

Improved detection of cystic fibrosis mutations in the heterogeneous U.S. population using an expanded, pan-ethnic mutation panel

Ruth A. Heim, PhD, Elaine A. Sugarman, MS, and Bernice A. Allitto, PhD

Purpose: To determine the comparative frequency of 93 *CFTR* mutations in U.S. individuals with a clinical diagnosis of cystic fibrosis (CF). **Methods:** A total of 5,840 CF chromosomes from Caucasians, Ashkenazi Jews, Hispanics, African Americans, Native Americans, Asians, and individuals of mixed race were analyzed using a pooled ASO hybridization strategy. **Results:** Sixty-four mutations provided a sensitivity of 70% to 95% in all ethnic groups except Asians, and at least 81% when the U.S. population was considered as a whole. **Conclusions:** For population-based carrier screening for CF in the heterogeneous U.S. population, which is characterized by increasing admixture, a pan-ethnic mutation panel of 50 to 70 *CFTR* mutations may provide a practical test that maximizes sensitivity. **Genetics in Medicine, 2001;3(3):168–176.**

Key Words: cystic fibrosis, *CFTR* mutations, carrier screening, molecular diagnosis, population genetics

Cystic fibrosis (CF) is one of the most common autosomal recessive disorders in Caucasian populations. A recent study of cases of CF in the United States estimated the incidence to be 1 in 3,200 white, 1 in 9,200 Hispanic, 1 in 10,900 Native American, 1 in 15,000 black, and 1 in 31,000 Asian American live births.¹ Since the identification of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene² more than 900 presumed mutations and 190 DNA sequence variants have been identified throughout the gene.³ The geographical distribution of mutations and their relative frequencies vary considerably, however. In a worldwide study of more than 43,000 CF chromosomes,⁴ only one mutation ($\Delta F508$) was present in the majority of CF chromosomes (66%), while only four others had relative frequencies of between 1.0% and 2.5%. Similarly, when the European population was considered as a whole, only 22 mutations had relative frequencies of $> 0.1\%$.⁵ The large number of different mutations found within populations reflects their genetic heterogeneity. To provide a cost-effective CF test and optimize the CF detection rate, the selection of a subset of mutations to be analyzed must be carefully considered and appropriate to the target population.

The selection of an appropriate mutation subset is confounded in the United States by the changing nature of the U.S. population. According to the U.S. Census Bureau (<http://www.census.gov>), in 1997, 11.1% of U.S. residents were Hispanic, 12.8% were black and 71.9% were white. One in 10 members of the population were foreign born. The Census Bureau projects that by 2010 the Hispanic population may

become the second largest racial/ethnic group in the United States. The Census data suggest that, despite the large European contribution to the U.S. population, the overall distribution of CF mutations is likely to be different from that in Europe and is likely to change over time.

To date, estimates of relative mutation frequencies and overall CF detection rates in specific ethnic groups in the United States have been determined in small, selected groups using a limited number of mutations. Since all individuals are seldom tested for the same mutations, overall sensitivity data are often extrapolated or assumed to be additive across studies. In small, recently published studies of U.S. ethnic groups, detection rates were 75% in African Americans,⁶ 58% in Hispanics,⁷ 33% in Asians,⁸ and 4% in Native Americans.⁹ In a large study of U.S. Caucasians in 1994, the detection rate was 78.6% (derived from ref. 4), but there have been no subsequent mutational surveys of the general U.S. testing population.

The limitations in studies of mutation frequencies in the United States have an impact on the debate regarding whether and how to implement widespread population-based carrier screening.¹⁰ In 1997, a National Institutes of Health Consensus Statement¹¹ recommended that CF carrier screening be offered to individuals with a family history of CF and their partners, as well as to all preconceptional and pregnant couples, particularly those in high-risk populations. Details regarding how to implement such a program were deferred. Questions that remain include how many and which mutations should comprise a core testing panel, and which populations should be targeted—the entire U.S. population or high-risk ethnic groups only.

To address these questions, we analyzed 5,840 CF chromosomes from individuals with a clinical diagnosis of CF referred to our diagnostic laboratory from throughout the United States. Mutations identified in this affected group may not reflect the actual proportion of these mutations in the general

From the Molecular Diagnostic Laboratory, Genzyme Genetics, Framingham, Massachusetts.
Ruth A. Heim, PhD, DNA Diagnostic Laboratory, Genzyme Genetics, 5 Mountain Road, Framingham, MA 01701-9322.

Received: December 27, 2000.

Accepted: February 21, 2001.

population. They do reflect mutations associated with a CF phenotype, however. Mutation frequency data derived from an affected population are used here to make recommendations for carrier screening, based on the expectation that carriers of these mutations have a rational basis for presuming that a child who inherits two such mutations will have CF.

We present here the relative mutation frequencies and distribution of 93 *CFTR* mutations. We show that an expanded, pan-ethnic panel of 64 mutations is more sensitive than any previously published mutation sets in all U.S. ethnic groups studied except Ashkenazi Jews. The results further our understanding of the distribution of CF alleles in the diverse U.S. population and facilitate the development of more sensitive tests for CF mutation analysis, for diagnostic testing or for population-based carrier screening in the United States.

