 24 that extended beyond the neighboring *CTTNBP2* gene on chromosome 7 (Figure 1). Two previously reported rearrangements—a deletion of exons 2 to 3 and a complex deletion of exons 17a to 17b that included a 62-bp insert—were detected in three and four probands, respectively (approximately 4% and approximately 5% of unidentified alleles). Two variants were identified in 101 of 140 (72%) probands overall, whereas in 39 the genotypes remained incomplete. Of these, 14 individuals (10%) had a single variant identified and 25 (17.8%) had none.

CF Patients with Two CFTR Variants in 2013 Compared with 2008

We compared the number of variants identified (0, 1, or 2) by ancestry in the current CF populations of 2008 and 2013, recorded in the CF Foundation Patient Registry. We used the concept of the current population, which included CF patients from 2008 or 2013 who were currently alive and seen at a CF Center in that year, to determine the potential population for genotyping. The comparison showed considerable progress between these years. In 2008 versus 2013, the percentage of whites (self-declared white-only ethnicity) not yet genotyped (individuals with zero or one identified variant) was 21% versus 9%, respectively, and for Hispanics (self-declared Hispanic-only ethnicity) was 29% versus 15%. Thirty-nine percent versus 19% of individuals with black, 29% versus 16% with Asian, and 19% versus 8% with Native American ancestry remained without two identified variants in those years (Table 3). Overall, between 2008 and the end of 2013, the total number of individuals whose genotype still needed to be elucidated was reduced to approximately half, aided substantially by the free *CFTR* sequencing provided by this study. Despite the overall decrease in the number of individuals without an identified molecular etiology, χ^2 analysis showed statistically significant differences between whites on the one hand and

Table 2 *CFTR* Rearrangements Identified in 12 Probands

Proband	Exon(s)	Type	Position	Size (kb)	Reference	Additional variant(s)	Ethnicity (as self-declared)
NCCF 7	2, 3	Del	c.54-5940_273+10250del21kb	21.1	26	Het p.Phe508del	Native American + unknown
NCCF 17	17a-18	Del	c.2988+1173_c.3468+2111del8898	8.9	27	Het p.G542X	Middle Eastern + unknown
NCCF 20	2, 3	Del	c.54-5940_273+10250del21kb	21.1	26	Het p.Phe508del	African American + white
NCCF 28	**1, 2**	**Del**	**c.-17,310_165-2332del43,916insAG**	**43.9**	**This study**	Het c.2988+1G>A (3120+1G>A)	African American
NCCF 4033643	**6b-8**	**Dup**	**c.744-272_1209+3261dup9094**	**9.1**	**This study**	Het p.R75X	Chinese
NCCF 74	**24**	**Del**	**c.4242+724_4443+1766del2585ins12bp**	**2.6**	**This study**	Het c.1393-1G>A(1525-1G>A)	Middle Eastern
NCCF 79	2, 3	Del	c.54-5940_273+10250del21kb	21.1	26	Het p.Phe508del	American Indian
NCCF 84	17a-17b	Del	c.2988+1615_3367+357del3796ins62bp	3.8	28	Het p.Phe508del	African American + white
	19-24	**Dup**	**chr7:117,266,922-117,760,437dup493,515**	**493.5**	**This study**		
NCCF 105	**4-7**	**Del**	**c.274-2973_1116+122del12541insCTTT**	**12.5**	**This study**	Het p.Phe508del	African American
NCCF 106	17a-17b	Del	c.2988+1615_3367+357del3796ins62bp	3.8	28	Het p.G551D	African American
NCCF 123	17a-17b	Del	c.2988+1615_3367+357del3796ins62bp	3.8	28	Het p.R553X	African American + white
NCCF 145	**11**	**Del**	**c.1585-6285_1679+741del7120**	**7.1**	**This study**		African American
	17a-17b	Del	c.2988+1615_3367+357del3796ins62bp	3.8	28		

Note: Exons are listed according to legacy numbering. Mutations in parentheses are according to legacy numbering for additional clarity. Bold type shows rearrangements newly characterized in this study.

Del, deletion; Dup, duplication.

Hispanics, blacks, and Asians on the other (Table 3). Interestingly, CF patients of Native American origin closely tracked the percentages of whites with identified genotypes in all of our analyses.

Genotype Frequency

We determined the type and frequency of genotypes by ethnicity and ancestry in the current 2013 CF population (Table 4) and observed that, although genotypes that include the c.1521_1523delCTT, p.Phe508del mutation strongly predominate in whites and in patients with Native American ancestry, affected individuals of other origins have a substantial percentage of non-p.Phe508del genotypes. Only 10% of white CF patients carry no p.Phe508del mutation at all, compared with 17%, 30%, 38%, and 40% of those with Native American, Hispanic, black, and Asian backgrounds. Thus, compared with whites, other ethnic groups were significantly more likely to carry no p.Phe508del mutations (Table 4).

The 23-mutation panel originally designed by the American College of Medical Genetics (now called the American College of Medical Genetics and Genomics)[3,17] was designed specifically for CF carrier screening in the general population, but it also is used for diagnostic and newborn screening purposes. With this panel, we found that the majority of white and Native American CF patients in the current 2013 CF population could be provided with a two-allele genotype, whereas less than half of the CF patients of other ethnicities would have their molecular etiology discovered. Conversely, the percentage of Hispanic, black, and Asian individuals who carry no mutations present on this panel is significantly larger than that in whites (Table 5).

