A Genotypic-Oriented View of *CFTR* Genetics Highlights Specific Mutational Patterns Underlying Clinical Macrocategories of Cystic Fibrosis

Marco Lucarelli,[1,2,3] Sabina Maria Bruno,[1] Silvia Pierandrei,[1,2] Giampiero Ferraguti,[1] Antonella Stamato,[3,4,5] Fabiana Narzi,[3,4,5] Annalisa Amato,[3,4,5] Giuseppe Cimino,[3,5] Serenella Bertasi,[3,5] Serena Quattrucci,[3,4,5] and Roberto Strom[1,3]

[1]Department of Cellular Biotechnologies and Hematology, Sapienza University of Rome, Rome, Italy; [2]Pasteur Institute, Cenci Bolognetti Foundation, Sapienza University of Rome, Rome, Italy; [3]Policlinico Umberto I Hospital, Rome, Italy; [4]Department of Pediatrics, Sapienza University of Rome, Rome, Italy; and [5]Cystic Fibrosis Reference Center of Lazio Region, Rome, Italy

Cystic fibrosis (CF) is a monogenic disease caused by mutations of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. The genotype–phenotype relationship in this disease is still unclear, and diagnostic, prognostic and therapeutic challenges persist. We enrolled 610 patients with different forms of CF and studied them from a clinical, biochemical, microbiological and genetic point of view. Overall, there were 125 different mutated alleles (11 with novel mutations and 10 with complex mutations) and 225 genotypes. A strong correlation between mutational patterns at the genotypic level and phenotypic macrocategories emerged. This specificity appears to largely depend on rare and individual mutations, as well as on the varying prevalence of common alleles in different clinical macrocategories. However, 19 genotypes appeared to underlie different clinical forms of the disease. The dissection of the pathway from the *CFTR* mutated genotype to the clinical phenotype allowed to identify at least two components of the variability usually found in the genotype–phenotype relationship. One component seems to depend on the genetic variation of *CFTR*, the other component on the cumulative effect of variations in other genes and cellular pathways independent from *CFTR*. The experimental dissection of the overall biological *CFTR* pathway appears to be a powerful approach for a better comprehension of the genotype–phenotype relationship. However, a change from an allele-oriented to a genotypic-oriented view of *CFTR* genetics is mandatory, as well as a better assessment of sources of variability within the *CFTR* pathway.
Online address: http://www.molmed.org
doi: 10.2119/molmed.2014.00229

INTRODUCTION

Cystic fibrosis (CF) (MIM 219700) is the most common lethal genetic disease among Caucasians. It is caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene (1–3). The *CFTR* is a transmembrane multifunctional protein expressed mainly at the apical membrane of epithelial cells (4).

The airway of CF patients with polysymptomatic forms is affected by a *CFTR* deficiency that causes anomalous ion transport, altered water absorption, sticky mucus, multiresistant bacterial infections and respiratory impairment (1–4). CF is a multiorgan disease with poly-, oligo- and mono-symptomatic forms (5,6) that are diagnosed by means of a combination of

genetic analysis, biochemical assessment and clinical presentation (7). The different forms of the disease can be grouped, according to recent guidelines and recommendations (7,8), as classic CF with or without pancreas sufficiency, *CFTR*-related disorders (CFTR-RD) and congenital bilateral absence of vas deferens (CBAVD). The most severe clinical findings are pulmonary symptoms. An as yet unclear relationship between the genotype and phenotype has been highlighted (1,5,9). Although almost 2,000 sequence variations of the *CFTR* gene are known (10), few of them have actually been functionally characterized. A considerable effort is currently being made to identify the disease-causing mutations (11) (Clinical and Functional Translation of *CFTR* [CFTR2] database) and to group them into mutational classes (12–14).

Although the model of an inverse correlation between protein residual function and phenotype severity is generally accepted, the poor functional characterization of most of the *CFTR* sequence variations that have been identified, especially with regard to their quantitative effect, hampers the practical application of this concept. In addition to these critical issues, the problems encountered in recognizing the different clinical forms of CF, as well as the influence of modifier genes, further complicate the framework.

The aim of the present work is to improve our understanding of the genotype–phenotype relationship in CF by means of a genotypic-oriented approach based on the selection of specific mutational patterns underlying different clinical forms of the disease. A general approach in which a specific pathogenic role is assigned to each *CFTR* sequence variation, also taking into account disease severity, is proposed. The results of this study also shed light on the sources of variability, acting at different levels, involved in the path from genotype to protein residual function and, eventually, to clinical phenotype.

MATERIALS AND METHODS

Case Series: Characterization and Diagnostic Criteria

We evaluated all patients already diagnosed and enrolled at the CF Reference Center of the Lazio Region up to 1996 and all subsequent new diagnoses from 1996 to 2012. This step yielded a consecutive case series comprising 692 patients. A total of 82 patients were excluded from this study because of incomplete genetic, biochemical, microbiological, clinical and/or family data. The remaining 610 patients (1,220 alleles) with complete data, mainly from central Italy, were enrolled in the study. According to generally accepted procedures, clinical (15,16), instrumental, laboratory (17) (Supplementary Table S1), microbiological (18,19) and biochemical and genetic (see below) evaluations were performed. Depending on these characterizations, the patients were

classified according to recent CF guidelines and recommendations (7,8,17,20) in the following four clinical macrocategories (also called "populations" in the text): (a) CF with pancreatic insufficiency (CF-PI) (354 patients, 708 alleles); (b) CF with pancreatic sufficiency (CF-PS) (138 patients, 276 alleles); (c) mono- or oligosymptomatic forms of CF, which for the purposes of this work included both *CFTR*-related disorders and atypical CF forms (here called CFTR-RD, 71 patients, 142 alleles); and (d) congenital bilateral absence of vas deferens (CBAVD), which for the purposes of this work was selected as the only clinical manifestation (with no other CF symptoms) (CBAVD, 47 patients, 94 alleles). When the diagnosis according to CF guidelines and recommendations was in contrast to clinical evidence, the latter prevailed.

Ethics Statement

Informed consent was obtained from every patient (or parents) before enrollment. The study was approved by the institutional ethics committee and carried out according to the Helsinki Declaration.

Biochemical Characterization

All patients underwent a sweat test, at least twice, performed by means of a quantitative pilocarpine iontophoresis method (21) by using the Macroduct device (Delcon, Milan, Italy) for sweat collection and the PCL M3 chloride analyzer Jenway (VWR International, Milan, Italy) for measurement. In accordance with recent guidelines (17), the sweat test in subjects up to 6 months of age was considered negative if [Cl$^-$] was <30 mEq/L, pathologic if ≥60 mEq/L and borderline if in the 30–59 mEq/L range; for all other subjects, the sweat test was considered negative if <40 mEq/L, pathologic if ≥60 mEq/L and borderline if in the 40–59 mEq/L range.

Exocrine pancreatic function was evaluated by the dosage of fecal elastase 1 (22) by using the immunometric pancreatic fecal elastase test (Meridian Bioscience, Milan, Italy). The status of pancreatic sufficiency for all CF-PS,

CFTR-RD and CBAVD patients was ascertained from the nonpathological levels of fecal elastase 1 (>200 µg/g) in at least two independent dosages as well as from the absence of steatorrhea. All the patients with elastase 1 level well below this threshold were characterized by reduced growth and were classified as CF-PI.

Mutational Search Strategy

DNA was extracted from peripheral blood by using the QIAamp DNA blood midi kit (Qiagen, Hilden, Germany). The mutational search on the *CFTR* gene (RefSeq NM_000492.3, NG_016465.3) was initially conducted by using a multistep approach, with the progressive application of five sequential steps for the analysis of the following:

(a) the 32 most common mutations worldwide, by means of the CF-OLA assay (Abbott, Wiesbaden, Germany);
(b) the 14 most frequent mutations in our geographical area, by means of our assay based on primer extension (CF-SNAP+20);
(c) the (TG)mTn variant tracts, specifically the (TG)13T5 (c.[1210-14TG[13]; 1210-12T[5]]), the (TG)12T5 (c.[1210-14TG[12];1210-12T[5]]) and the (TG)11T5 (c.[1210-14TG[11];1210-12T[5]]), by means of our assay (23) based on DNA sequencing;
(d) the proximal 5'-flanking, all exons and adjacent intronic zones, by means of our assay based on DNA sequencing (24), always applied to completion when included (this step is referred to as "SEQ" in this article); and
(e) the seven most frequent macrodeletions worldwide, by means of the FC del assay (Nuclear Laser Medicine, Milan, Italy) (this step is referred to as "DEL" in this article).

The mutational search was usually interrupted when the first two *CFTR* mutations already characterized as disease-causing were found on different alleles. Those genotypes reported to have at least one unknown allele (see Results) underwent all the steps, including DEL step.

To shed further light on the relationship between genotype and phenotype, for

specific genotypes, the mutational search was extended up to the SEQ step, even if two mutations on different alleles had already been found. These specific genotypes were the following (see Results): the 19 genotypes found in different clinical macrocategories, the 19 genotypes involving the (TG)13T5 (c.[1210-14TG[13]; 1210-12T[5]]) , (TG)12T5 (c.[1210-14TG[12]; 1210-12T[5]]) or (TG)11T5 (c.[1210-14TG[11]; 1210-12T[5]]) variant tracts and all genotypes involving mutations with a controversial functional effect, as listed in Supplementary Table S2. Furthermore, among these specific genotypes, those found in CF-PI also underwent the DEL step. The controversial complex allele [1249-8A>G; G576A;R668C] (c.[1117-8A>G;1727G>C; 2002C>T]) and mutations G1069R (p.Gly1069Arg), D614G (p.Asp614Gly), S42F (p.Ser42Phe) and S912L (p.Ser912Leu) should also be considered as part of this extension, even if not found in CF-PI but studied up to the DEL step because they are found in genotypes with an unknown allele. For the specific protocol of mutational search applied to each genotype, see Supplementary Table S4.

The mutational search from step (a) to step (d) [CF-OLA, CF-SNAP+20, (TG)mTn and SEQ] was performed in a 96-well format, using a semi-automated platform made up of a robotic system (Microlab Starlet; Hamilton) for the reaction setup and two genetic analyzers (ABI PRISM 3100 *Avant* and ABI PRISM 3130 *xl*; Applied Biosystems [Thermo Fisher Scientific Inc., Waltham, MA, USA]) for the development of the electropherograms. For data analysis, the specific CF-OLA template (Abbott) and our specific CF-SNAP+20 template, based, respectively, on the Genotyper and GeneMapper software (Applied Biosystems [Thermo Fisher Scientific]) were used for the CF-OLA and CF-SNAP+20 steps, respectively. The results of the (TG)mTn tracts were analyzed as previously described (23). Sequences obtained in the SEQ step were analyzed by using our specific template based on Seqscape software (Applied Biosystems [Thermo Fisher Scientific]) (25). The segregation of

all mutated alleles was ascertained by analysis of parents.