SUBJECTS AND METHODS

Test population

The test population included 2,920 individuals with a clinical diagnosis of CF, referred from all 50 states, with 28.3% from the Northeast, 24.8% from the Southeast, 16.6% from the West, and 30.3% from the Central States. Referring physicians provided information regarding clinical status and ethnicity. Ethnic classification was based on self-identification, country of origin, or race (for example, individuals referred to as "white" were classified as "Other Caucasian"). The ethnic origins of the individuals tested are summarized in Table 1. Approximately one-third of the individuals in this study were of mixed, unknown, or not readily classifiable ancestry. They are included in two categories. First, the Mixed/Other Caucasian category, in which the majority of individuals were of mixed Northern and Southern European ancestry. Second, the Other/Mixed Race/Unknown category, which included all non-

Europeans with more than one ethnic background, all individuals whose ethnic background was not provided, and all individuals who could not be classified as Hispanic, African American, Asian, or Native American. Individuals in the Hispanic category were of Latin American origin, either self-identified as Hispanic or of Caribbean, Central, or South American heritage. Affected siblings are excluded from all analyses of mutation frequencies.

Mutations

Samples were analyzed using one of two mutation panels comprising 70 or 86 mutations, as noted in Table 2 and in Results. The mutations in the 70-mutation panel were selected from the literature on the basis of known contribution to CF, or predicted deleterious effect on the *CFTR* gene product. The 70-mutation panel was replaced with an 86-mutation panel based on data from more than 30,000 individuals referred for diagnostic or carrier testing to our laboratory. The 86-mutation panel included 63 mutations in common with the 70-mutation panel, included an additional 23 mutations, and excluded seven mutations not identified in the 70-mutation panel (see Results). In total, 5,840 alleles were tested with the 63 mutations using either the 70- or the 86-mutation panels. The additional 23 mutations were tested in a subset of 1,512 alleles using the 86-mutation panel only. The seven mutations unique to the 70-mutation panel were tested in a nonoverlapping subset of 4,328 alleles. A total of nine mutations had not been detected in the 70-mutation panel, but two of these were retained in the 86-mutation panel. The mutation 3662delA¹² was retained, because it was originally identified in the African American population, and 3849 + 4A>G¹³ was retained, because it has generally been considered relevant to the Caucasian population. In both cases, we wished to collect more data. The 23 additional mutations included in the 86-mutation panel were considered likely to be relevant to the test populations. They were selected according to the criteria noted above, with the additional requirement for each mutation of a published frequency of > 1% in at least one of the ethnic groups constituting the test population described in Table 1. In total, 93 mutations were analyzed by using both mutation panels.

The 70-mutation assay distinguished between the $\Delta F508$ mutation and the F508C sequence variant. The 86-mutation assay distinguished $\Delta F508$ from the F508C, I506V, I506M, and I507V sequence variants.

Mutation analysis

Specimens received included anticoagulated peripheral blood, buccal brushes, and other human tissues. Genomic DNA was isolated by standard methods (high salt or organic extraction), or by a commercially available DNA extraction kit (Qiagen). Mutation analysis was performed using a pooled allele-specific oligonucleotide (ASO) strategy as described previously.¹⁴ This strategy allowed simultaneous analysis of hundreds of samples. Nineteen regions of the *CFTR* gene were amplified in two multiplex reactions. Am-

Table 1

Ethnic origins of individuals tested for *CFTR* mutations

Ethnic/race origin	n ^a (%)
Northern European Caucasian (NEC)	1,498 (51.3)
Southern European Caucasian (SEC)	97 (3.3)
Mixed Caucasian/Other Caucasian	722 (24.7)
Other/Mixed Race/Unknown	325 (11.1)
Hispanic	124 (4.3)
African American	101 (3.5)
Ashkenazi Jewish	24 (0.8)
Native American	21 (0.7)
Asian	8 (0.3)
Total	2,920 (100.0)

^aNumber of individuals with a clinical diagnosis of CF prior to mutation analysis; in brackets the percentage of all individuals tested. Percentages are rounded.

Table 2
Frequency and ethnic distribution of 64 *CFTR* mutations in 5,840 U.S. CF chromosomes^c