The CFTR Variant Spectrum in White and Nonwhite US CF Patients

After the p.Phe508del mutation, which has the highest frequency in every analyzed ethnic group in the United States,

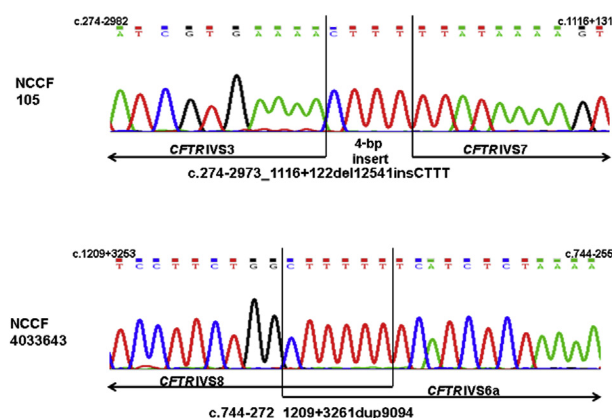

Figure 1 Newly characterized *CFTR* gene rearrangements. NCCF 105: This sequence depicts a novel complex deletion of exons 4 to 7 with a small, 4-bp insertion. NCCF 4033643: This sequence illustrates a novel duplication of exons 6b to 8. Exon listing is according to legacy numbering. *CFTR* rearrangements are relatively common in nonwhites and often are different from those in whites. As such, they are an important genotyping consideration in nonwhite CF patients.

the 49 most common sequence changes for each population were ranked in order of frequency (Table 6). Known benign changes were excluded, as were the various alleles of the polymorphic polyTG-T region in legacy IVS8 (intron 8) immediately preceding legacy exon 9 (IVS 9 and exon 10 in sequential numbering). We also excluded exon deletions and duplications because we could not verify the individual breakpoints for Registry entries. For this analysis, we used the total CF population in the CF Foundation Registry, which included all patients, living and deceased, in the Registry through the end of 2013. This population most completely captures the *CFTR* genotypes identified in individuals with nonwhite backgrounds. For comparison, the top 50 sequence changes for CF patients who declared white-only or Hispanic-only origins also were included. The variants of lowest frequencies in the nonwhite, non-Hispanic categories of Table 6 are not based on relative importance and, because we used an arbitrary cut-off value of 50 mutations, mutations of the same frequency are not all listed (unlisted variants are included in Supplemental Table S1). Variant number 50 (Table 6) corresponds to 38 patients in the white category, 9 patients in the Hispanic group, 6 patients in the black group, and 1 patient each in the Asian and Native American columns. In Hispanics, six variants were each seen in nine patients, and in black individuals eight variants were seen in six patients each. In blacks, one of these variants falls below the cut-off value of 50 listed variants. In Asians, 60 variants were seen in only one patient, and 37 of those were below the variant number 50. In Native Americans, 29 variants in the top 50 were seen in only one individual, whereas 12 additional ones were also at that frequency.

The bolded variants in Table 6 were present in a single ancestry group only. These variants may be present at very low frequencies in other ethnic groups and should not be interpreted as specific to any one ancestry because such conclusions cannot be drawn based on self-declared ethnicities and relatively low numbers in some groups. Nevertheless, this representation illustrates the differences in mutation frequencies between populations and especially highlights the divergence from the most common mutations identified in whites, who have 4 of 50 unique sequence changes among this list, compared with 20, 14, 18, and 7 of 50 in Hispanics, blacks, Asians, and Native Americans, respectively. By χ^2 analysis, these numbers are different from whites with the following P values: 0.001 for Hispanics, 0.02 for blacks, 0.002 for Asians, and 0.5 for Native Americans. The latter group had a spectrum quite

Table 3 Genotype Identification by Ethnic Ancestry in the CF Populations of 2008 and 2013

	White only		Hispanic only		Black, any		Asian, any		Native American, any	
Variants identified	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
2008										
2	17,327	79	1252	71	658	61	91	71	116	81
1	2194	10	254	15	209	19	17	13	15	10
0	2513	11	250	14	216	20	21	16	13	9
Total	22,034	100%	1756	100%	1083	100%	129	100%	144	100%
Individuals not yet genotyped	4707	21%	504	29%	425	39%	38	29%	28	19%
P value	NA		<0.0001		<0.001		0.02		0.58	
2013										
2	20,809	91	1738	85	1028	81	137	84	159	92
1	1208	5	175	9	136	11	15	9	9	5
0	903	4	126	6	98	8	11	7	4	2
Total	22,920	100%	2039	100%	1262	100%	163	100%	172	~100%
Individuals not yet genotyped	2111	9%	301	15%	234	19%	26	16%	13	8%
P value	NA		<0.0001		<0.0001		0.003		0.44	

Totals include patients who were never genotyped. Percentages are rounded and may not add up to exactly 100%, and are indicated, where applicable, by "~100%."

NA, not applicable.

ARTICLE IN PRESS

Table 4 Type and Frequency of *CFTR* Variants in the Genotyped CF Population of 2013 by Ethnic Ancestry

Genotype	White only		Hispanic only		Black, any		Asian, any		Native American, any	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
p.Phe508del/p.Phe508del	11,133	50	532	27	234	19	33	21	82	48
p.Phe508del/other	7889	36	729	37	456	38	49	31	55	32
p.Phe508del/unknown	886	4	106	5	66	5	11	7	5	3
Other/other	1776	8	477	24	337	28	55	35	22	13
Other/unknown	332	1	69	4	70	6	4	3	4	2
Unknown/unknown	190	1	42	2	51	4	4	3	3	2
Total	22,206	100%	1955	∼100%	1214	100%	156	100%	171	100%
0 p.Phe508del mutations	2298	10	588	30	458	38	63	40	29	17
1 p.Phe508del mutation	8775	40	835	43	522	43	60	38	60	35
2 p.Phe508del mutations	11,133	50	532	27	234	19	33	21	82	48
Total	22,206	100%	1955	100%	1214	100%	156	∼100%	171	100%
P value*	Reference group		<0.0001		<0.0001		<0.0001		0.005	

Percentages are rounded and may not add up to exactly 100%, and are indicated, where applicable, by "∼100%."
*Carrying 0 p.Phe508del mutations.

similar to whites. The Asian mutation spectrum is the most divergent from the other ethnic groups, even among the more common sequence changes. Of the 12 entries with frequencies greater than 1% in the Asian group, 4 were not present in the top 50 of any other ethnic category.