Mutations are reported with both the old (legacy name) and the new nomenclature (HGVS name) in all of the tables and in the text; for practical purposes, the legacy name alone is used in the figures.

Pathogenic Classification of Mutated Alleles

The general principle applied for the clinical classification of alleles was that the clinical effect is determined by the overall residual functionality of the CFTR protein and thus, ultimately, by the allele with the highest functionality. A set of three rules was established to assign a phenotypic effect to each mutated allele found in patients and to determine its ability to induce clinical manifestations belonging to one (or more) of the four clinical macrocategories identified. It was possible to apply this procedure because the diagnosis in patients had previously been made on the basis not only of guidelines and recommendations, but also of a conclusive clinical assessment. The rules were subsequently applied and their performance was experimentally validated according to the clinical classification of patients. Upon the application of each rule, a suitability control was applied to the previously classified alleles, which in some cases led to alleles being reclassified. The first rule consisted in the assignment of all alleles found in homozygosis to the specific macrocategory in which they had been identified. The second rule consisted in the classification, as CF-PI–causing alleles, of all alleles found in compound heterozygosis in the CF-PI macrocategory. The third rule consisted in the classification of all alleles found in CF-PS, CFTR-RD and/or CBAVD in compound heterozygosis with a previously classified allele (see the Supplementary Materials for the algorithm, Supplementary Figure S1 and examples). By applying these rules, it proved possible to assign each allele to (a) a single macrocategory and label it as an allele with a unique phenotypic ef-

fect; (b) more than one macrocategory and label it as an allele with a variable effect; and (c) no category and label it with an uncertain classification.

Statistical Analysis

Contingency tables, analysis of variance (ANOVA), Student *t* test and Bonferroni multiple comparison test were used for the statistical analysis of experimental data, by using the SPSS software (SPSS [IBM, Armonk, NY, USA]).

All supplementary materials are available online at www.molmed.org.

RESULTS

Allele Frequencies Reveal Genetic Heterogeneity between Clinical Macrocategories

The results on allele frequencies are reported in Figure 1A and Supplementary Table S3. We identified 125 different *CFTR* mutated alleles. Eleven mutations were novel (described below). We also identified 10 complex alleles (with two or more mutations in *cis* on the same allele, described below), two of which included two of the novel mutations. The different mutated alleles found were 69 in CF-PI, 60 in CF-PS, 37 in CFTR-RD and 24 in CBAVD. Forty-three (34.4%) of the 125 mutated alleles were found in at least two different macrocategories (4 alleles in all 4 populations, 14 alleles in 3 different populations, 25 alleles in 2 different populations) (Figure 1B), whereas 82 alleles (65.6%) were exclusive to a single population (Supplementary Table S3). Among the latter, 39 were found exclusively in CF-PI (56.5% of CF-PI alleles), 21 exclusively in CF-PS (35.0% of CF-PS alleles), 16 exclusively in CFTR-RD (43.2% of CFTR-RD alleles) and 6 exclusively in CBAVD (25.0% of CBAVD alleles). By summing all the populations (Supplementary Table S3), the frequency of the F508del (p.Phe508del) mutation was 0.400; the number of additional moderately frequent mutated alleles, with a prevalence ≥0.008 in CF (PI + PS), was 16 of 125 (12.8%), with an overall prevalence of 0.343; the number of rare nonindividual
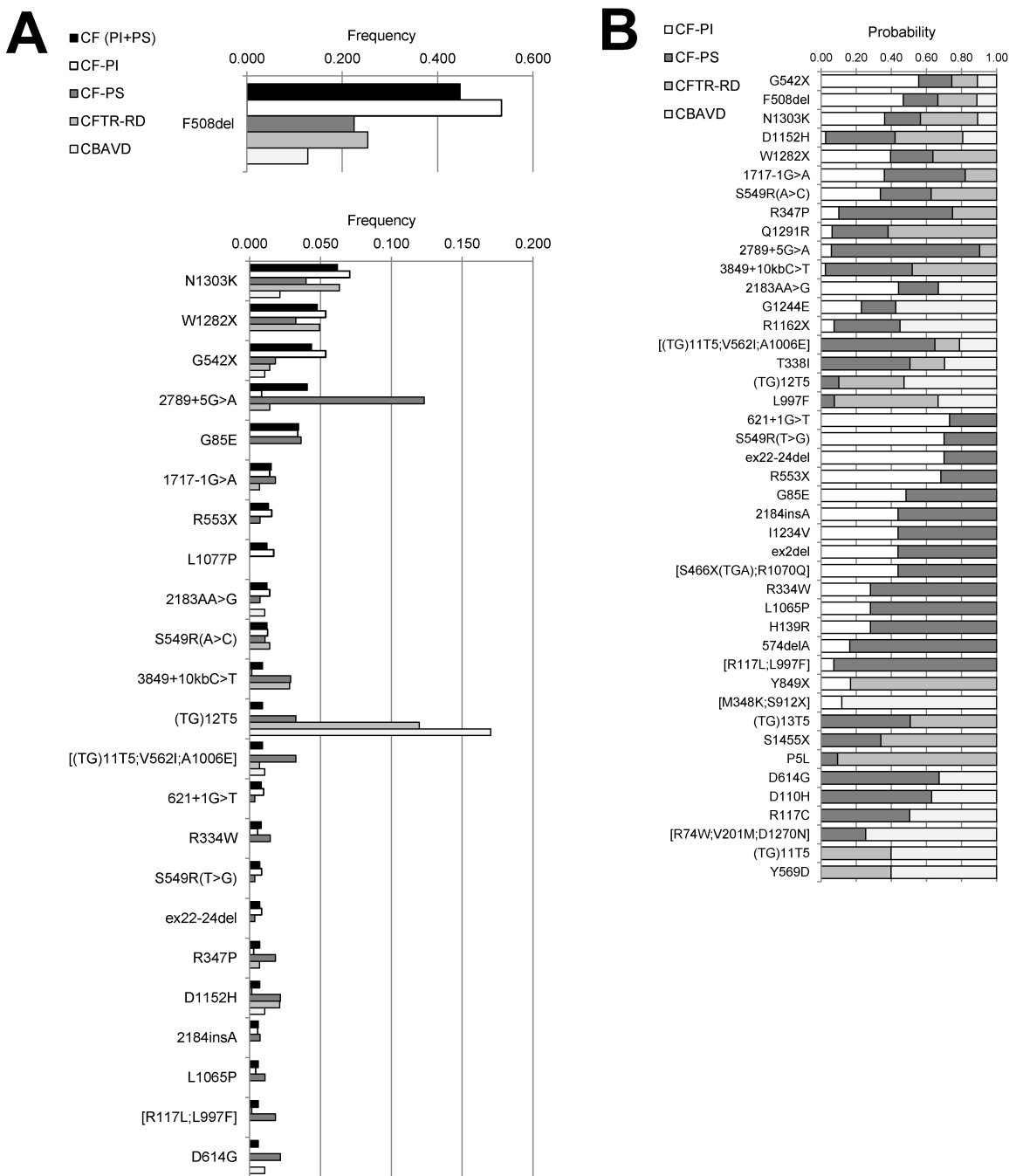
Figure 1. Allele frequencies and probabilities. (A) Frequencies of mutated alleles, with a prevalence ≥0.006 in the CF (PI + PS) population, are reported in frequency decreasing order according to CF (PI + PS); the last mutation with a prevalence = 0.008 in the CF (PI + PS) population is the R334W (p.Arg334Trp). (B) Allele distribution between populations. Only alleles found in at least two different populations are shown. The length of the bars is proportional to the frequency of each allele in the specific population; it represents the probability that the specific allele is found in each clinical form. See Supplementary Table S3 for allele HGVS name.

(found in at least two unrelated patients) mutations was 48 of 125 (38.4%), with an overall prevalence of 0.152; lastly, the number of individual mutations (found in only one patient or in siblings from one family) was 60 of 125 (48.0%), with an overall prevalence of 0.056. Each population displayed a peculiar mutational pattern (overall χ^2 $p < 0.0001$; for each population pair χ^2 $p < 0.0001$, with the exception of the CFTR-RD versus CBAVD comparison, which was χ^2 $p < 0.05$).

When the CF (PI + PS) population alone was considered (Figure 1A, Supplementary Table S3), 101 mutations were found. The most frequent mutation was F508del (p.Phe508del), with an overall frequency of 0.447, and a well-differentiated frequency of 0.534 in CF-PI and 0.225 in CF-PS. Only 28 of the 101 mutations were found in both CF-PI and CF-PS, including 14 mutations also found in other populations (Figure 1B). Another 41 mutations were found in CF-PI but not in CF-PS, including 39 that were found exclusively in CF-PI; the other two were found also in other populations. Another 32 mutations were found in CF-PS, although not in CF-PI, including 21 that were found exclusively in CF-PS and 11 found also in other populations.

A total of 37 mutations were found in the CFTR-RD population (Figure 1A, Supplementary Table S3). The most frequent mutation in this population was again F508del (p.Phe508del), with a frequency of 0.254, which is similar to that of CF-PS. Sixteen of the 37 mutations found were exclusive to CFTR-RD. By contrast, the other 21 mutations were also found in other populations (Figure 1B).

Only 24 mutations were identified in the CBAVD population (Figure 1A, Supplementary Table S3). The most frequent mutation was the (TG)12T5 (c.[1210-14TG[12];1210-12T[5]]) variant allele, with a frequency of 0.170; F508del (p.Phe508del) was the second most prevalent mutation, with a frequency of 0.128. Six of the 24 mutations found were exclusive to the CBAVD population, whereas the other 18 were also found in other populations (Figure 1B).

Genetic Heterogeneity between Populations Is Amplified at a Genotypic Level

The results on genotype frequencies are reported in Figure 2A and Supplementary Table S4. A total of 225 different *CFTR* mutated genotypes were identified, 11 and 20 of which include, respectively, the novel mutations and the complex alleles found. The different mutated genotypes found were 115 in CF-PI, 77 in CF-PS, 44 in CFTR-RD and 12 in CBAVD. Nineteen

(8.4%) of the 225 different genotypes were found in at least two different populations (Figure 2B), whereas the remaining 206 (91.6%) were exclusive to a single population (Supplementary Table S4). In particular, 105 were exclusively found in CF-PI (91.3% of CF-PI genotypes), 58 exclusively in CF-PS (75.3% of CF-PS genotypes), 35 exclusively in CFTR-RD (79.5% of CFTR-RD genotypes) and 8 exclusively in CBAVD (66.7% of CBAVD genotypes). Fifty-nine (26.2%) of the 225 genotypes were found in at least two unrelated individuals. These 59 nonindividual genotypes included 4 and 15 (in total, accounting for 32.2%) that were found in three and two different populations, respectively (Figure 2B), and that consequently underwent a specific extensive genetic analysis (see Materials and Methods); the other 40 genotypes (67.8%) proved to be associated with a single population. The remaining 166 genotypes (73.8%) were found to be individual genotypes found only once in single patients (146 genotypes) or only in siblings from a single family (20 genotypes), associated with a single population.