[Column reference] Mutation	[1] NEC ^d	[2] SEC ^e	[3] Mixed/ Other Caucasian	[1 + 2 + 3] Combined Caucasian ^f	[4] Other/ Mixed Race/ Unknown	[5] Ashkenazi Jewish	[6] Hispanic	[7] African American	[8] Native American	[9] Asian	
	n ^a n ^b	2,996 782	194 52	1,444 382	4,634 1,216	650 146	48 10	248 74	202 58	42 4	16 4
E60X ^a		6 (0.2)		10 (0.7)	16 (0.3)		1 (0.4)				
G85E ^a		10 (0.3)	2 (1.0)	2 (0.1)	14 (0.3)	4 (0.6)	2 (0.8)				
394delTT ^b		1 (0.1)		1 (0.3)	2 (0.2)	1 (0.7)					
405+1G>A ^a											1 (6.3)
405+3A>C ^b		1 (0.1)			1 (0.1)						1 (6.3)
444delA ^a								3 (1.5)			
R117C ^b				2 (0.5)	2 (0.2)						
R117H ^a		22 (0.7)	1 (0.5)	8 (0.6)	31 (0.7)	1 (0.2)		2 (1.0)	1 (2.4)		
I148T ^a		3 (0.1)			3 (0.1)		1 (0.4)				
621+1G>T ^a		27 (0.9)	4 (2.1)	14 (1.0)	45 (1.0)	6 (0.9)		1 (0.5)			
G178R ^a		4 (0.1)		3 (0.2)	7 (0.2)						1 (6.3)
711+1G>T ^a		3 (0.1)	1 (0.5)	2 (0.1)	6 (0.1)		1 (0.4)				
L206W ^b		1 (0.1)			1 (0.1)						
ΔF311 ^b						1 (0.7)	1 (1.4)				
I078delT ^a							1 (0.4)				
G330X ^a						1 (0.2)		3 (1.5)			
R334W ^a		7 (0.2)	1 (0.5)	2 (0.1)	10 (0.2)	4 (0.6)	5 (2.0)	1 (0.5)			
R347H ^a		5 (0.2)		2 (0.1)	7 (0.2)						
R347P ^a		12 (0.4)	1 (0.5)	6 (0.4)	19 (0.4)	1 (0.2)					
R352Q ^a		1		5 (0.3)	6 (0.1)		1 (0.4)				
S364P ^a						1 (0.2)					
A455E ^a		7 (0.2)		5 (0.3)	12 (0.3)	2 (0.3)	1 (2.1)				
G480C ^b								1 (1.7)			
Q493X ^a		7 (0.2)		5 (0.3)	12 (0.3)	1 (0.2)		1 (0.5)			
ΔI507 ^a		17 (0.6)		11 (0.8)	28 (0.6)	1 (0.2)	6 (2.4)				
ΔF508 ^a		1983 (66.2)	119 (61.3)	895 (62.0)	2997 (64.7)	359 (55.2)	14 (29.2)	80 (32.3)	58 (28.7)	28 (66.7)	3 (18.8)
1677delTA ^a				2 (0.1)	2		2 (0.8)				
V520P ^a		4 (0.1)	1 (0.5)	2 (0.1)	7 (0.2)						
1717-1G>A ^a		27 (0.9)	3 (1.5)	9 (0.6)	39 (0.8)	6 (0.9)					
G542X ^a		52 (1.7)	2 (1.0)	23 (1.6)	77 (1.7)	14 (2.2)	4 (8.3)	7 (2.8)			
S549N ^a		1			1	1 (0.2)	3 (1.2)				
S549R ^a		2 (0.1)		1 (0.1)	3 (0.1)	1 (0.2)					
G551D ^a		76 (2.5)	4 (2.1)	39 (2.7)	119 (2.6)	10 (1.5)	1 (0.4)	2 (1.0)	1 (2.4)	1 (6.3)	
R553X ^a		32 (1.1)	2 (1.0)	20 (1.4)	54 (1.2)	10 (1.5)	1 (2.1)	3 (1.2)	1 (0.5)		
A559T ^a								2 (1.0)			
R560T ^a		14 (0.5)		7 (0.5)	21 (0.5)						
1812-1G>A ^a			1 (0.5)		1		1 (0.4)	2 (1.0)			
1898+1G>A ^a		16 (0.5)	1 (0.5)	8 (0.6)	24 (0.5)	1 (0.2)	1 (0.4)	2 (1.0)			

—Continued

Table 2
Continued

[Column reference] Mutation	[1] NEC ^d	[2] SEC ^e	[3] Mixed/ Other Caucasian	[1 + 2 + 3] Combined Caucasian ^f	[4] Other/ Mixed Race/ Unknown	[5] Ashkenazi Jewish	[6] Hispanic	[7] African American	[8] Native American	[9] Asian	
	n ^a n ^b	2,996 782	194 52	1,444 382	4,634 1,216	650 146	48 74	248 74	202 58	42 4	16 4
2143delT ^b		1 (0.1)		1 (0.1)							
2183delAA>G ^a		10 (0.3)	2 (1.0)	7 (0.5)	19 (0.4)	1 (0.2)	1 (0.4)				
2184delA ^a		11 (0.4)		5(0.3)	16 (0.3)	1 (0.2)					
K710X ^a		2 (0.1)			2						
2307insA ^a								1 (0.5)			
2789+5G>A ^a		23 (0.8)	1 (0.5)	13 (0.9)	37 (0.8)	6 (0.9)	1 (0.4)				
Q890X ^a				1 (0.1)	1						
2869insG ^a						1 (0.2)	1 (0.4)				
3120G>A ^a		7 (0.2)			7 (0.2)	1 (0.2)					
3120+1G>A ^a		4 (0.1)			4 (0.1)	7 (1.1)	5 (2.0)	27 (13.8)			
R1066C ^b		2 (0.3)	1 (1.9)		3 (0.2)	2 (1.4)		1 (1.7)			
W1089X ^a							3 (1.2)				
Y1092X ^a		10 (0.3)		3 (0.2)	13 (0.3)	1 (0.2)	1 (0.4)				
M1101K ^b		5 (0.6)	1 (1.9)		6 (0.5)						
D1152H ^b							1 (10.0)				
R1158X ^a		4 (0.1)	1 (0.5)	2 (0.1)	7 (0.2)	1 (0.2)		4 (2.0)			
R1162X ^a		11 (0.4)	1 (0.5)	6 (0.4)	18 (0.4)	1 (0.2)	3 (1.2)	1 (0.5)			
3659delC ^a		15 (0.5)		7 (0.5)	22 (0.5)	3 (0.5)					
3791delC ^a								4 (2.0)			
S1195X ^a		2 (0.1)			2						
3849+10kbC>T ^a		14 (0.5)	2 (1.0)	5 (0.3)	21 (0.5)		2 (4.2)	1 (0.4)		2 (4.8)	
S1255X ^a								2 (1.0)			
3905insT ^a		14 (0.5)		2 (0.1)	16 (0.3)	1 (0.2)		1 (0.5)	1 (2.4)		
D1270N ^a		1			1		3 (1.2)				
W1282X ^a		21 (0.7)	4 (2.1)	30 (2.1)	55 (1.2)	2 (0.3)	16 (33.3)	3 (1.2)			
N1303K ^a		39 (1.3)	3 (1.5)	27 (1.9)	69 (1.5)	2 (0.3)	3 (6.3)	3 (1.2)		1 (2.4)	
Undetected		459	35	252	746	194	6	106	82	8	10