Novel CFTR Sequence Changes in Nonwhite US CF Patients

Among the total CF population in the CF Foundation Registry there were 34 previously unreported sequence changes identified in 49 CF patients that could be characterized unambiguously. Two of these had amino acid changes that each could result from one of two nucleotide changes, and therefore the cDNA numbering reflects this possibility (c.1695T>A or G for p.D565E and c.2885C>G or A for p.S962X) (Table 7). Of the 34 unique novel variants, 26 were identified in blacks, 8 in Asians, and 2 in Native Americans. One of the variants seen in an Asian individual also was present in a white-only patient in the Registry. Two other sequence variants were identified in multiple ethnicities: one in the black and Native American category, and one in the black and Asian category. The latter two could be assignment artifacts, however, because patients with any black, Asian, or Native American backgrounds were assigned to each of these categories, and someone with mixed ethnicity could be counted twice.

Eight sequence changes present in blacks were observed more than once, to the best of our knowledge these previously were unreported. Three of these variants (p.F17Lfs, c.51delC; c.579+4delT; and p.S962X, c.2885C>G/A) were seen at least three times, indicating that some of these changes may be present relatively frequently in that single group. In the 140 probands who were sequenced as part of our study, seven novel variants were identified in 7 patients (Table 7).

Discussion

In contrast to thoroughly analyzed white CF populations, the *CFTR* variant spectrum and prevalence in black, Asian, Native American, and Middle Eastern CF patients have not been elucidated completely. Such knowledge gaps can lead to racial-ethnic disparities in the clinical sensitivity of

Table 5 ACMG 23 Carrier Screening Panel Genotype Frequencies by Ethnic Ancestry in the Genotyped CF Population of 2013

Mutations on ACMG 23	White only		Hispanic only		Black, any		Asian, any		Native American, any	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
2	16,782	76	933	48	454	37	48	31	116	68
1	4775	22	712	36	536	44	64	41	42	25
0	649	3	310	16	224	18	44	28	13	8
Total	22,206	∼100%	1955	100%	1214	∼100%	156	100%	171	∼100%
P value*	Reference group		<0.0001		<0.0001		<0.0001		0.0003	

Percentages are rounded and may not add up to exactly 100%, indicated, where applicable, by "∼100%."
*Carrying 0 mutations of the ACMG 23 mutation panel.
ACMG, American College of Medical Genetics and Genomics.