Taking all the populations together (Supplementary Table S4), the frequency of the homozygous F508del/F508del (p.[Phe508del];[Phe508del]) genotype was 0.180. The number of additional moderately frequent genotypes, with a prevalence ≥0.008 in CF (PI + PS), was 15 out of 225 (6.7%), with an overall prevalence of 0.259. The number of rare nonindividual (found in at least two unrelated patients) genotypes was 43 out of 225 (19.1%), with an overall prevalence of 0.185. Lastly, the number of individual genotypes (found in only one patient or in siblings from one family) was 166 out of 225 (73.8%), with an overall prevalence of 0.305. As for mutated alleles, each population displayed a peculiar pattern of genotypes (overall χ^2 $p < 0.0001$; for each population pair χ^2 $p < 0.0001$).

If we consider the CF (PI + PS) population alone (Figure 2A, Supplementary Table S4), 182 different genotypes were found with the homozygote F508del/F508del (p.[Phe508del];[Phe508del]), which was the

most frequent genotype, with an overall frequency of 0.224. However, although this genotype was also the most frequent in the CF-PI population, with a frequency of 0.311, it was never found in the CF-PS population, in which the most frequent genotype was the F508del/2789+5G>A (c.[1521_1523delCTT];[2657+5G>A]), with a frequency of 0.101. Only 10 of the 182 genotypes found in this mixed population were found in both CF-PI and CF-PS (Figure 2B), but were never found in either CFTR-RD or CBAVD. An additional 105 genotypes were found exclusively in CF-PI but were never found in CF-PS or the other populations. Another 67 genotypes were found in CF-PS but not in CF-PI, with 58 of them being found exclusively in CF-PS and nine also found in CFTR-RD and/or CBAVD.

A total of 44 genotypes were found in the CFTR-RD population (Figure 2A, Supplementary Table S4). The F508del/F508del (p.[Phe508del];[Phe508del]) genotype was not found in this population either (nor was it found in CF-PS). The most frequent genotype was the F508del/(TG)12T5 (c.[1521_1523delCTT];[1210-14TG[12];1210-12T[5]]) with a frequency of 0.070. The 44 genotypes found included 35 that were exclusive to CFTR-RD. By contrast, nine were also found in CF-PS and/or CBAVD (Figure 2B). No genotype of this population was also found in CF-PI.

Only 12 genotypes were identified in the CBAVD population (Figure 2A, Supplementary Table S4). As occurred in the CF-PS and CFTR-RD populations, the F508del/F508del (p.[Phe508del];[Phe508del]) genotype was not found in CBAVD either. The most frequent genotype was the F508del/(TG)12T5 (c.[1521_1523delCTT];[1210-14TG[12];1210-12T[5]]), with a frequency of 0.213. Eight of the 12 genotypes found were exclusive to the CBAVD population. The remaining four were also found in both CF-PS and CFTR-RD (Figure 2B), whereas no genotype of CBAVD population was also found in CF-PI.

Phenotypic Description of the 11 Novel Mutations

The characteristics of the 11 novel mutations found are reported in Tables 1–3
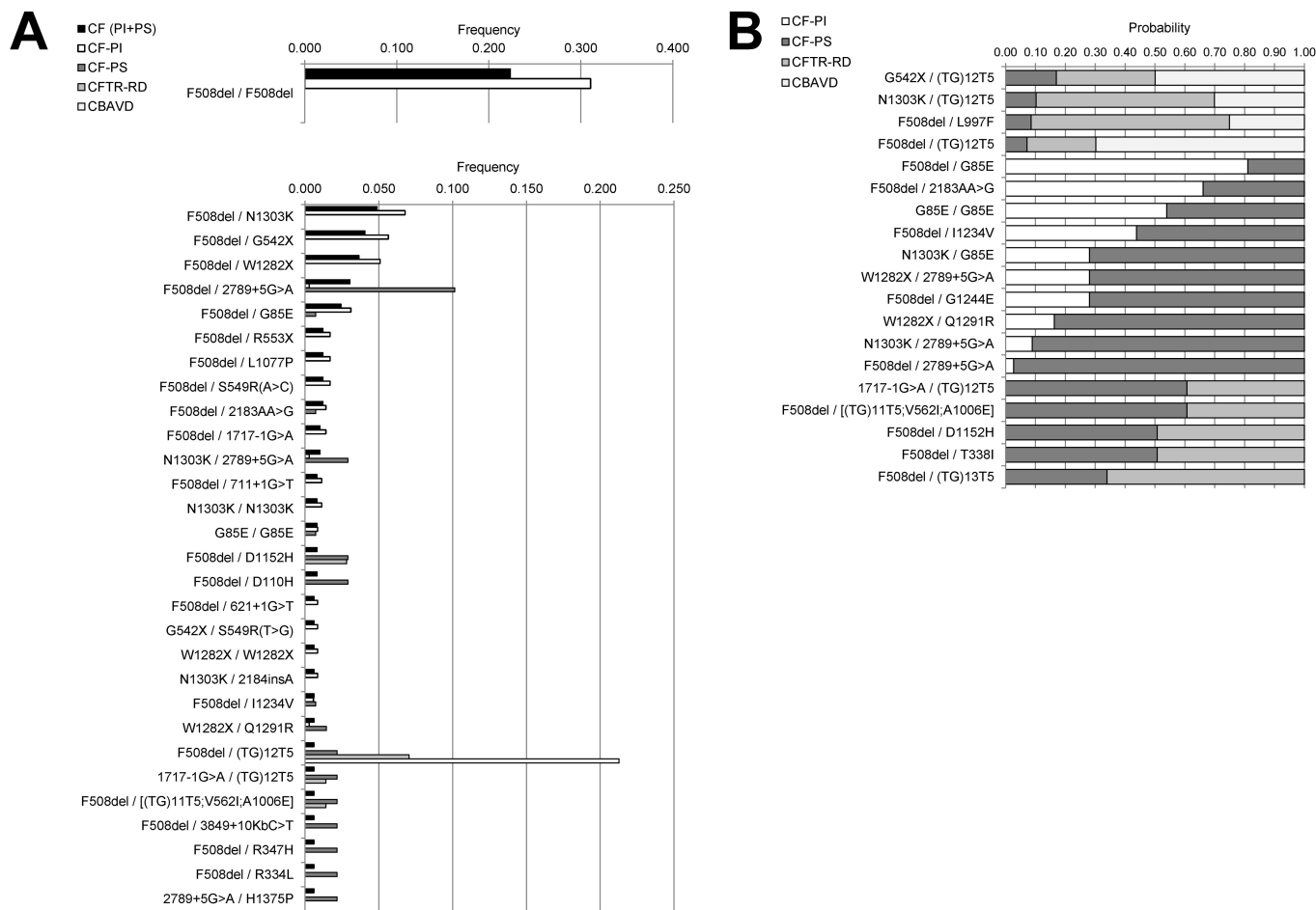
Figure 2. Genotype frequencies and probabilities. (A) Frequencies of mutated genotypes, with a prevalence ≥0.006 in the CF (PI + PS) population, are reported in frequency decreasing order according to CF (PI + PS); the last genotype with a prevalence = 0.008 in the CF (PI + PS) population is the F508del/D110H (p.[Phe508del];[Asp110His]). (B) Genotype distribution between populations. Only genotypes found in at least two different populations are shown. The length of the bars is proportional to the frequency of each genotype in the specific population; it represents the probability that the specific genotype is found in each clinical form. See Supplementary Table S4 for genotype HGVS name.

and summarized below. Their absence in at least 100 subjects (200 alleles) from the general population was verified.

The E479X (p.Glu479*) mutation was found in a novel complex allele [E479X;V754M] (p.[Glu479*;Val754Met]) in a CF-PI male patient, enrolled at 1.5 years of age on the basis of symptoms, with a F508del/[E479X;V754M] (p.[Phe508del];[Glu479*;Val754Met]) genotype. The average sweat test was 106 ± 9 mEq/L. Respiratory manifestations were already present at diagnosis, although with no pulmonary bacterial isolates. The patient is now 37 years old, with worsened respiratory symptoms

and chronic bacterial colonization.

The K442X (p.Lys442*) mutation was found in a CF-PI male patient with a F508del/K442X (p.[Phe508del];[Lys442*]) genotype. The average sweat test was 82 ± 8 mEq/L. The patient was enrolled at 2 months of age on the basis of neonatal screening (26,27) with no symptoms or bacterial pulmonary isolates. He is now 5 years old and displays respiratory manifestations, although with no bacterial isolates.

The D529N (p.Asp529Asn) mutation was found in a CF-PI female patient with a F508del/D529N (p.[Phe508del];[Asp529Asn]) genotype. The average sweat test was 42 ±

5 mEq/L. A late diagnosis was performed at 32 years of age, when severe respiratory manifestations as well as pulmonary bacterial isolates were already present. The patient is now 39 years old and has been displaying persistent severe pulmonary manifestations with chronic bacterial colonization.

The T465N (p.Thr465Asn) mutation was found in a CF-PI male patient with a W1282X/T465N (p.[Trp1282*];[Thr465Asn]) genotype. The average sweat test was 83 ± 7 mEq/L. Meconium ileus was present at diagnosis, which was made at 3 months of age. No other symptoms, respiratory manifestations or pulmonary

bacterial isolates were present. He died at 33 years of age, with severe pulmonary manifestations, chronic bacterial colonization, liver disease and cholelithiasis.

The W19X(TAG) (p.Trp19*) mutation was found in a CF-PI male patient with a G542X/W19X(TAG) (p.[Gly542*];[Trp19*]) genotype. The average sweat test was 58 ± 5 mEq/L. Diagnosis was performed at birth, when the patient exhibited meconium ileus. No other symptoms, respiratory manifestations or pulmonary bacterial isolates were present. The patient is now 3 years old and displays cholelithiasis and mild respiratory manifestations with intermittent bacterial colonization.

The H1375P (p.His1375Pro) mutation was found in 3 CF-PS patients (a brother and sister and a third unrelated male patient) with the same 2789+5G>A/H1375P (c.[2657+5G>A];[4124A>C]) genotype. The average sweat tests ranged from 63 ± 2 to

Table 1. Genetic, biochemical, microbiological and clinical characterization of the patients with the 11 novel mutations found: position and nomenclature of novel mutations found.