^aMutation included in both the 70- and 86-mutation panels (n^a = number of chromosomes tested in specific ethnic group).

^bMutation included in the 86-mutation panel only (n^b = number of chromosomes tested in specific ethnic group).

^cNumbers are the number of mutation-positive chromosomes; brackets contain relative frequencies $\geq 0.1\%$ for each mutation as a percentage of the total number of chromosomes tested; where no number is given, the mutation was not identified in the specified number of chromosomes; mutations are listed by nucleotide/ amino acid position in the CFTR gene/protein, from 5' to 3'/NH₂ to COOH. All percentages are rounded.

^dNorthern European Caucasian.

^eSouthern European Caucasian.

^fSum of NEC, SEC, and Mixed/Other Caucasian.

plified products were immobilized on dot blots and hybridized to different pools of radiolabeled ASO probes. Pool-positive samples were identified by autoradiography and comparison of hybridization intensities with cloned positive controls. To determine specific mutation identity, pool-positive samples were further analyzed by hybridization with individual ASOs.

RESULTS

Of the 93 different mutations analyzed, 64 were detected at least once and 29 were not detected at all in the U.S. populations studied. Table 2 shows the frequency and distribution of the 64 different mutations detected. Fifty-four of the 64 were included in both the 86-mutation panel and the 70-mutation

Table 3
Comparison of *CFTR* mutation detection rates by ethnic group^a

Ethnic group	Mutations in 70- and 86-mutation panels			Mutations in 86-mutation panel only			All mutations
	No. of CF chromosomes detected	No. of CF chromosomes analyzed	Detection rate (1)	No. of CF chromosomes detected	No. of CF chromosomes analyzed	Detection rate (2)	Overall detection rate [sum of (1) and (2)]
Caucasian							
Northern European (NEC) [1] ^b	2,526	2,996	84.3%	11	782	1.4%	85.7%
Southern European (SEC) [2]	157	194	80.9%	2	52	3.8%	84.7%
Mixed/Other [3]	1,189	1,444	82.3%	3	382	0.8%	83.1%
Combined ([1], [2], and [3])	3,872	4,634	83.6%	16	1,216	1.3%	84.9%
Other/Mixed Race/Unknown [4]	452	650	69.5%	4	146	2.7%	72.2%
Ashkenazi Jewish [5]	41	48	85.4%	1	10	10.0%	95.4%
Hispanic [6]	141	248	56.9%	1	74	1.4%	58.3%
African American [7]	118	202	58.4%	2	58	3.5%	61.9%
Native American [8]	34	42	81.0%	0	4	0	81.0%
Asian [9]	6	16	37.5%	0	4	0	37.5%
All	4,664	5,840	79.8%	24	1,512	1.6%	81.4%

^aDetection rates are calculated as the proportion of mutation-positive chromosomes among the total number of chromosomes analyzed expressed as a percentage. Percentages are rounded.

^bNumbers in brackets refer to the corresponding column in Table 2.

panel. These mutations were, therefore, analyzed in 5,840 chromosomes and are designated in Table 2 with the superscript "a." Ten of the 64 mutations were included in the 86-mutation panel only. They were, therefore, analyzed in 1,512 chromosomes and are designated in Table 2 with the superscript "b." Since the 64 mutations were analyzed in two sample sets of different size, detection rates were calculated separately and then added together for overall detection rates (Table 3). The numbers of different mutations detected among the various ethnic groups analyzed varied considerably, and are summarized in Table 4.

Mutations not detected

Twenty-nine mutations of the 93 analyzed were not detected at all. Seven mutations included in the 70-mutation panel were not detected in 4,328 chromosomes analyzed. These mutations, Y122X, 556delA, 2909delT, 3358delAC, 3750delAG, W1310X, and W1316X, were subsequently excluded from the 86-mutation panel. Nine mutations included in both the 70- and 86-mutation panels were not detected in 5,840 chromosomes. These mutations, 574delA, C524X, Y563D, P574H, 2043delG, 3662delA, 3821delT, Q1238X, and 3849 + 4A>G, as well as the previous seven mutations, make no detectable contribution to mutation detection in the general U.S. CF population. Thirteen mutations included in the 86-mutation panel were not detected in 1,512 chromosomes. These were G91R, 711 + 5G>A, T338I, 712-1G>T, Q359K/T360K, 1161delC, 1609delCA, S549I, Q552X, 1949del84, 1989 + 5G>T, S1251N, and R1283M. Since they were analyzed in a relatively small sample set, it will be necessary to collect more data to deter-

Table 4
Total numbers of different *CFTR* mutations detected

Ethnic group	Total no. of different mutations detected ^a
Caucasian	
Northern European (NEC) [1] ^b	46
Southern European (SEC) [2]	23
Mixed/Other European [3]	37
Combined Caucasian ([1], [2], and [3])	50
Other/Mixed Race/Unknown [4]	37
Ashkenazi Jewish [5]	8
Hispanic [6]	28
African American [7]	21
Native American [8]	6
Asian [9]	4
Combined Mixed Race/Other/Unknown ([3] and [4])	56
All	64

^aTotal number of different mutations analyzed = 93.