Table 6 The 50 Most Frequent *CFTR* Sequence Variants in US CF Patients

Numbering for sequence variants	White only *N* (patients) = 31,286, *N* (mutations) = 59,713		Hispanic only *N* (patients) = 2578, *N* (mutations) = 4816		Black, any *N* (patients) = 1551, *N* (mutations) = 2841		Asian, any *N* (patients) = 203, *N* (mutations) = 376		Native American, any *N* (patients) = 249, *N* (mutations) = 477	
	Variant	%	Variant	%	Variant	%	Variant	%	Variant	%
1	p.F508del	73.4	p.F508del	54.0	p.F508del	46.5	p.F508del	43.4	p.F508del	65.2
2	p.G551D	2.6	p.G542X	5.2	3120+1G>A	10.9	3849+10kbC>T	4.5	p.R1162X	4.4
3	p.G542X	2.3	3849+10kbC>T	2.0	p.A559T	2.4	p.S549N	4.0	3849+10kbC>T	3.4
4	p.R117H	1.8	p.R334W	1.8	**2307insA**	2.2	p.R334W	2.4	p.G542X	2.1
5	p.W1282X	1.5	**3876delA**	1.6	p.G542X	1.6	**p.V456A**	2.1	p.R117H	1.3
6	p.N1303K	1.4	p.N1303K	1.5	p.G551D	1.4	**p.R709X**	1.9	p.R75X	1.3
7	p.R553X	1.0	p.I507del	1.3	p.R553X	1.2	p.R347H	1.3	3659delC	1.0
8	621+1G>T	0.9	p.S549N	1.2	p.D1270N	0.9	**p.L218X**	1.3	3272-26A>G	1.0
9	1717-1G>A	0.9	3120+1G>A	0.9	p.S549N	0.8	p.G551D	1.1	p.G551D	0.8
10	3849+10kbC>T	0.8	**406-1G>A**	0.8	p.N1303K	0.8	p.W1282X	1.1	3120+1G>A	0.8
11	2789+5G>A	0.7	p.W1282X	0.8	p.R117H	0.7	p.R1066C	1.1	p.A455E	0.8
12	3659delC	0.4	**p.W1089X**	0.8	621+1G>T	0.7	**1525-1G>A**	1.1	p.R1066H	0.8
13	1898+1G>A	0.4	p.R1162X	0.7	p.R1162X	0.7	p.R1162X	0.8	p.R553X	0.6
14	p.I507del	0.4	p.R553X	0.7	p.G330X	0.7	p.A455E	0.8	p.N1303K	0.4
15	p.G85E	0.4	p.G551D	0.7	**p.R74W**	0.6	p.R1158X	0.8	1717-1G>A	0.4
16	p.R1162X	0.3	p.R75X	0.6	p.R1158X	0.6	**457TAT>G**	0.8	2789+5G>A	0.4
17	p.D1152H	0.3	p.D1152H	0.6	3120G>A	0.6	**p.Q1352H**	0.8	1898+1G>A	0.4
18	**p.R347P**	0.3	p.R117H	0.6	**p.S466X**	0.6	**p.A1319E**	0.8	p.R560T	0.4
19	2184insA	0.3	p.G85E	0.6	p.S1255X	0.6	p.G542X	0.5	1154insTC	0.4
20	p.R560T	0.3	1717-1G>A	0.6	**1812-1G>A**	0.5	p.R553X	0.5	394delTT	0.4
21	p.A455E	0.3	1811+1643G>T	0.6	3849+10kbC>T	0.5	1717-1G>A	0.5	p.R1158X	0.4
22	3272-26A>G	0.3	**2105-2117del13insAGAAA**	0.6	1717-1G>A	0.5	p.S945L	0.5	621+1G>T	0.2
23	p.Q493X	0.3	p.R1066C	0.5	**3791delC**	0.5	p.E585X	0.5	p.I507del	0.2
24	1154insTC	0.2	p.L206W	0.5	p.W1282X	0.4	**1677delTA**	0.5	p.G85E	0.2
25	**2184delA**	0.2	621+1G>T	0.5	p.I507del	0.4	**p.Y569D**	0.5	2184insA	0.2
26	p.E60X	0.2	2789+5G>A	0.5	p.R560T	0.4	**p.G622D**	0.5	p.Q493X	0.2
27	3905insT	0.2	**2055del9>A**	0.5	p.R1066H	0.4	**p.D979A**	0.5	p.E60X	0.2
28	2183del AA>G	0.2	**p.W1204X**	0.5	p.S434X	0.4	p.N1303K	0.3	2183del AA>G	0.2
29	**p.P67L**	0.2	**663delT**	0.5	1898+1G>A	0.3	3120+1G>A	0.3	p.S549N	0.2
30	p.R347H	0.2	**935delA**	0.4	p.R1066C	0.3	3659delC	0.3	p.R347H	0.2
31	p.R334W	0.2	p.S945L	0.4	p.L467P	0.3	3272-26A>G	0.3	p.Y1092X	0.2
32	394delTT	0.2	**3199del6**	0.4	p.Q98X	0.3	p.Q493X	0.3	p.V520F	0.2
33	p.L206W	0.2	p.D1270N	0.4	**405+3A>C**	0.3	3905insT	0.3	p.R117C	0.2
34	p.Y1092X	0.1	**1288insTA**	0.4	**p.Y913X**	0.3	p.L206W	0.3	p.G576A	0.2
35	**p.M1101K**	0.1	2183delAA>G	0.3	**p.F1099L**	0.3	p.Y1092X	0.3	p.R668C	0.2
36	p.V520F	0.1	**1811+1.6kbA>G**	0.3	p.R334W	0.3	3120G>A	0.3	2622+1G>A	0.2
37	p.S945L	0.1	**p.Q890X**	0.3	3272-26A>G	0.3	p.A559T	0.3	p.S1235R	0.2
38	711+1G>T	0.1	**712-1G>T**	0.3	**p.Y275X**	0.3	p.R75X	0.3	p.L997F	0.2
39	p.R117C	0.1	3272-26A>G	0.2	**p.G480C**	0.3	p.L997F	0.3	2789+2insA	0.2
40	p.G576A	0.1	p.Y1092X	0.2	2789+5G>A	0.2	1078delT	0.3	p.L467P	0.2
41	p.R1158X	0.1	**p.F311del**	0.2	p.G85E	0.2	**p.S492F**	0.3	**p.S549R**	0.2
42	2622+1G>A	0.1	**1248+1G>A**	0.2	p.E585X	0.2	p.S1255X	0.3	**2143delT**	0.2
43	p.R668C	0.1	**p.P205S**	0.2	**444delA**	0.2	**p.Q98R**	0.3	**2585delT**	0.2
44	3120G>A	0.1	**p.H199Y**	0.2	**p.F191V**	0.2	**p.R31C**	0.3	p.G330X	0.2
45	p.S1235R	0.1	711+1G>T	0.2	2184insA	0.2	p.S434X	0.3	**4005+2T>C**	0.2
46	p.R1066C	0.1	p.R352Q	0.2	p.Y1092X	0.2	**p.R851X**	0.3	**p.R1070W**	0.2

(*table continues*)

ARTICLE IN PRESS

Table 6 (*continued*)

Numbering for sequence variants	White only *N* (patients) = 31,286, *N* (mutations) = 59,713		Hispanic only *N* (patients) = 2578, *N* (mutations) = 4816		Black, any *N* (patients) = 1551, *N* (mutations) = 2841		Asian, any *N* (patients) = 203, *N* (mutations) = 376		Native American, any *N* (patients) = 249, *N* (mutations) = 477	
	Variant	%	Variant	%	Variant	%	Variant	%	Variant	%
47	3120+1G>A	0.1	1078delT	0.2	p.R75X	0.2	p.I807M	0.3	p.Q98X	0.2
48	2789+2insA	0.1	**p.V232D**	0.2	1811+1643G>T	0.2	**p.L568X**	0.3	**2957delT**	0.2
49	p.L997F	0.1	**3271delGG**	0.2	**3500-2A>G**	0.2	**p.S1255P**	0.3	**p.T1053I**	0.2
50	p.R352Q	0.1	**p.H609R**	0.2	**1548delG**	0.2	**1898+5G>T**	0.3	p.I807M	0.2
	4		**20**		**14**		**18**		**7**	