Old nomenclature (legacy name)			New nomenclature (HGVS name)		
Position	Nucleotidic notation	Aminoacidic notation	Position	Nucleotidic notation	Aminoacidic notation
Exon 10	1567G>T	E479X	exon 11	c.1435G>T	p.Glu479*
Exon 9	1456A>T	K442X	exon 10	c.1324A>T	p.Lys442*
Exon 11	1717G>A	D529N	exon 12	c.1585G>A	p.Asp529Asn
Exon 10	1526C>A	T465N	exon 11	c.1394C>A	p.Thr465Asn
Exon 2	188G>A	W19X(TAG)	exon 2	c.56G>A	p.Trp19*(TAG)
Exon 22	4256A>C	H1375P	exon 25	c.4124A>C	p.His1375Pro
Exon 13	2467C>T	Q779X	exon14	c.2335C>T	p.Gln779*
Exon 20	3871G>C	G1247R(G>C)	exon 23	c.3739G>C	p.Gly1247Arg
Exon 20	3862G>A	G1244R	exon 23	c.3730G>A	p.Gly1244Arg
Intron 7	1249-8A>G	—	intron 8	c.1117-8A>G	—
Exon 3	299A>G	E56G	exon 3	c.167A>G	p.Glu56Gly

91 ± 8 mEq/L. The diagnosis of the male sibling was performed at 32 years of age on the basis of symptoms, when some pulmonary manifestations were present, although with no bacterial isolates, and de-

hydration occurred. He is now 41 years old, no longer displays pulmonary symptoms but has intermittent bacterial colonization. The female sibling was enrolled at 33 years of age because of familiarity

Table 2. Genetic, biochemical, microbiological and clinical characterization of the patients with the 11 novel mutations found: *CFTR* genotypes, sweat test values, seminal evaluation, cause of enrollment and final diagnosis of patients.

Patient	Genotype[a] Legacy name	HGVS name	Gender	Average sweat test value[b] (mEq/L)	Semen analysis	Cause of enrollment	Diagnosis
1	F508del/[E479X;V754M]	c.[1521_1523delCTT];[1435G>T;2260G>A] p.[Phe508del];[Glu479*;Val754Met]	M	106 ± 9	nd	Symptoms	CF-PI
2	F508del/K442X	c.[1521_1523delCTT];[1324A>T] p.[Phe508del];[Lys442*]	M	82 ± 8	Too young	Neonatal screening	CF-PI
3	F508del/D529N	c.[1521_1523delCTT];[1585G>A] p.[Phe508del];[Asp529Asn]	F	42 ± 5	—	Symptoms	CF-PI
4	W1282X/T465N	c.[3846G>A];[1394C>A] p.[Trp1282*];[Thr465Asn]	M	83 ± 7	nd	Symptoms	CF-PI
5	G542X/W19X(TAG)	c.[1624G>T];[56G>A] p.[Gly542*];[Trp19*]	M	58 ± 5	Too young	Symptoms	CF-PI
6[c]	2789+5G>A/H1375P	c.[2657+5G>A];[4124A>C]	F	91 ± 8	—	Familiarity	CF-PS
7[c]	2789+5G>A/H1375P	c.[2657+5G>A];[4124A>C]	M	76 ± 9	OA	Symptoms	CF-PS
8	2789+5G>A/H1375P	c.[2657+5G>A];[4124A>C]	M	63 ± 2	OA	Symptoms	CF-PS
9[d]	[(TG)11T5;V562I;A1006E]/Q779X	c.[1210-14TG[11];1210-12T[5];1684G>A;3017C>A];[2335C>T]	M	70 ± 15	Too young	Neonatal screening	CF-PS
10[d]	[(TG)11T5;V562I;A1006E]/Q779X	c.[1210-14TG[11];1210-12T[5];1684G>A;3017C>A];[2335C>T]	F	62 ± 17	—	Neonatal screening, familiarity	CF-PS
11	W1282X/G1247R(G>C)	c.[3846G>A];[3739G>C] p.[Trp1282*];[Gly1247Arg]	F	78 ± 20	—	Neonatal screening	CF-PS
12[e]	3849+10kbC>T/G1244R	c.[3717+12191C>T];[3730G>A]	M	54 ± 1	Too young	Symptoms	CF-PS
13	Unknown/[1249-8A>G;G576A;R668C]	c.[?];[1117-8A>G;1727G>C;2002C>T]	F	72 ± 4	—	Symptoms	CF-PS
14	F508del/E56G	c.[1521_1523delCTT];[167A>G] p.[Phe508del];[Glu56Gly]	M	48 ± 2	OA	Symptoms	CBAVD

OA, Obstructive azoospermia; —, not applicable because female; nd, not determined.
[a]In this column, the new mutations found are reported as the second allele.
[b]Each sweat test value is the average of repeated sweat test measurements (from 2 to 4) upon enrollment and during follow-up.
[c]Patients number 6 and 7 are siblings.
[d]Patients number 9 and 10 are siblings.
[e]This patient has already been partially described (see text for quotation) and is included here to show the follow-up to 14 years.

Table 3. Genetic, biochemical, microbiological and clinical characterization of the patients with the 11 novel mutations found: clinical, pulmonary and microbiological characteristics of patients, upon enrollment and at follow-up.

| | | Upon enrollment (without therapy) | | | | At follow-up (with therapy whenever necessary) | | | | |
| | | | Respiratory manifestations | | Pulmonary bacterial isolates | | | Respiratory manifestations | | Pulmonary bacterial isolates |
Patient	Age	Clinical symptoms	By FEV1	By Rx		Age	Clinical symptoms	By FEV1	By Rx	
1	1.5 years	Diarrhea	Too young	Mild	Absent	37 years	Bronchopneumonias	Mild	Severe	*P.aeruginosa* (CC), *S. aureus* (CC)
2	2 months	Absent	Too young	Absent	Absent	5 years	Absent	Too young	Severe	Absent
3	32 years	Bronchopneumonia	Severe	Severe	*P.aeruginosa*, atypical mycobacteria	39 years	Bronchopneumonias	Severe	Severe	*P.aeruginosa* (CC), atypical mycobacteria (CC)
4	3 months	Meconium ileus	Too young	Absent	Absent	33 years[a]	Liver disease, cholelithiasis	Severe	Severe	*P.aeruginosa* (CC)
5	At birth	Meconium ileus	Too young	Absent	Absent	3 years	Cholelithiasis	Too young	Mild	*P.aeruginosa* (IC), *S. aureus* (IC)
6[b]	33 years	Bronchopneumonias, ABPA	Absent	Moderate	*S. aureus*	48 years	Bronchopneumonias, ABPA	Absent	Moderate	*P.aeruginosa* (IC), *S. aureus* (CC)
7[b]	32 years	Dehydratation, bronchopneumonia, bronchitis, pharyngitis	Absent	Absent	Absent	41 years	Absent	Absent	Absent	*P.aeruginosa* (IC), *S. aureus* (IC)
8	33 years	Pancreatitis, cholelithiasis	Absent	Mild	Absent	45 years	Absent (after cholecystectomy)	Absent	Mild	Absent
9[c]	2 months	Absent	too young	Absent	Absent	11 years	Absent	Absent	Mild	*S. aureus* (IC)
10[c]	2 months	Absent	Too young	Mild	*P.aeruginosa*, *S. aureus*	5 years	Pancreatitis, liver disease	Too young	Moderate	*S. aureus* (IC)
11	6 months	Absent	Too young	Absent	Absent	21 years	Rhinosinusitis, nasal polyposis	Mild	Moderate	*P.aeruginosa* (IC), *S. aureus* (CC)
12[d]	14 months	Bronchopneumonia	Too young	Mild	Absent	14 years	Recurrent pancreatitis, nasal polyposis	Absent	Moderate	*P.aeruginosa* (CC), *S. aureus* (CC)
13	7 years	Dehydration	Absent	Mild	Absent	20 years	Recurrent dehydration	Absent	Moderate	Absent
14	33 years	Absent	Absent	Absent	Absent	35 years	Absent	Absent	Absent	Absent

CC, chronic colonization; IC, intermittent colonization; ABPA, allergic bronchopulmonary aspergillosis. Classification of pulmonary symptoms by FEV1 is as follows: absent: >90%, mild: from 70 to 90%, moderate: from 40 to 70%, severe: <40%. Classification of pulmonary symptoms by chest X-ray is as follows: absent: no radiological signs; mild: limited air trapping or peribronchial infiltration; moderate: dense areas or bronchiectasis restricted to one lobe; severe: dense areas or bronchiectasis in both hemithoraxes.
[a]Age at death.
[b]Patients number 6 and 7 are siblings.
[c]Patients number 9 and 10 are siblings.
[d]This patient has already been partially described (see text for quotation) and is included here to show the follow-up to 14 years.

and displayed stronger pulmonary symptoms at the diagnosis than the brother, with bacterial isolates. She is now 48 years old and has worsened pulmonary symptoms with chronic bacterial colonization. The unrelated male patient was diagnosed at 33 years of age. He displayed pulmonary symptoms, pancreatitis and cholelithiasis with no bacterial isolates. He is now 45 years old and exhibits the same pulmonary conditions as those present at enrollment. He no longer has pancreatitis and underwent a cholecystectomy.

The Q779X (p.Gln779*) mutation was found in a CF-PS brother and sister with a [(TG)11T5; V562I; A1006E]/Q779X (c.[1210-14TG[11];1210-12T[5];1684G>A; 3017C>A];[2335C>T]) genotype. Their average sweat tests were, respectively, 70 ± 15 and 62 ± 17 mEq/L (both very variable). The enrollment of the male sibling was on the basis of neonatal screening (26,27) at 2 months of age, with no clinical symptoms and no pulmonary bacterial isolates. The enrollment of the female sibling was because of familiarity and neonatal screening at 2 months age, with pulmonary symptoms and bacterial isolates already present. At follow-up, which was respectively up to 11 and 5 years, both patients showed intermittent bacterial colonization. Pulmonary symptoms appeared in the male sibling. The female sibling exhibited worsened pulmonary symptoms as well as pancreatitis and liver disease.

The G1247R(G>C) (p.Gly1247Arg) mutation was found in a CF-PS female pa-

tient with a W1282X/G1247R(G>C) (p.[Trp1282*];[Gly1247Arg]) genotype. The average sweat test was 78 ± 20 mEq/L (very variable). The diagnosis was made at 6 months of age on the basis of neonatal screening (26,27), with no other symptoms. The patient is now 21 years old and displays pulmonary symptoms with chronic bacterial colonization, as well as rhinosinusitis and nasal polyposis.

The G1244R (p.Gly1244Arg) mutation was already published by us (24) when the patient was 7 years old; here we provide a further 7-year follow-up report following that description. The G1244R (p.Gly1244Arg) mutation was found in a CF-PS male patient, diagnosed at 14 months of age on the basis of symptoms, with a 3849+10kbC>T/G1244R

(c.[3717+12191C>T];[3730G>A]) genotype. The average sweat test was 54 ± 1 mEq/L. Respiratory manifestations were already present at diagnosis, although with no pulmonary bacterial isolates. The patient is now 14 years old, his respiratory symptoms worsened and he displays chronic bacterial colonization. The patient has also suffered from recurrent pancreatitis and nasal polyposis.