^bNumbers in brackets refer to the corresponding column in Table 2.

mine what contribution these thirteen mutations make to mutation detection in the United States. It is possible that some of these mutations are associated with milder CF phenotypes and would be more frequent among populations selected for atypical CF phenotypes.

Overall detection rate

When all ethnic groups in this study are considered, the 64 different mutations detected were distributed among 4,688 CF chromosomes from the United States, accounting for an overall detection rate of 81.4% (summarized in Tables 2, 3 and 4).

DISCUSSION

Development of an appropriate mutation test panel for population-based carrier screening requires consideration of the extreme heterogeneity of the U.S. population, admixture among ethnic groups in the United States, and the differing prevalence of individual mutations among various ethnic, demographic, and racial groups. By analyzing mutations not previously studied in diverse ethnic groups or previously identified only in small sample sets, and by analyzing all groups of individuals with standard panels of mutations, we obtained new and sometimes surprising information regarding mutation frequencies and distribution.

Caucasian mutation detection

The observed detection rate of 85% in 2,317 U.S. Caucasians is higher than the detection rate of 78.6% reported by the Cystic Fibrosis Genetic Analysis Consortium (CFGAC)⁴ in 4,124 U.S. Caucasians. Fifty different mutations were required to achieve the observed detection rate, compared with 23 different mutations in the CFGAC study.⁴ Therefore, the increase in sensitivity in the present study is likely to be due to the expanded mutation panel.

The observed frequency of $\Delta F508$ (64.7%) is slightly lower than that reported by the CFGAC study⁴ (66%). This difference can be explained by ascertainment bias. We have evidence that our referral population is depleted of $\Delta F508$ homozygotes because of prior testing for this common mutation (data not shown). Our referral pattern predicts that both the frequency of $\Delta F508$ in the Caucasian CF population and the overall detection rate are underestimates. Conversely, it predicts an over-ascertainment of non- $\Delta F508$ mutations, although the relative mutation frequencies provided in Table 2 are unaffected by ascertainment.

Our data show that mutations relatively frequent in Europe are not necessarily frequent in U.S. Caucasians. Examples include 394delTT (identified in only two U.S. Caucasian alleles, although it has a frequency of 1.1–28.8% in Northern Europe⁵), T338I (not identified in the United States, although it has a frequency of 9.9% in Sardinia and nearly 0.1% in Italy⁵), and 1609delCA (not identified in 434 Southern or mixed European Caucasian alleles and 74 Hispanic alleles, although it has a frequency of 4.1% in the east of Spain⁵). Conversely, mutations frequent in genetic isolates or specific ethnic groups are surprisingly frequent in the general Caucasian population. Examples include M1101K (identified in 6/1,216 Caucasian alleles including one of Southern European origin, but originally described as specific to the Hutterite population¹⁵), and

3120 + 1G>A (identified in 4/4,634 Caucasian alleles, but reported to be specific to African Americans⁶).

Ashkenazi Jewish mutation detection

The overall detection rate in Ashkenazi Jews from the U.S. (95.4%) is slightly lower than that determined by Abeliovich et al.¹⁶ (97%). More strikingly, the five mutations reported to account for 97% of Ashkenazi Jewish chromosomes ($\Delta F508$, G542X, W1282X, N1303K, and 3849 + 10 kbC>T)¹⁶ accounted for only 39/48 chromosomes in this study (81.3%). This difference is unlikely to be due to ascertainment, since the observed frequency of $\Delta F508$ (29%) is equivalent to the frequency of 30% reported by Abeliovich et al.¹⁶ In this study, the detection of an additional three mutations was required to bring the overall detection rate to 95.4% (A455E, R553X, and D1152H). Our sample size is small, but the results suggest that admixture may be a factor to consider when developing mutation panels for CF analysis in Ashkenazi Jews. An expanded panel beyond the widely accepted panel of five mutations may be of value to the U.S. Ashkenazi Jewish population.

Hispanic mutation detection

The observed detection rate of 58.3% in 246 Hispanics from throughout the U.S. is consistent with the detection rate of 58% in 129 Hispanics geographically localized to the Southwest.⁷ However, we believe that the observed frequency of $\Delta F508$ (32.3%) is an underestimate because of under-ascertainment of $\Delta F508$ homozygotes, as described in the Caucasian mutation detection section. When $\Delta F508$ is excluded, the detection rate in the present study is 26.0% compared with 12.4%.⁷ This increase in sensitivity was achieved by detecting 28 different mutations, compared with the eight detected by Grebe et al.⁷ In the absence of referral bias, based on a $\Delta F508$ frequency of 46%,⁷ we predict that a panel of 28 mutations would provide a Hispanic detection rate of 72%.