Bolded entries and numbers in bottom row represent *CFTR* sequence variants that occur in the top 50 mutations solely in the indicated ethnic group. These variants may be represented at a lower ranking in other ethnic groups. Noncoding sequence variants and frameshifts are listed according to legacy nomenclature for recognition clarity. The corresponding consensus nomenclature is as follows: 394delTT = c.262_263delTT, p.L88Ifs; 405+3A>C = c.273+3A>C; 406-1G>A = c.274-1G>A; 444delA = c.313delA, p.I105Sfs; 457TAT>G = c.325_327delTATinsG, p.Y109Gfs; 621+1G>T = c.489+1G>T; 663delT = c.531delT, p.I177Mfs; 711+1G>T = c.579+1G>T; 712-1G>T = c.580-1G>T; 935delA = c.803delA, p.N268Ifs; 1078delT = c.948delT, p.F316Lfs; 1154insTC = c.1022_1023insTC, p.F342Hfs; 1248+1G>A = c.1116+1G>A; 1288insTA = c.1155_1156het_dupTA, p.N386Ifs; 1525-1G>A = c.1393-1G>A; 1548delG = c.1418delG, p.G473Efs; 1677delTA = c.1545_1546delTA, p.Y515X; 1717-1G>A = c.1585-1G>A; 1811+1643G>T = c.1679+1643G>T; 1811+1.6kbA>G = c.1679+1.6kbA>G; 1812-1G>A = c.1680-1G>A; 1898+1G>A = c.1766+1G>A; 1898+5G>T = c.1766+5G>T; 2055del9>A = c.1923_1931del9insA, p.S641Rfs; 2105-2117del13insAGAAA = c.1973_1985del13insAGAAA, p.R658Kfs; 2143delT = c.2012delT, p.L671X; 2183delAA>G = c.2051_2052delAAinsG, p.K684Sfs; 2184delA = c.2052delA, p.K684Nfs; 2184insA = c.2052_2053insA, p.Q685Tfs; 2307insA = c.2175_2176insA, p.E726Rfs; 2585delT = c.2453delT, p.L818Wfs; 2622+1G>A = c.2490+1G>A; 2789+2insA = c.2657+2_2657+3insA; 2789+5G>A = c.2657+5G>A; 2957delT = c.2825delT, p.I942Tfs; 3120G>A = c.2988G>A; 3120+1G>A = c.2988+1G>A; 3199del6 = c.3067_3072delATAGTG, p.I1023_V1024del; 3271delGG = c.3139_3139+1delGG, p.G1047Qfs; 3272-26A>G = c.3140-26A>G; 3500-2A>G = c.3368-2A>G; 3659delC = c.3528delC, p.K1177Sfs; 3791delC = c.3659delC, p.T1220Kfs; 3849+10kbC>T = c.3717+12191C>T; 3876delA = c.3744delA, p.K1250Rfs; 3905insT = c.3773_3774insT, p.L1258Ffs; 4005+2T>C = c.3873+2T>C.

N, Total number.

neonatal screening algorithms and molecular diagnostic testing. With the implementation of CF newborn screening programs in all 50 US states, the median age at which CF is diagnosed has been reduced to approximately 2 to 4 weeks of age overall,[29] however, nonwhite patients remain at risk of late identification. All newborn screening algorithms begin with the immunoreactive trypsinogen enzyme test on dried blood spots.[30] Increased immunoreactive trypsinogen levels trigger second-tier testing, which varies from state to state but may consist of another immunoreactive trypsinogen test, DNA testing (typically involving a panel of the most common mutations), or some combination of the two. A newborn with a second increased immunoreactive trypsinogen and/or a single *CFTR* panel mutation usually is deemed by the program to be screening test positive. Upon clinical referral, sweat chloride testing typically is performed. Inequitable identification of CF in nonwhite versus white groups persists because mutation analyses often used for newborn screening and/or diagnostic testing are not as effective overall in detecting the CF variants prevalent among nonwhites (Tables 4 and 5). Our results confirm the widely held notion that the American College of Medical Genetics list of 23 mutations that was designed specifically for carrier screening is inadequate for diagnostic testing, even though it is used widely. A delay in the molecular and clinical diagnosis of CF can affect morbidity, mortality, and overall quality of life.

The diversity of the spectrum of sequence changes as well as the frequency of certain alleles in different ethnic populations should be important considerations in the design of any screening program or diagnostic test for the ethnically diverse US population. To enable the consideration of sequence changes that are prevalent in nonwhites for inclusion, we performed a nationwide characterization of *CFTR* variants in nonwhite patients with CF. Although benign changes were excluded, our study did not assign pathogenicity to these sequence changes and should not be interpreted as a recommendation or guideline for the development of mutation panels. Our study relied on entries into the CF Foundation Patient Registry, which is used by CF Centers as a repository for clinical and laboratory patient information. At the end of 2013, the percentage of nonwhite patients with fewer than two identified *CFTR* variants was reduced to approximately half of that in 2008; reducing the disparity between white and nonwhite CF patients therefore is possible. Nevertheless, at present, the mutation detection differences with white patients remain statistically significant (Table 3).