The 1249-8A>G (c.1117-8A>G) mutation was found in a novel complex allele [1249-8A>G;G576A;R668C (c.[1117-8A>G;1727G>C;2002C>T]) in a CF-PS female patient with an unknown mutation (after all mutational search steps, including the DEL step) on the other allele. The average sweat test was 72 ± 4 mEq/L. Upon enrollment, the patient was 7 years old, displayed pulmonary symptoms with no bacterial isolates and had already had a dehydration event. At the 20-year follow-up, she exhibited worsened pulmonary manifestations, although without bacterial isolates, and recurrent dehydration events.

The E56G (p.Glu56Gly) mutation was found in a CBAVD male subject with a F508del/E56G (p.[Phe508del];[Glu56Gly]) genotype. The average sweat test was 48 ± 2 mEq/L. At both the diagnosis (33 years) and follow-up (35 years), no symptoms other than CBAVD were present.

Phenotypic Description of the 10 Complex Alleles

Although the protocol used for the mutational search was not specifically aimed at the selection of complex alleles, 10 such alleles were found. The following five complex alleles encompassed mutations found only within respective complex alleles (never found separately).

The [E479X;V754M] (p.[Glu479*; Val754Met]) novel complex allele was found once in a CF-PI patient with a F508del (p.Phe508del) mutation on the other allele and an average sweat test of 106 ± 13 mEq/L. The E479X (p.Glu479*) is a novel mutation (described above and in Tables 1–3) found exclusively in this novel complex allele.

The [L24F;296+2T>G] (c.[72G>C; 164+2T>G]) complex allele was found once in a CF-PS patient with the (TG)13T5 (c.[1210-14TG[13];1210-12T[5]]) variant tract on the other allele and an average sweat test of 68 ± 11 mEq/L.

The [M348K;S912X] (p.[Met348Lys; Ser912*]) complex allele was found in 2 patients (1 CF-PI and 1 CBAVD). The CF-PI patient had an F508del (p.Phe508del) mutation on the other allele, whereas no mutation was detected on the other allele in the CBAVD patient. Sweat test values were 15 ± 3 mEq/L for the CBAVD patient and 107 ± 12 mEq/L for the CF-PI patient.

The [S466X(TAG);R1070Q] (p.[Ser466*; Arg1070Gln]) complex allele was found in 3 patients (2 CF-PI and 1 CF-PS). These patients had the following mutations on the other allele: F508del (p.Phe508del) (1 CF-PI), G542X (p.Gly542*) (1 CF-PI), 2789+5G>A (c.2657+5G>A) (1 CF-PS). Sweat test values ranged from 78 ± 3 to 79 ± 11 mEq/L.

The [R74W;V201M;D1270N] (p.[Arg74Trp; Val201Met;Asp1270Asn]) complex allele was found in 2 patients (1 CF-PS and 1 CBAVD). These patients had the following mutations on the other allele: S1206X (p.Ser1206*) (1 CF-PS), D1152H (p.Asp1152His) (1 CBAVD). Sweat test values were 46 ± 2 mEq/L for the CF-PS patient and 31 ± 1 mEq/L for the CBAVD patient.

As the mutations described above were only found in *cis* in the five complex alleles, it was impossible to evaluate the specific clinical effects due to the presence of more than one mutation on the same allele. By contrast, at least one of the mutations found in *cis* for the other five complex alleles was also found separately from the complex allele, thereby allowing the following speculation about their *cis*-acting effect.

The [(TG)11T5;V562I;A1006E] (c.[1210-14TG[11];1210-12T[5];1684G>A;3017C>A]) complex allele was found in 11 patients (9 CF-PS, 1 CFTR-RD and 1 CBAVD). These patients had the following mutations on the other allele: F508del (p.Phe508del) (3 CF-PS and 1 CFTR-RD), W1282X (p.Trp1282*) (2 CF-PS), Q779X (p.Gln779*) (2 CF-PS siblings), D110H (p.Asp110His)

(1 CF-PS), D614G (p.Asp614Gly) (1 CF-PS), unknown (1 CBAVD). Sweat test values ranged from 17 ± 3 to 76 ± 11 mEq/L, with an average value of 61 ± 20 mEq/L. The V562I (p.Val562Ile) and the A1006E (p.Ala1006Glu) were only found within the complex allele. By contrast, the (TG)11T5 (c.[1210-14TG[11];1210-12T[5]]), with no other mutations in *cis*, was found in 1 CFTR-RD patient and 1 CBAVD patient with, respectively, the 3849+10kbC>T (c.3717+12191C>T) and the D110H (p.Asp110His) mutations on the other allele. The sweat tests were 60 ± 7 mEq/L for the CFTR-RD patient and 23 ± 3 mEq/L for the CBAVD patient, with an average value of 42 ± 26 mEq/L. The CF-PS was only found in the group of patients with the complex allele, who also exhibited higher average sweat test values than patients without the complex allele, although with no statistical significance due to variability (Student *t* test, *p* = 0.26). This highlights the effect of the three mutations in *cis*.

The [R117L;L997F] (p.[Arg117Leu; Leu997Phe]) complex allele (28) was found in 6 patients (1 CF-PI and 5 CF-PS). These patients had the following mutations on the other allele: F508del (p.Phe508del) (1 CF-PI), G85E (p.Gly85Glu) (1 CF-PS), R334W (p.Arg334Trp) (2 CF-PS siblings) and W1282X (p.Trp1282*) (2 CF-PS). Sweat test values ranged from 70 ± 2 to 102 ± 4 mEq/L, with an average value of 84 ± 13 mEq/L. The R117L (p.Arg117Leu) was only found in the complex allele. The L997F (p.Leu997Phe), with no R117L (p.Arg117Leu) in *cis*, was found in 13 patients (2 CF-PS, 8 CFTR-RD and 3 CBAVD). These patients had the following mutations on the other allele: F508del (p.Phe508del) (1 CF-PS, 4 CFTR-RD and 1 CBAVD, including 2 siblings), G85E (p.Gly85Glu) (1 CF-PS), W1282X (p.Trp1282*) (2 CFTR-RD siblings), L320V (p.Leu320Val) (1 CFTR-RD), S549R(A>C) (p.Ser549Arg) (1 CFTR-RD), 711+5G>A (c.579+5G>A) (1 CBAVD) and unknown (1 CBAVD). Sweat test values ranged from 15 ± 2 to 77 ± 5 mEq/L, with an average value of 32 ± 18 mEq/L. Part of this case series has been described previously (28). Here we confirm that patients

with the complex allele had more severe diagnoses and significantly (Student *t* test, *p* < 0.0001) higher average sweat test values than the patients without the complex allele.

The [1249-8A>G;G576A;R668C] (c.[1117-8A>G;1727G>C;2002C>T]) complex allele was found once in a CF-PS patient with an unknown mutation on the other allele and an average sweat test of 72 ± 4 mEq/L. The 1249-8A>G (c.1117-8A>G) is a novel mutation (described above and in Tables 1–3) found exclusively in this novel complex allele. The complex allele [G576A;R668C] (p.[Gly576Ala;Arg668Cys]) (without the first mutation in *cis*) was found in two CFTR-RD patients with F508del (p.Phe508del) and S1235R (p.Ser1235Arg) on the other allele, with sweat tests, respectively, of 19 ± 2 and 17 ± 1 mEq/L (average 18 ± 1 mEq/L). These three mutations were never found alone. The complex allele with the three mutations in *cis* in the CF-PS patient was found in a more severe form of CF than the complex allele with only two mutations in *cis*. Accordingly, the average sweat test value was significantly (Student *t* test, *p* < 0.05) higher in the complex allele with three mutations than in the other allele.

The [359insT;(TG)12T5] (c.[227_228insT; 1210-14TG[12];1210-12T[5]]) complex allele was found once in a CFTR-RD patient with a (TG)12T5 (c.[1210-14TG[12];1210-12T[5]]) variant tract on the other allele and an average sweat test of 46 ± 5 mEq/L. The 359insT (p.Trp79LeufsX32) mutation was found only within the complex allele, whereas the (TG)12T5 (c.[1210-14TG[12]; 1210-12T[5]]) mutation was also found alone, with no other mutations in *cis*, in 41 patients (9 CF-PS, 16 CFTR-RD and 16 CBAVD). These patients had the following mutations on the other allele: F508del (p.Phe508del) (3 CF-PS, 5 CFTR-RD and 10 CBAVD), N1303K (p.Asn1303Lys) (1 CF-PS, 3 CFTR-RD and 1 CBAVD), 1717-1G>A (c.1585-1G>A) (3 CF-PS and 1 CFTR-RD), W1282X (p.Trp1282*) (3 CFTR-RD), G542X (p.Gly542*) (1 CF-PS, 1 CFTR-RD and 1 CBAVD), Y849X (p.Tyr849*) (1 CFTR-RD), 3849+10kbC>T (c.3717+12191C>T) (1 CFTR-RD), R1162X

(p.Arg1162*) (1 CBAVD), S549R(A>C) (p.Ser549Arg) (1 CFTR-RD) and unknown (1 CF-PS and 3 CBAVD). Sweat test values ranged from 11 ± 2 to 104 ± 43 mEq/L, with an average value of 47 ± 22 mEq/L (highly variable). In this case, the overlapping clinical presentations and average sweat test values between the patient with the complex allele and the other patients may depend on the varying effect of the (TG)12T5 (c.[1210-14TG[12];1210-12T[5]]) tract, which is also on the other allele in the patient with the complex allele and lowers the overall phenotypic effect.

Clinical Classification of Mutated Alleles and Genotypes

The three rules described in Materials and Methods led to the classification of 109 alleles (Table 4); 16 alleles could not, despite being identified as disease-causing, be univocally assigned to one (or more) macrocategories on the basis of our experimental data. It was also possible to assign 87 of the 109 classified alleles to a single macrocategory (56 to CF-PI, 15 to CF-PS, 14 to CFTR-RD and 2 to CBAVD). The remaining 22 alleles were classified as causing variable phenotypes (11 CF-PI and CF-PS; 4 CF-PS and CFTR-RD; 2 CFTR-RD and CBAVD; 2 CF-PI, CF-PS and CFTR-RD; 3 CF-PS, CFTR-RD and CBAVD). No allele was classified as causing all four phenotypes, nor was any allele found to cause very different phenotypes (for example, CF-PI and CBAVD). According to the principle that the prevalent clinical effect depends on the allele with the highest residual functionality, the adherence of our model regarding the clinical effect of the combination of classified alleles was not only verified on experimentally available allele combinations but also inferred from allele combinations that are not experimentally available (Supplementary Table S5).