Our data show that mutations frequent among Hispanic individuals are generally not specific to the Hispanic population, suggesting that a Hispanic-specific mutation panel is not warranted. Only two mutations are unique to the Hispanic population in this study. Of the two, W1089X is relatively frequent (1.2%) but was originally reported in one Turkish and one Egyptian Jewish individual¹⁷ and had not been described in Hispanics until now. The other, 1078delT,¹⁸ was identified on one Hispanic chromosome and no others. Several mutations previously identified in other geographical regions were unexpectedly identified in Hispanics. Examples include 711 + 1G>T (frequency of 6.9% in Tunisia⁵ and 9% in Quebec¹⁹), 1677delTA (originally described in a small Georgian ethnic group in the U.S.S.R.²⁰), and Y1092X (originally identified in the Canadian population²¹ and in the homozygous state in a CF patient of Jewish Egyptian origin³).

With the exception of W1089X, 1078delT, and 2869insG, all other mutations identified in Hispanics were also identified in Caucasians. Eight mutations, including the African American mutation 3120 + 1G>A,⁶ were identified in both Hispanics and African Americans. The mutation $\Delta F311$, found only in an

Hispanic individual, is considered to be African American specific.²² These data support the contention that individuals identified as “Hispanic” are genetically extremely heterogeneous.²³ To increase the detection rate among Hispanics, an expanded, pan-ethnic mutation panel containing both “Caucasian” and “African American” mutations was required.

African American mutation detection

The observed detection rate of 61.9% in 202 African Americans from throughout the United States is lower than the detection rate of 75% reported in 148 African Americans.⁶ However, the observed frequency of $\Delta F508$ (28.7%) is most likely an underestimate because of under-ascertainment of $\Delta F508$ homozygotes, as described in the Caucasian mutation detection section. When $\Delta F508$ is excluded, the detection rate in the present study is 33.2% compared with 27%.⁶ Twenty-one different mutations account for the increase in sensitivity, compared with the 14 mutations detected by Macek et al.⁶ In the absence of referral bias, based on a $\Delta F508$ frequency of 48%,⁶ we predict that a panel of 21 mutations would provide an African American detection rate of 81%. This rate of CF mutation detection approaches that in Caucasians.

As expected, after $\Delta F508$ the most common African American mutation was 3120 + 1G>A, with a frequency (13.8%) comparable to that reported by Macek et al. (12.3%).⁶ Five mutations that are considered to be African American specific (405 + 3A>C, $\Delta F311$, S364P, Y563D, and 3662delA)^{6,12,22} were not observed in any African Americans included in this study, however. It was unexpected that six of the next most common mutations after 3120 + 1G>A would be of Caucasian origin (R1158X, R117H, G551D, 1812–1G>A, 1898 + 1G>A, and R1066C). Of these, R1066C has a frequency of 3.1% in Portugal,⁵ 1812–1G>A was originally identified in 1/50 Spanish CF chromosomes,²⁴ and R1158X was originally identified in an Italian CF patient.¹³ Our detection of R1158X on four African American chromosomes (2.0%) was not anticipated.

Of the 20 mutations that account for the overall detection rate in African Americans when $\Delta F508$ is excluded, nine that account for 23.6% of the chromosomes analyzed are considered to be “African” mutations⁶ (444delA, G330X, G480C, R553X, A559T, 2307insA, 3120 + 1G>A, 3791delC, and S1255X). By comparison, eight “African” mutations accounted for a similar percentage of the chromosomes analyzed (23%) in the study by Macek et al.⁶ In contrast, 11 of the 20 mutations detected in this study are considered to be “Caucasian” mutations and account for 10.5% of the chromosomes analyzed (R117H, 621 + 1G>T, R334W, Q493X, G551D, 1812–1G>A, 1898 + 1G>A, R1066C, R1158X, R1162X, and 3905insT). By comparison, only 4% of the chromosomes analyzed were accounted for by six Caucasian mutations in the study by Macek et al.⁶ We conclude, therefore, that the observed increase in sensitivity in this study was achieved not by increasing the number of African American-specific mutations tested but by screening all African American alleles for “Caucasian” mutations.

Asian and Native American mutation detection

The observed detection rate of 38% in 16 Asian-Americans from throughout the United States is comparable to the detection rate of 33% reported in three Asians.⁸ Our data show that a panel of CF mutations that increases detection rates in U.S. Hispanics, Caucasians, and African Americans does not provide a similar increase in Asians.

By contrast, the detection rate of 81% for Native Americans in this study is strikingly higher than the 4.2% obtained by Grebe et al.⁹ The 12 individuals reported by Grebe et al.⁹ were Pueblo and Navajo Native Americans, but the specific origins of the 21 Native Americans included in the present study are unknown. Our detection rate was achieved with six different mutations compared with the single mutation reported by Grebe et al.⁹ All six mutations were also relatively frequent in Caucasian populations. We did not detect the R1162X mutation, although it has a carrier frequency of 6.7% among Zuni Native Americans.²⁵ Despite its small size, our sample set is the largest reported to date. Our Native American detection rate rivals that found in Caucasians and was achieved by screening for common “Caucasian” mutations.