Apart from enhanced analysis of small sequence changes by more appropriate panels or sequencing, diagnostic genotyping for nonwhite patients can be improved further by inclusion of deletion and duplication testing, such as MLPA. Deletions and duplications in the *CFTR* gene appear to be relatively common in nonwhite CF patients and accounted for approximately 17% of unidentified alleles after Sanger sequencing in our cohort of 140 probands. This percentage is consistent with other studies in which rearrangements constituted 11% to 24% of unidentified CF alleles,

Table 7 Novel *CFTR* Sequence Variants in Nonwhite US CF Patients

Amino acid	Nucleotide	Black	Asian	Native American	Total
p.F17Lfs	c.51delC	4			4
p.Q30P	c.89A>C	1			1
p.N48Yfs	*c.142_145delAATC*	1			1
p.A62P	c.184G>C			1	1
p.W79Lfs	*c.234dupC*	1		1	2
p.L123Pfs	c.366_367insC	1			1
p.H146delCAT	c.436_438delCAT	2			2
p.H147delCAC	*c.438_440delTCA*	2			2
p.M150Ifs	c.450delG	1			1
N/A	*c.579+4delT*	4			4
p.G253R	c.757G>C		1		1
p.E292Tfs	c.874_875delGA	1			1
N/A	c.1001+25A>G	1	1		2
N/A	c.1209+1G>C		1		1
N/A	c.1210-1G>T	1			1
p.K522E	c.1564A>G	1			1
p.D565E	c.1695T>A/G		1		1
p.L570Ffs	c.1710delA	1			1
p.V603Sfs	c.1807delG	2			2
p.S624R	c.1870A>C	1			1
p.D648Vfs	c.1943delA		1		1
p.S795Yfs	c.2384delC	1			1
p.Q799X	c.2395C>T	1			1
N/A	c.2619+2T>C	1			1
p.S962X	c.2885C>G/A	3			3
N/A	c.2908+1G>A	1			1
N/A	*c.2982_2988+2delCATCCAGGT*	1			1
p.P1021T	*c.3061C>A*		1		1
p.S1248X	*c.3743C>G*	1			1
p.G1265V	c.3794G>C	2			2
p.E1266X	c.3796G>T	2			2
p.Q1330X	c.3989C>T		1*w		1
p.G1343Afs	c.4028delG	1			1
p.P1372H	c.4115C>A		1		1
Total		39	8	2	49

Novel sequence variants identified in the 140 probands in this study are indicated in italics.

*w = this sequence change also was listed in one CF patient who had indicated white-only ancestry.

depending on the patient population.[31] Six of the nine different deletions and duplications detected among our probands were not reported previously (Table 2 and Figure 1). Such a high proportion of novel rearrangements strongly suggests that the copy number mutations contributing to the nonwhite CF spectrum differ from those found in white CF populations. It previously was shown that deletion/duplication testing is a helpful step in the testing process for Hispanics,[31] and it also seems especially fruitful in black and Asian patients. For example, we identified the deletion of legacy exons 17a and b (19 and 20 by sequential numbering; c.2988+1615_3367+357del3796ins62bp) in four individuals, all of whom declared African American origins (Table 2). Among the entries in the CF Foundation Patient Registry, Asians had the highest percentage, with 2.4% of their variants overall being rearrangements.

The deletion of exons 2 and 3 (c.54-5940_273+10250del21kb) was identified in two of our probands of Native American background but has been described in whites as well. Overall, patients with Native American roots were genetically the most similar to whites, which likely reflects extensive white admixture with various Native American nations or tribes. Interestingly, a high incidence of CF is observed in the geographically isolated Pueblo Native Americans of the southwestern United States, particularly among the Zuni (approximately 1:1580).[32,33] Four mutations—p.G542X, p.R1162X, 3849+10kbC>T (c.3717+12191C>T), and p.D648V—seemingly account for all of the CF alleles identified in the Pueblo populations thus far. The first three are among the most common *CFTR* sequence variants in whites, whereas the last one has been described only in the Pueblo Jemez nation.

The deletion of legacy exons 17a to 18 (19 to 21 by sequential numbering, c.2988+1173_c.3468+2111del8898) was identified in a patient of Middle Eastern descent, and previously described in a patient of similar origin.[27] In our analysis of 51 individuals in whom sequencing did not show two variants, 2 of 12 identified deletions were in Middle Eastern patients, which may indicate that such rearrangements are relatively common in that population. Unfortunately, since 2008, the Middle Eastern designation in the CF Foundation Registry was absorbed in the white category, which prevented further analysis. It is clear, however, that our study contributes considerably to knowledge of the *CFTR* deletion and duplication mutation spectrum in nonwhite patients overall. MLPA should be applied routinely in diagnostic *CFTR* testing, and is an especially valuable part of the testing algorithm in nonwhite and mixed-ethnicity individuals affected with CF.

Upon review of the rearrangements recorded in the CF Foundation Patient Registry, it became clear that we would not be able to determine whether rearrangements with the same name (eg, CFTRdele22-24) were, in fact, correctly named and breakpoints fully characterized. This limited further analysis, except for the determination that deletions and duplications constituted 0.2% of white alleles, 0.1% of Hispanic alleles, 0.6% of black alleles, 2.4% of Asian alleles, and 0.4% of Native American alleles. These numbers appear relatively low,[13] but the deletions identified by MLPA in our study were confirmed and fully characterized in a research setting, however, because they were not reported to the CF Center would not be entered into the Registry. In addition, it is likely that the majority of patients did not have testing for deletions and duplications as part of their molecular work-up.