Relationship between Genotype, Residual Functionality and Clinical Presentation

The sweat test is an *in vivo* measurement of *CFTR* residual functionality. A general significant correlation between

the sweat test and clinical manifestations emerged (ANOVA, *p* < 0.001; Figure 3A). The average values of the sweat test were 87 ± 19 mEq/L for CF-PI, 73 ± 22 mEq/L for CF-PS, 47 ± 24 mEq/L for CFTR-RD and 27 ± 13 mEq/L for CBAVD. However, a considerable degree of biological variability within each population and a wide overlap of values between different populations were observed. Consequently, a genotype-specific analysis of sweat test values in the 59 nonindividual genotypes (found in at least two unrelated individuals) was performed (Figures 3B, C). The 19 genotypes found in at least two different populations yielded significantly different overall sweat test values (Figure 3B; ANOVA, *p* < 0.0001). However, Bonferroni multiple comparison test revealed that this overall significant difference is due to only 13 pairs of genotypes out of a total of 171 possible comparisons. In addition, a marked interindividual biological variability emerged for the four genotypes found in three different populations (Figure 3B, the four leftmost genotypes) as well as for the 15 genotypes found in two different populations (Figure 3B, the 15 rightmost genotypes). The general effect observed is that the same genotype can give rise to a wide range of sweat test values. Furthermore, no evident correlation was detected between the different sweat test values (obtained from the same genotype in these 19 nonindividual genotypes from different populations) and the severity of the clinical presentation. Moreover, even the 40 nonindividual genotypes found only in one population yielded significantly different overall sweat test values (Figure 3C; ANOVA, *p* < 0.0002). However, Bonferroni multiple comparison test did not detect any statistically significant difference when the 780 possible comparisons between each genotype pair were performed. To sum up, it is clear that highly similar sweat test values may be observed in different populations.

DISCUSSION

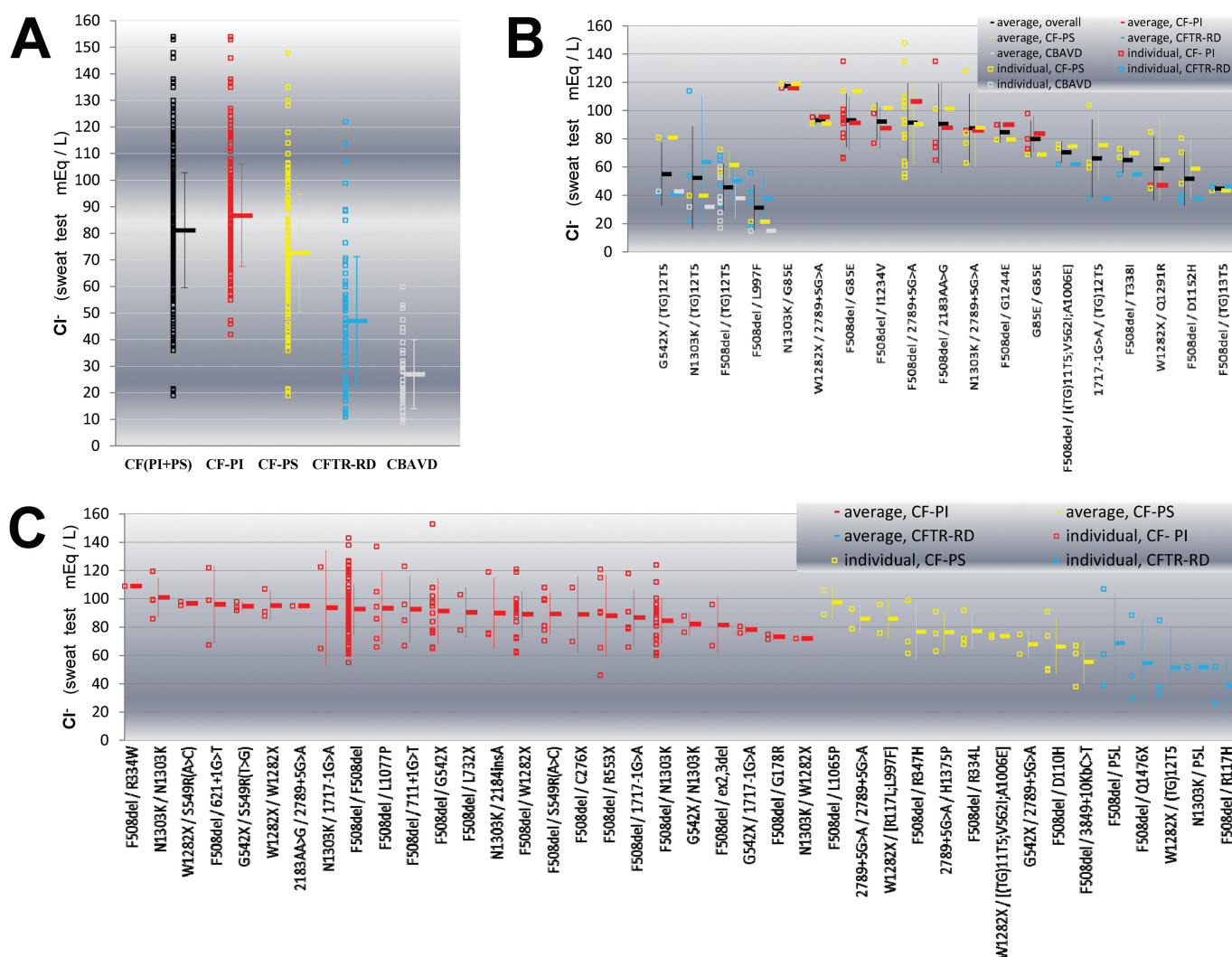A high degree of heterogeneity was observed between the mutational patterns of

Figure 3. Relationship between clinical presentation and sweat test. The empty squares represent the individual average sweat test values (average of at least two measurements); the filled rectangles represent the average sweat test value of each population: red for CF-PI, yellow for CF-PS, blue for CFTR-RD and gray for CBAVD. The bars represent the standard deviation (SD). (A) All genotypes (black empty squares and black filled rectangle for CF (PI + PS)). (B) Nonindividual genotypes found in at least two different populations (black filled rectangles for average of each genotype). (C) Nonindividual genotypes found only in one population. See Supplementary Table S4 for genotype HGVS name. See text for further explanations.

the clinical macrocategories analyzed. The mutational patterns appeared to be specific to each population, with only 34.4% of mutated alleles shared by at least two populations and 65.6% of population-specific mutated alleles (Figure 1, Supplementary Table S3). This specificity may be quantified as 56.5% of mutated alleles that were exclusive to CF-PI, 35.0% to CF-PS, 43.2% to CFTR-RD and 25.0% to CBAVD. This heterogeneity and mutational pattern specificity was enhanced at the genotypic

level, with only 8.4% of genotypes shared by at least two populations and 91.6% of population-specific mutated genotypes (Figure 2, Supplementary Table S4). This genotype specificity may be quantified as 91.3% of genotypes that were exclusive to CF-PI, 75.3% to CF-PS, 79.5% to CFTR-RD and 66.7% to CBAVD. In addition, alleles and genotypes found in different populations also displayed well-differentiated frequencies that were specific to each population. Most of the 125 different mutated

alleles identified were individual, found in only one patient or in siblings from a single family (48.0%), or were rare (38.4%), with a prevalence <0.008. Among the 225 genotypes, 73.8% were individual genotypes found in single patients or in siblings from a single family, and 19.1% were rare (frequency <0.008). It is noteworthy that 8.8% of the mutated alleles identified were novel (with an overall prevalence of 0.011), giving rise to 4.9% of the genotypes (with an overall prevalence

Table 4. Clinical classification of mutated alleles.

Allele legacy name	Allele HGVS name	Clinical classification	CFTR2
M1V	c.1A>G p.Met1Val	CF-PI	CF-causing
P5L	c.14C>T p.Pro5Leu	CF-PS, CFTR-RD	nd
ex2,3del	c.54-5940_273+10250del21080	CF-PI	CF-causing
ex2del	c.54-1161_164+1603del2875	CF-PI	nd
W19X(TAG)	c.56G>A p.Trp19*	CF-PI	nd
[L24F;296+2T>G]	c.[72G>C;164+2T>G]	uncertain: CF-PI and/or CF-PS	L24F nd; 296+2T>G nd
R31C	c.91C>T p.Arg31Cys	CFTR-RD	non CF-causing
S42F	c.125C>T p.Ser42Phe	uncertain: found only with an unknown allele in *trans*	nd
E56G	c.167G>A p.Glu56Lys	CBAVD	nd
[R74W;V201M;D1270N]	c.[220C>T;601G>A;3808G>A] p.[Arg74Trp;Val201Met;Asp1270Asn]	uncertain: CF-PS and/or CFTR-RD and/or CBAVD	R74W varying clinical consequence; V201M nd; D1270N varying clinical consequence
[359insT;(TG)12T5]	c.[227_228insT;1210-14TG[12];1210-12T[5]]	uncertain: CF-PI and/or CF-PS and/or CFTR-RD	359insT nd; T5 varying clinical consequence
G85E	c.254G>A p.Gly85Glu	CF-PI, CF-PS	CF-causing
D110H	c.328G>C p.Asp110His	CF-PS	CF-causing
R117C	c.349C>T p.Arg117Cys	CF-PS	CF-causing
R117H	c.350G>A p.Arg117His	CFTR-RD	varying clinical consequence
[R117L;L997F]	c.[350G>T;2991G>C] p.[Arg117Leu;Leu997Phe]	CF-PI, CF-PS	R117L nd; L997F non CF-causing
G126D	c.377G>A p.Gly126Asp	uncertain: CF-PI and/or CF-PS	nd
H139R	c.416A>G p.His139Arg	CF-PI, CF-PS	nd
574delA	c.442delA p.Ile148LeufsX5	CF-PI	CF-causing
621+1G>T	c.489+1G>T	CF-PI	CF-causing
621+3A>G	c.489+3A>G	CFTR-RD	nd
G178R	c.532G>A p.Gly178Arg	CF-PI	CF-causing
D192G	c.575A>G p.Asp192Gly	CF-PS	nd
E193K	c.577G>A p.Glu193Lys	CBAVD	nd
711+1G>T	c.579+1G>T	CF-PI	CF-causing
711+3A>G	c.579+3A>G	CF-PS	CF-causing
711+5G>A	c.579+5G>A	uncertain: CF-PI and/or CF-PS and/or CFTR-RD and/or CBAVD	CF-causing
H199R	c.596A>G p.His199Arg	CF-PI	nd
L206W	c.617T>G p.Leu206Trp	CFTR-RD	CF-causing
Q220X	c.658C>T p.Gln220*	CF-PI	CF-causing
852del22	c.720_741delAGGGAGAATGATGATGAAGTAC p.Gly241GlufsX13	CF-PI	CF-causing
907delCins29	c.775delCinsTCTTCCTCAGATTCATTGTGATTACCTCA	uncertain: CF-PI and/or CF-PS	nd
C276X	c.828C>A p.Cys276*	CF-PI	CF-causing

Table 4. Continued.