Rationale for a pan-ethnic mutation panel

Approximately one-third of the individuals in this study were of mixed race or unknown ancestry. These included 722 Mixed/Other Caucasians and 325 Other/Mixed Race/Unknown individuals. To achieve mutation detection rates of 83% and 72%, respectively, a combined total of 56 mutations were required. Our experience shows that, in practice, a pan-ethnic approach to mutation detection is necessary to provide maximal information to individuals of mixed race or unknown ancestry. Our data also show that, even when ethnic background is specified, there are no nonoverlapping “core” mutation panels appropriate to specific ethnic groups. This is a reflection of the genetic heterogeneity intrinsic to the U.S. population and the admixture among ethnic groups. In many cases, individuals are themselves not fully aware of their ethnic backgrounds. It is impractical to expect comprehensive assessment of the ethnic background of every individual referred for testing. We expect that this problem will be magnified when population-based carrier testing becomes the standard-of-care. Our conclusion is similar to that of the ACMG Subcommittee on Cystic Fibrosis Screening, which recently recommended a pan-ethnic mutation panel.²⁶ Our data show that the increase in sensitivity seen in all groups except Ashkenazi Jews compared with previous studies was achieved in part by a pan-ethnic approach to mutation analysis. As the constitution of the U.S. population continues to change, a pan-ethnic approach will be essential.

Optimal CF mutation test panel

Our data show that mutations must be tested in all ethnic groups constituting a heterogeneous population to determine their relative frequencies. Frequencies cannot be predicted by extrapolation. Knowledge of relative frequencies then allows

determination of how many and which CF mutations are appropriately included in a population-screening program. While this study provides the most comprehensive information about mutation distribution in the United States to date, it has the same limitations as other mutation surveys—only characterization of the *CFTR* gene by sequencing would clarify the complete spectrum of mutations likely to be present in the United States. Nevertheless, from the current data set, a core mutation panel can be established that will provide a cost-effective screening alternative to a sequencing test. A conservative mutation panel based on the data presented here would include, at minimum, all mutations detected with a frequency of at least 0.5% at least once in Table 2 (48 mutations). A more comprehensive panel would include, at minimum, all mutations detected with a frequency of at least 0.1% at least once in Table 2 (64 mutations). An optimal panel would also include other relevant mutations, for example, the recently reported 3876delA mutation that is frequent among Hispanic individuals.²⁷ A panel containing mutations with an allele frequency of $\geq 0.1\%$ is consistent with the ACMG Subcommittee recommendations.²⁶ The sizes of the panels suggested here are in line with the 75-mutation panel reported to account for 90% of Spanish chromosomes.²⁸

The ACMG Subcommittee's recommended Standard Mutation Panel²⁶ comprises 25 mutations, all of which were included in the present study. The 25-mutation panel would have detected 4,398 CF chromosomes (75.3%) in our study, whereas the 64-mutation panel detected 4,668 CF chromosomes (81.4%). If the panels were used for general population carrier screening this difference in sensitivity has the practical implication that the recommended panel would miss 1 in 17 carriers of CF in the general population compared with the 64-mutation panel. In the Hispanic or African American populations the recommended panel would miss 1 in 7 or 1 in 11 carriers of CF, respectively, compared with the 64-mutation panel. In practice, missing 1 in 17 carriers would result in 1 in 1,156 ($1/17 \times 1/17 \times 1/4$) unexpected CF births. This is relevant in the context of millions of U.S. births each year.

The observed range of detection in this study is 58 to 95% among all ethnic groups except Asians. These are underestimates, however, because samples referred to our laboratory are depleted of $\Delta F508$ homozygotes. When ascertainment bias is taken into account, the range of detection is 72 to 95%. One argument for offering testing only to Caucasians is supported by data showing screening sensitivities in the range of 80 to 90%. We conclude that a mutation panel containing 50 to 70 mutations will achieve similar sensitivities in all ethnic groups except Asians, and will accommodate a population characterized by increasing admixture. We propose that the approach of using a pan-ethnic, expanded mutation panel to account for genetic admixture and ethnic heterogeneity may provide a practical model for population-based diagnostic testing or carrier screening of other genes as well as the *CFTR* gene.

Acknowledgments

We are grateful for the excellent technical contributions of Nancy Cataldo, Michelle Durwin, Joanne Garver, Nichole Napolitano, and Rebekah Silka.