In the 140 probands whose samples were laboratory analyzed as part of our study, two or more variants were identified in 72.1% (101 of 140), and in 27.9% (39 of 140) fewer than two variants were found. Of these, 14 individuals (10%) had a single variant identified and 25 (17.9%) had none. One can only speculate regarding the reasons for these missing

ARTICLE IN PRESS

variants. They could be present in areas of the gene that are not usually included in sequence analyses, such as introns. Alternatively, they could reflect the possibility that some patients followed up at CF centers have a constellation of symptoms in the CF spectrum that mimics aspects of the condition but is not rooted in *CFTR* etiology.

A cross-sectional study such as ours inevitably has limitations. For example, the United States has much ethnic diversity and admixture, and we depended on self-identification of ethnic groups. Individuals who were aware of only white or only Hispanic ancestry may have unknown genetic admixtures of other ethnic groups. For nonwhite patients, we assigned them to categories based on any black, Asian, or Native American ancestry. This assignment may be too limited for a group with few patients, such as the Asian category. More than any other nonwhite, non-Hispanic category, Asians had sequence changes in their most frequent 50 variants (Table 6) that were not seen in other groups. This likely reflects the relatively high number of individual or private sequence changes on the list (each seen in one patient only). The actual number of 18 unique top 50 changes, however, is arbitrary because, of the 60 variants seen in a single patient in the Asian cohort, only 23 are listed as part of the 50 most frequent variants. In addition, the assignment of Asian ancestry is very broad and includes divergent geographies and backgrounds. Finally, patients with mixed ethnic backgrounds (eg, Asian-black), were assigned to both categories. At least one other race was listed for 15% of blacks, 32% of Asians, and 51% of Native Americans. For such patients, we ultimately attributed one mutation to each of these two groups, realizing that this is an imperfect assignment. Nevertheless, through the CF Foundation Patient Registry, we obtained the currently most complete and reliable *CFTR* sequence change information available, and recognize that future refinements will be necessary.

In conclusion, this is the first large-scale and in-depth look into *CFTR* sequence variants in nonwhite US patients and an important step toward future studies that correlate *CFTR* sequence changes (and combinations thereof) observed in nonwhite individuals with clinical severity of symptoms. Knowledge of sequence variants in each population and improved genotype-phenotype correlations can affect results reporting, counseling, prognosis predictions, and therapeutic decisions. This information can be used to optimize newborn screening programs in the United States based on the ethnic composition of state populations, resulting in earlier diagnosis and intervention, timely clinical treatment, and enhanced prognosis. For both screening and diagnostic testing it could propel equity in mutation detection for white and nonwhite CF patients.

Acknowledgments

We thank the CF Centers with nonwhite patients for their cooperation, and the participating CF patients, without whom no progress would be possible.

Supplemental Data

Supplemental material for this article can be found at *http://dx.doi.org/10.1016/j.jmoldx.2015.07.005*.

References

1. Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, Durie PR, Legrys VA, Massie J, Parad RB, Rock MJ, Campbell PW 3rd; Cystic Fibrosis Foundation: Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. J Pediatr 2008, 153:S4–S14
2. Abeliovich D, Lavon IP, Lerer I, Cohen T, Springer C, Avital A, Cutting GR: Screening for five mutations detects 97% of cystic fibrosis (CF) chromosomes and predicts a carrier frequency of 1:29 in the Jewish Ashkenazi population. Am J Hum Genet 1992, 51:951–956
3. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ; Subcommittee on Cystic Fibrosis Screening, Accreditation of Genetic Services Committee, ACMG. American College of Medical Genetics: Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. Genet Med 2001, 3:149–154
4. Palomaki GE, FitzSimmons SC, Haddow JE: Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. Genet Med 2004, 6:405–414
5. Hamosh A, FitzSimmons SC, Macek M Jr, Knowles MR, Rosenstein BJ, Cutting GR: Comparison of the clinical manifestations of cystic fibrosis in black and white patients. J Pediatr 1998, 132:255–259
6. Welsh MJ, Ramsey BW, Accurso F, Cutting GR: Cystic Fibrosis. Edited by Valle D, Beaudet AL, Vogelstein B, Kinzler KW, Antonarakis SE, Ballabio A, Gibson K, Mitchell G. OMMBID - the online metabolic and molecular bases of inherited diseases. New York, NY, McGraw-Hill, 2013
7. Zielenski J, Rozmahel R, Bozon D, Kerem B, Grzelczak Z, Riordan JR, Rommens J, Tsui LC: Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Genomics 1991, 10:214–228
8. Anderson MP, Rich DP, Gregory RJ, Smith AE, Welsh MJ: Generation of cAMP-activated chloride currents by expression of CFTR. Science 1991, 251:679–682
9. Anderson MP, Gregory RJ, Thompson S, Souza DW, Paul S, Mulligan RC, Smith AE, Welsh MJ: Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. Science 1991, 253:202–205
10. Hull J: Cystic fibrosis transmembrane conductance regulator dysfunction and its treatment. J R Soc Med 2012, 105 Suppl 2:S2–S8
11. Dhooghe B, Noël S, Huaux F, Leal T: Lung inflammation in cystic fibrosis: pathogenesis and novel therapies. Clin Biochem 2014, 47:539–546
12. Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, Ramalho AS, Amaral MD, Dorfman R, Zielenski J, Masica DL, Karchin R, Millen L, Thomas PJ, Patrinos GP, Corey M, Lewis MH, Rommens JM, Castellani C, Penland CM, Cutting GR: Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat Genet 2013, 45:1160–1167
13. Dequeker E, Stuhrmann M, Morris MA, Casals T, Castellani C, Claustres M, Cuppens H, des Georges M, Ferec C, Macek M, Pignatti PF, Scheffer H, Schwartz M, Witt M, Schwarz M, Girodon E: Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders–updated European recommendations. Eur J Hum Genet 2009, 17:51–65
14. Bobadilla JL, Macek M Jr, 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
15. Alper OM, Wong LJ, Young S, Pearl M, Graham S, Sherwin J, Nussbaum E, Nielson D, Platzker A, Davies Z, Lieberthal A, Chin T,