991del5	c.859_863delAACTT p.Asn287LysfsX19	CF-PI	nd
L320V	c.958T>G p.Leu320Val	uncertain: CF-PI and/or CF-PS and/or CFTR-RD	nd
R334W	c.1000C>T p.Arg334Trp	CF-PI, CF-PS	CF-causing
R334L	c.1001G>T p.Arg334Leu	CF-PS	nd
T338I	c.1013C>T p.Thr338Ile	CF-PS, CFTR-RD, CBAVD	CF-causing
R347P	c.1040G>C p.Arg347Pro	CF-PI, CF-PS	CF-causing
R347H	c.1040G>A p.Arg347His	CF-PS	CF-causing
[M348K;S912X]	c.[1043T>A;2735C>A] p.[Met348Lys;Ser912*]	CF-PI	M348K nd; S912X CF-causing
[1249-8A>G;G576A;R668C]	c.[1117-8A>G;1727G>C;2002C>T]	uncertain: found only with an unknown allele in *trans*	1249-8A>G nd; G576A non CF-causing; R668C non CF-causing
1259insA	c.1127_1128insA p.Gln378AlafsX4	CF-PI	CF-causing
E379X	c.1135G>T p.Glu379*	CF-PI	nd
M394R	c.1181T>G p.Met394Arg	CF-PI	nd
(TG)11T5	c.[1210-14TG[11];1210-12T[5]]	CFTR-RD, CBAVD	T5 varying clinical consequence
(TG)12T5	c.[1210-14TG[12];1210-12T[5]]	CF-PS, CFTR-RD, CBAVD	T5 varying clinical consequence
(TG)13T5	c.[1210-14TG[13];1210-12T[5]]	CF-PS, CFTR-RD	T5 varying clinical consequence
[(TG)11T5;V562I;A1006E]	c.[1210-14TG[11];1210-12T[5];1684G>A;3017C>A]	CF-PS, CFTR-RD	T5 varying clinical consequence; V562I nd; A1006E nd
K442X	c.1324A>T p.Lys442*	CF-PI	nd
T465N	c.1394C>A p.Thr465Asn	CF-PI	nd
[S466X(TGA);R1070Q]	c.[1397C>G;3209G>A] p.[Ser466*;Arg1070Gln]	CF-PI	S466X(TGA) CF-causing; R1070Q varying clinical consequence
[E479X;V754M]	c.[1435G>T;2260G>A] p.[Glu479*;Val754Met]	CF-PI	E479X nd; V754M non CF-causing
F508del	c.1521_1523delCTT p.Phe508del	CF-PI	CF-causing
1717-8G>A	c.1585-8G>A	CF-PI	CF-causing
1717-1G>A	c.1585-1G>A	CF-PI	CF-causing
D529N	c.1585G>A p.Asp529Asn	CF-PI	nd
G542X	c.1624G>T p.Gly542*	CF-PI	CF-causing
S549R(A>C)	c.1645A>C p.Ser549Arg	CF-PI	CF-causing
S549N	c.1646G>A p.Ser549Asn	CF-PI	CF-causing
S549R(T>G)	c.1647T>G p.Ser549Arg	CF-PI	CF-causing
G551D	c.1652G>A p.Gly551Asp	CF-PI	CF-causing
Q552X	c.1654C>T p.Gln552*	CF-PI	CF-causing
R553X	c.1657C>T p.Arg553*	CF-PI	CF-causing
L558S	c.1673T>C p.Leu558Ser	CF-PI	unknown significance
Y569D	c.1705T>G p.Tyr569Asp	CFTR-RD, CBAVD	unknown significance

Table 4. Continued.

L571S	c.1712T>C p.Leu571Ser	CF-PI	nd
[G576A;R668C]	c.[1727G>C;2002C>T] p.[Gly576Ala;Arg668Cys]	CFTR-RD	G576A non CF-causing; R668C non-CF causing
D579G	c.1736A>G p.Asp579Gly	CF-PS	varying clinical consequence
E585X	c.1753G>T p.Glu585*	CF-PI	CF-causing
H609L	c.1826A>T p.His609Leu	CFTR-RD	nd
A613T	c.1837G>A p.Ala613Thr	CF-PS	nd
D614G	c.1841A>G p.Asp614Gly	CF-PS	unknown significance
2143delT	c.2012delT p.Leu671*	CF-PS	CF-causing
2183AA>G	c.2051_2052delAAinsG p.Lys684SerfsX38	CF-PI, CF-PS	CF-causing
2184insA	c.2052_2053insA p.Gln685ThrfsX4	CF-PI	CF-causing
R709X	c.2125C>T p.Arg709*	CF-PI	CF-causing
L732X	c.2195T>G p.Leu732*	CF-PI	CF-causing
R764X	c.2290C>T p.Arg764*	CF-PI	CF-causing
Q779X	c.2335C>T p.Gln779*	uncertain: CF-PI and/or CF-PS	nd
E831X	c.2491G>T p.Glu831*	CF-PS	CF-causing
Y849X	c.2547C>A p.Tyr849*	CF-PI	CF-causing
ex14b-17bdel	c.2620–674_3367+198del9858	CF-PI	nd
2789+5G>A	c.2657+5G>A	CF-PI, CF-PS	CF-causing
2790-2A>G	c.2658-2A>G	CF-PS	nd
S912L	c.2735C>T p.Ser912Leu	uncertain: found only with an unknown allele in *trans*	nd
S945L	c.2834C>T p.Ser945Leu	CF-PS	CF-causing
S977F	c.2930C>T p.Ser977Phe	CFTR-RD	varying clinical consequence
L997F	c.2991G>C p.Leu997Phe	CF-PS, CFTR-RD, CBAVD	non CF-causing
ex17a-18del	c.2988+1173_3468+2111del8600	CF-PI	nd
P1013L	c.3038C>T p.Pro1013Leu	CFTR-RD	nd
Y1032C	c.3095A>G p.Tyr1032Cys	CFTR-RD	nd
3272-26A>G	c.3140-26A>G	CF-PS	CF-causing
L1065P	c.3194T>C p.Leu1065Pro	CF-PI, CF-PS	CF-causing
L1065R	c.3194T>G p.Leu1065Arg	uncertain: CF-PI and/or CF-PS	nd
R1066C	c.3196C>T p.Arg1066Cys	CF-PI	CF-causing
R1066H	c.3197G>A p.Arg1066His	CF-PI	CF-causing
G1069R	c.3205G>A p.Gly1069Arg	uncertain: found only with an unknown allele in *trans*	varying clinical consequence

Continued on next page

Table 4. *Continued.*

L1077P	c.3230T>C p.Leu1077Pro	CF-PI	CF-causing
Y1092X(C>A)	c.3276C>A p.Tyr1092*	CF-PI	CF-causing
M1137V	c.3409A>G p.Met1137Val	CFTR-RD	nd
D1152H	c.3454G>C p.Asp1152His	CF-PI, CF-PS, CFTR-RD	varying clinical consequence
R1162X	c.3484C>T p.Arg1162*	CF-PI	CF-causing
D1168G	c.3503A>G p.Asp1168Gly	CFTR-RD	nd
3667ins4	c.3535_3536insTCAA p.Thr1179IlefsX17	CF-PI	CF-causing
S1206X	c.3617C>A p.Ser1206*	uncertain: CF-PI and/or CF-PS	nd
I1234V	c.3700A>G p.Ile1234Val	CF-PI, CF-PS	CF-causing
S1235R	c.3705T>G p.Ser1235Arg	CFTR-RD	non CF-causing
3849+10kbC>T	c.3717+12191C>T	CF-PI, CF-PS	CF-causing
V1240G	c.3719T>G p.Val1240Gly	CFTR-RD	nd
G1244R	c.3730G>A p.Gly1244Arg	uncertain: CF-PI and/or CF-PS	nd
G1244E	c.3731G>A p.Gly1244Glu	CF-PI, CF-PS	CF-causing
G1247R(G>C)	c.3739G>C p.Gly1247Arg	CF-PS	nd
W1282X	c.3846G>A p.Trp1282*	CF-PI	CF-causing
Q1291R	c.3872A>G p.Gln1291Arg	CF-PI, CF-PS, CFTR-RD	nd
4016insT	c.3884_3885insT p.Ser1297PhefsX5	CF-PI	CF-causing
4040delA	c.3908delA p.Asn1303ThrfsX25	CF-PI	nd
N1303K	c.3909C>G p.Asn1303Lys	CF-PI	CF-causing
ex22-24del	c.3964–3890_4443+3143del9454ins5	CF-PI	nd
ex22,23del	c.3964–78_4242+577del1532	CF-PI	CF-causing
4168delCTAAGCC	c.4036_4042del p.Leu1346MetfsX6	CF-PI	nd
G1349D	c.4046G>A p.Gly1349Asp	CF-PI	CF-causing
H1375P	c.4124A>C p.His1375Pro	uncertain: CF-PI and/or CF-PS	nd
S1455X	c.4364C>G p.Ser1455*	CF-PS, CFTR-RD	nd
Q1476X	c.4426C>T p.Gln1476*	CFTR-RD	nd

nd, Not determined. According to the three rules described (see Materials and Methods), each mutated allele was classified according to its clinical outcome. It was impossible to univocally assign 16 of the 125 different mutated alleles to one or more macrocategories. A comparison with the CFTR2 project (11) (http://www.cftr2.org) is shown. The alleles are ordered according to their nucleotidic position.

of 0.021). When taken together, these results revealed a peculiarity of the mutational pattern within each clinical macrocategory that appeared to be largely dependent on rare and individual muta-

tions and genotypes, as well as on the varying prevalence of common alleles and genotypes.

After the extended search in the CF-PI, CF-PS and CFTR-RD macrocategories, a

low frequency of unknown alleles (0.007, 0.033 and 0.035, respectively) and of patients with at least one unknown allele (0.014, 0.064 and 0.042, respectively) was left. This highlights the fact that *CFTR*

molecular lesions underlie the vast majority of these clinical forms. By contrast, after the extended search in CBAVD, 0.436 of the alleles and 0.551 of the patients remained uncharacterized. The role of *CFTR* in the reproductive apparatus (29–36) and the involvement of *CFTR* mutations in reduced male (37–41) and female (42–44) fertility are strongly debated issues. It is noteworthy that 31.9% of CBAVD patients had a genotype with two unknown alleles. We may conclude that the strictly mono-symptomatic form of CBAVD, which has no other CF manifestations, is frequently caused by molecular lesions other than those in the *CFTR*.

Although systematic studies have not yet been performed, over 40 complex alleles (with two or more mutations in *cis* on the same allele) of *CFTR* have so far been described. By using our approach, which was only partially aimed at the selection of complex alleles, we were able to identify 10 such alleles, two of which are novel. The fact that widely used protocols designed for mutational searches are usually interrupted after the first two mutations on different alleles have been found may greatly limit the interpretation of genetic data. The true functional significance of a sequence variation in *cis* with another variation (undetected and with some functional consequence) may be biased, as may the relationship between genotype and phenotype. For 4 out of 10 complex alleles found in our case series, an evaluation of the *cis*-acting effect was possible by comparing alleles with mutations in *cis* with those with the same mutations disjointed. The ability of complex alleles to give rise to more severe forms of CF and higher sweat test values was highlighted for three of these complex alleles. These data, together with those in the literature (see Lucarelli *et al.* [1]) for a review) indicate that complex alleles may account for a greater degree of variability than is usually acknowledged. To be meaningful, mutational search protocols should be designed to search for complex alleles, at least in cases in which the clinical presentation varies even when the genotype is apparently identical. At the

very least, they should be planned in such a way as to complete the characterization of known complex alleles when one of the mutations already known to be in *cis* is found.