References

1. Hamosh A, Fitzsimmons SC, Macek M, 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.
2. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC. Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989;245:1072–1080.
3. Cystic Fibrosis Genetic Analysis Consortium, 2000 <http://www.genet.sickkids.on.ca/cftr>
4. Cystic Fibrosis Genetic Analysis Consortium. Population variation of common cystic fibrosis mutations. *Hum Mutat* 1994;4:167–177.
5. Estivill X, Bancells C, Ramos C, and the Biomed CF Mutation Analysis Consortium. Geographic distribution and regional origin of 272 cystic fibrosis mutations in European populations. *Hum Mutat* 1997;10:135–154.
6. 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.
7. Grebe TA, Seltzer WK, DeMarchi J, Silva DK, Doane WW, Gozal D, Richter SF, Bowman CM, Norman RA, Rhodes SN, Hernried LS, Murphy S, Harwood IR, Accurso FJ, Jain KD. Genetic analysis of Hispanic individuals with cystic fibrosis. *Am J Hum Genet* 1994;54:443–446.
8. Macek M Jr, Mercier B, Mackova A, Miller PW, Hamosh A, Ferec C, Cutting GR. Sensitivity of the denaturing gradient gel electrophoresis technique in detection of known mutations and novel Asian mutations in the *CFTR* gene. *Hum Mutat* 1997;9:136–147.
9. Grebe TA, Doane WW, Richter SF, Clericuzio C, Norman RA, Seltzer WK, Rhodes SN, Goldberg BE, Hernried LS, McLure M, Kaplan G. Mutation analysis of the cystic fibrosis transmembrane regulator gene in Native American populations of the Southwest. *Am J Med Genet* 1992;51:736–740.
10. Grody WG. Cystic fibrosis. Molecular diagnosis, population screening, and public policy. *Arch Pathol Lab Med* 1999;123:1041–1046.
11. Genetic Testing for Cystic Fibrosis. NIH Consensus Statement 1997;15:1–37.
12. Cutting GR, Curristan SM, Nash E, Rosenstein BJ, Lerer I, Abeliovich D, Hill A, Graham C. Analysis of a four diverse population groups indicates that a subset of cystic fibrosis mutations occur in common among Caucasians. *Am J Hum Genet* 1992;50:1185–1194.
13. Ronchetto P, Orriols JTT, Fanen P, Cremonesi L, Ferrari M, Magnani C, Seia M, Goossens M, Romeo G, Devoto M. A nonsense mutation (R1158X) and a splicing mutation (3849+4A>G) in exon 19 of the cystic fibrosis transmembrane conductance regulator gene. *Genomics* 1991;12:417–418.
14. Shuber AP, Skoletsky J, Stern R, Handelin B. Efficient 12-mutation testing in the *CFTR* gene: a general model for complex mutation analysis. *Hum Mol Genet* 1993; 2:153–158.
15. Zielenski J, Fujiwara TM, Markiewicz, Paradis AJ, Anacleto AI, Richards B, Schwartz RH, Klinger KW, Tsui L-C, Morgan K. Identification of the M1101K mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene and complete detection of cystic fibrosis mutations in the Hutterite population. *Am J Hum Genet* 1993;52:609–615.
16. Abeliovich D, Lavon IP, Lerer I, Cohen T, Springer C, Avital A, Cutting G. Screening for five mutations detects 97% of cystic fibrosis chromosomes and predicts a carrier frequency of 1:29 in the Jewish Ashkenazi population. *Am J Hum Genet* 1992;51: 951–956.
17. Shoshani T, Augarten A, Yahav J, Gazit E, Kerem B. Two novel mutations in the *CFTR* gene: W1089X in exon 17b and 4010delTATT in exon 21. *Hum Mol Genet* 1994;3:657–658.
18. Claustres M, Gerrard B, White MB, DesGeorges M, Kjellberg P, Rollin B, Dean M. A new mutation (1078delT) in exon 7 of the *CFTR* gene in a southern French adult with cystic fibrosis. *Genomics* 1992;13:907–908.
19. Rozen R, De Braekeleer M, Daigneault J, Ferreira-Rajabi L, Gerdes M, Lamoureux L, Aubin G, Simard F, Fujiwara TM, Morgan K. Cystic fibrosis mutations in French Canadians: three *CFTR* mutations are relatively frequent in a Quebec population with an elevated incidence of cystic fibrosis. *Am J Med Genet* 1992;42:360–364.
20. Ivaschenko TE, White MB, Dean M, Baranov VS. A deletion of two nucleotides in exon 10 of the *CFTR* gene in a Soviet family with cystic fibrosis causing early infant death. *Genomics* 1991;10:298–299.

21. Bozon D, Zielenski J, Rininsland F, Tsui L-C. Identification of four new mutations in the cystic fibrosis transmembrane conductance regulator gene: I148T, L1077P, Y1092X, 2183AA>G. *Hum Mutat* 1994;3:330–332.
22. Friedman KJ, Leigh MW, Czarnecki, Feldman GL. Cystic fibrosis transmembrane-conductance regulator mutations among African Americans. *Am J Hum Genet* 1998;62:195–196.
23. Arzimanoglou I, Tuchman A, Li Z, Gilbert F, Denning C, Valverde K, Zar H, Quittel L. Cystic fibrosis carrier screening in Hispanics. *Am J Hum Genet* 1995; 56:544–547.
24. Chillon M, Casals T, Gimenez J, Nunes V, Estivill X. Analysis of the CFTR gene in the Spanish population: SSCP-screening for 60 known mutations and identification of four new mutations (Q30X, A120T, 1812–1G>A, and 3667del4). *Hum Mutat* 1994; 3:223–230.
25. Kessler D, Moehlenkamp C, Kaplan G. Determination of cystic fibrosis carrier frequency for Zuni native Americans of New Mexico. *Clin Genet* 1996;49:95–97.
26. Subcommittee on Cystic Fibrosis Screening, Accreditation of Genetic Services Committee, American College of Medical Genetics. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet in Med* 2001. In press.
27. Wang J, Bowman MC, Hsu E, Wertz K, Wong CL-J. A novel mutation in the CFTR gene correlates with severe clinical phenotype in seven Hispanic patients. *J Med Genet* 2000;37:215–218.
28. Casals T, Ramos MD, Gimenez J, Larriba S, Nunes V, Estivill X. High heterogeneity for cystic fibrosis in Spanish families: 75 mutations account for 90% of chromosomes. *Hum Genet* 1997;101:365–370.