Shay G, Hardy K, Kharrazi M: Identification of novel and rare mutations in California Hispanic and African American cystic fibrosis patients. Hum Mutat 2004, 24:353
16. Macek M Jr, Mackova A, Hamosh A, Hilman BC, Selden RF, Lucotte G, Friedman KJ, Knowles MR, Rosenstein BJ, Cutting GR: Identification of common cystic fibrosis mutations in African-Americans with cystic fibrosis increases the detection rate to 75%. Am J Hum Genet 1997, 60:1122−1127
17. Watson MS, Cutting GR, Desnick RJ, Driscoll DA, Klinger K, Mennuti M, Palomaki GE, Popovich BW, Pratt VM, Rohlfs EM, Strom CM, Richards CS, Witt DR, Grody WW: Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. Genet Med 2004, 6:387−391
18. Rosenstein BJ, Cutting GR: The diagnosis of cystic fibrosis: a consensus statement. Cystic Fibrosis Foundation Consensus Panel. J Pediatr 1998, 132:589−595
19. Accurso FJ, Sontag MK, Wagener JS: Complications associated with symptomatic diagnosis in infants with cystic fibrosis. J Pediatr 2005, 147:S37−S41
20. Campbell PW 3rd, White TB: Newborn screening for cystic fibrosis: an opportunity to improve care and outcomes. J Pediatr 2005, 147: S2−S5
21. Spencer DA, Venkataraman M, Weller PH: Delayed diagnosis of cystic fibrosis in children from ethnic minorities. Lancet 1993, 342:238
22. Spencer DA, Venkataraman M, Higgins S, Stevenson K, Weller PH: Cystic fibrosis in children from ethnic minorities in the West Midlands. Respir Med 1994, 88:671−675
23. Chillon M, Dork T, Casals T, Gimenez J, Fonknechten N, Will K, Ramos D, Nunes V, Estivill X: A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811+1.6kbA–>G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. Am J Hum Genet 1995, 56: 623−629
24. Highsmith WE, Burch LH, Zhou Z, Olsen JC, Boat TE, Spock A, Gorvoy JD, Quittel L, Friedman KJ, Silverman LM, Boucher RC, Knowles MR: 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
25. Schrijver I, Ramalingam S, Sankaran R, Swanson S, Dunlop CL, Keiles S, Moss RB, Oehlert J, Gardner P, Wassman ER, Kammesheidt A: Diagnostic testing by CFTR gene mutation analysis in a large group of Hispanics: novel mutations and assessment of a population-specific mutation spectrum. J Mol Diagn 2005, 7:289−299
26. Dörk T, Macek M Jr, Mekus F, Tümmler B, Tzountzouris J, Casals T, Krebsová A, Koudová M, Sakmaryová I, Macek M Sr, Vávrová V, Zemková D, Ginter E, Petrova NV, Ivaschenko T, Baranov V, Witt M, Pogorzelski A, Bal J, Zékanowsky C, Wagner K, Stuhrmann M, Bauer I, Seydewitz HH, Neumann T, Jakubiczka S: Characterization of a novel 21-kb deletion, CFTRdele2,3(21 kb), in the CFTR gene: a cystic fibrosis mutation of Slavic origin common in Central and East Europe. Hum Genet 2000, 106:259−268
27. Lerer I, Laufer-Cahana A, Rivlin JR, Augarten A, Abeliovich D: A large deletion mutation in the CFTR gene (3120+1Kbdel8.6Kb): a founder mutation in the Palestinian Arabs. Mutation in brief no. 231. Hum Mutat 1999, 13:337
28. Tang S, Moonnumakal SP, Stevens B, Douglas G, Mason S, Schmitt ES, Eng CM, Katz M, Fang P: Characterization of a recurrent 3.8kb deletion involving exons 17a and 17b within the CFTR gene. J Cyst Fibros 2013, 12:290−294
29. Sanders DB, Lai HJ, Rock MJ, Farrell PM: Comparing age of cystic fibrosis diagnosis and treatment initiation after newborn screening with two common strategies. J Cyst Fibros 2012, 11:150−153
30. Ross LF: Newborn screening for cystic fibrosis: a lesson in public health disparities. J Pediatr 2008, 153:308−313
31. Schrijver I, Rappahahn K, Pique L, Kharrazi M, Wong LJ: Multiplex ligation-dependent probe amplification identification of whole exon and single nucleotide deletions in the CFTR gene of Hispanic individuals with cystic fibrosis. J Mol Diagn 2008, 10:368−375
32. Mercier B, Raguénès O, Estivill X, Morral N, Kaplan GC, McClure M, Grebe TA, Kessler D, Pignatti PF, Marigo C, Bombieri C, Audrézet MP, Verlingue C, Férec C: Complete detection of mutations in cystic fibrosis patients of Native American origin. Hum Genet 1994, 94:629−632
33. Genetic testing for cystic fibrosis. National Institutes of Health Consensus Development Conference Statement on genetic testing for cystic fibrosis. Arch Intern Med 1999, 159:1529−1539