Some unusual results on the clinical outcome of (TG)mTn tracts and of some stop mutations are described in the Supplementary Materials.

A crucial issue in CF is the assignment of a possible pathological role to sequence variations. The algorithm we applied in this work (described in Materials and Methods and in Supplementary Figure S1 with examples) allowed 87.2% of the mutated alleles identified (109 of 125) to be assigned to clinical macrocategories. These mutated alleles comprised 79.8% (87 of 109) that could be considered to cause restricted clinical manifestations (only one specific clinical macrocategory) and the remaining 20.2% (22 of 109) that could be considered to have a varying effect (more than one clinical macrocategory) (Table 4). Our approach left some uncertainty with regard to which clinical form(s) causes the remaining 12.8% of alleles (16 of 125), although their ability to induce disease is unequivocal. For a comment on the clinical classification of these 16 alleles, see the Supplementary Materials.

The best approach recently made to characterize *CFTR* mutations is the CFTR2 study (11) (http://www.cftr2.org). This and our approach have both common and distinctive features. The main common characteristic is that both use a phenotypic-driven approach. The main distinctive characteristics are that the CFTR2 is focused on the most common *CFTR* mutations worldwide and on classic forms of CF (with a positive sweat test), whereas our study also includes nonclassic CF forms (with also borderline sweat test) and rare mutations. Consequently, a greater mutational heterogeneity in the *CFTR* gene was observed in our study. A direct consequence is that the 43.2% of the alleles we identified (54 of 125, also taking into account complex alleles) were not included in the CFTR2 study. Another three alleles we classified

had been recognized in the CFTR2 study as being of unknown significance. If the two alleles with an uncertain significance in our study were also excluded, 66 alleles classified in both studies, and which could consequently be compared, were left. For these alleles, the level of agreement between the two characterization approaches was excellent, with 95.5% of them (63 of 66) being classified similarly. In particular, 54 alleles were classified as causing CF-PI and/or CF-PS in our study in perfect agreement with a classification as CF-causing in the CFTR2 study. Another four alleles were classified as belonging to two or more different macrocategories (including at least one of the nonclassic, namely CFTR-RD and/or CBAVD) in our study and with varying clinical consequences in the CFTR2 study, also in this case with an excellent match. A good level of agreement may also be recognized for mutations R117H (p.Arg117His) and S977F (p.Ser977Phe), classified in our study as CFTR-RD–causing, and the mutation D579G (p.Asp579Gly), classified in our study as CF-PS–causing, and both recognized in the CFTR2 study with varying clinical consequences, therefore also including our phenotypic findings. A similar good match may also be assumed for two mutations classified as non–CF-causing in CFTR2 [S1235R (p.Ser1235Arg) and R31C (p.Arg31Cys)] but as CFTR-RD–causing in our study, owing to the fact that CFTR2 is mainly aimed at classic CF and is more prone to a classification as non-causing for those mutations originating nonclassic clinical and biochemical phenotypes. The three actually discrepant alleles were L997F (p.Leu997Phe), without the R117L (p.Arg117Leu) in *cis*, L206W (p.Leu206Trp) and T338I (p.Thr338Ile). The L997F (p.Leu997Phe) allele can, according to the findings that emerge both from this work and previous studies (28), also give rise to CF-PS, whereas in the CFTR2 study, it was classified as non–CF-causing. The L206W (p.Leu206Trp), which in our study was classified as a CFTR-RD–causing mutation, was classified as CF-causing in the CFTR2 study.

The T338I (p.Thr338Ile) is classified in the CFTR2 study as CF-causing. From our study resulted the origination not only of CF-PS (in agreement with the CFTR2 study) but also of CFTR-RD and CBAVD. This result extends the phenotypic consequences of this mutation also in accord with previous findings (45,46). These discrepancies may be linked to the degree of variability that is independent of *CFTR* and that accounts for about 4.5% (3 of 66) of the mutations considered in both studies. We found 52 individual mutated alleles that, in combination with nonindividual alleles, originated 146 (64.9%) genotypes found only once in a single patient. The assignment of a pathological role to these individual alleles was quite easy. On the other hand, there were no other patients in our case series to confirm the assignment. However, for 19 (of 52) individual alleles, a comparison with the CFTR2 study was possible. In this case, for these individual alleles, we found a perfect adherence between our classification and that of the CFTR2 study. Overall, the excellent agreement on common, rare and individual alleles included in both studies lends further support to our conclusions regarding the other mutations not included in the CFTR2 study. Our classification method represents a good example of how it is possible to deduce the phenotypic consequence of a mutated allele on the basis of a well-defined clinical classification and a reasonable extended mutational analysis, even in the absence of experimental functional studies. Obviously, the limitations of every attempt of pathogenicity inference by a phenotypic-driven approach considering a limited number of cases should be taken into account. The final goal should be to achieve a large case series as a starting point for *CFTR* sequence variations experimental functional analysis. This approach would be more powerful but also more complex and time-consuming.

The approaches designed to assign a phenotypic outcome to *CFTR* sequence variations are generally based on an allele-oriented view. However, it is widely accepted that overall residual functionality depends on both alleles and, ultimately, on the allele with the higher residual functionality. To address this issue, we elaborated a two-allele combinatorial view of the clinical outcome that is to be expected, starting from previously classified alleles (Supplementary Table S5). This approach may be considered as a genotypic-oriented prediction tool of clinical outcome, in part experimentally validated in this work and in part inferred (to be experimentally verified when the specific genotypes are identified).

The molecular mechanisms underlying the variability between genotype and phenotype are as yet unclear. At least two steps may be involved. The first step may be defined as the transition from the *CFTR*-mutated genotype to CFTR protein residual function (genotype → residual functionality step). It is reasonable to assume that this transition is more likely to be influenced by intragenic (*CFTR*-dependent) variability. It may originate from the large number of sequence variations and be markedly enhanced by their combination in *trans* and in *cis* as complex alleles (1,28), as well as by a regulatory posttranscriptional and posttranslational impairment that often escapes recognition. The second step may be defined as the transition from the CFTR protein residual function to clinical phenotype (residual functionality → clinical step). This transition is more likely to be influenced by extragenic variability due to genes other than *CFTR*, such as the so-called modifier genes (9,47) and the *CFTR* interactome (48,49). Significant differences emerged in the average sweat test values, which may be considered an *in vivo* measurement of the CFTR protein residual function between the populations analyzed. This result highlights that at least a general correlation between residual functionality and clinical macrocategories exists. However, the marked variability within each population results in a considerable overlap between the sweat tests. This intra-population variability seems to arise from the presence of several different genotypes, each giving rise to its own range of residual functionality, although narrower than that observed in the overall populations. It may be argued that at least some of the variability arises in the genotype → residual functionality step. On the other hand, it is to be expected that the different functionalities that arise from the same genotype are correlated with different clinical manifestations (the higher the sweat test, the more severe the clinical manifestations). However, no such correlation was detected between different sweat test values obtained from the same genotype and the clinical presentation. Furthermore, similar sweat test values associated to different genotypes may be observed in different populations. It may also be argued that another part of the variability arises in the residual functionality → clinical step. The relative contribution of the two steps to the overall variability is still largely unknown and deserves further quantitative studies.

CONCLUSION

The full clinical and mutational characterization of CF patients reveals a genetic heterogeneity that underlies a strong correlation between genotypic patterns and phenotypic macrocategories. This specificity appears to largely depend on rare and individual mutations, as well as on the varying prevalence of common alleles in the populations analyzed. A pathogenic classification of sequence variations on the basis of rigorous clinical studies and an extended mutational search may be a rapid and meaningful way to initially characterize sequence variations for which an experimental functional characterization is still lacking, as well as a starting point for subsequent experimental quantitative functional characterizations. The experimental dissection of the overall biological *CFTR* pathway appears to be a powerful approach for a better comprehension of the sources of variability in the genotype–phenotype relationship. Overall, our findings call

REFERENCES

1. Lucarelli M, Pierandrei S, Bruno SM, Strom R. (2012) The Genetics of CFTR: Genotype – Phenotype Relationship, Diagnostic Challenge and Therapeutic Implications [Internet]. In: *Cystic Fibrosis - Renewed Hopes Through Research*. Sriramulu D (ed.) Intech, Rijeka, Croatia, [cited 2015 Jun 15], pp. 91–122. Available from: http://www.intechopen.com/books/cystic-fibrosis-renewed-hopes-through-research/the-genetics-of-cftr-genotype-phenotype-relationship-diagnostic-challenge-and-therapeutic-implicatio
2. Davies JC, Alton EW, Bush A. (2007) Cystic fibrosis. *BMJ.* 335:1255–9.
3. O'Sullivan BP, Freedman SD. (2009) Cystic fibrosis. *Lancet.* 373:1891–904.
4. Riordan JR. (2008) CFTR function and prospects for therapy. *Annu. Rev. Biochem.* 77:701–26.
5. Drumm ML, Ziady AG, Davis PB. (2012) Genetic variation and clinical heterogeneity in cystic fibrosis. *Annu. Rev. Pathol.* 7:267–82.
6. Paranjape SM, Zeitlin PL. (2008) Atypical cystic fibrosis and CFTR-related diseases. *Clin. Rev. Allergy Immunol.* 35:116–23.
7. Dequeker E, *et al.* (2009) Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders: updated European recommendations. *Eur. J. Hum. Genet.* 17:51–65.
8. Bombieri C, *et al.* (2011) Recommendations for the classification of diseases as CFTR-related disorders. *J. Cyst. Fibros.* 10 Suppl 2:S86–102.
9. Bombieri C, Seia M, Castellani C (2015) Genotypes and phenotypes in cystic fibrosis and cystic fibrosis transmembrane regulator-related disorders. *Semin. Respir. Crit. Care Med.* 36:180–93.
10. Cystic Fibrosis Mutation Database [Internet]. [updated 2011 Apr 25; cited 2015 June 22]. Available from: http://www.genet.sickkids.on.ca/cftr/
11. Sosnay PR, *et al.* (2013) Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. *Nat. Genet.* 45:1160–7.
12. Amaral MD, Kunzelmann K. (2007) Molecular targeting of CFTR as a therapeutic approach to cystic fibrosis. *Trends Pharmacol. Sci.* 28:334–41.
13. Becq F, Mall MA, Sheppard DN, Conese M, Zegarra-Moran O. (2011) Pharmacological therapy for cystic fibrosis: from bench to bedside. *J. Cyst. Fibros.* 10(Suppl 2):S129–45.
14. Rogan MP, Stoltz DA, Hornick DB. (2011) Cystic fibrosis transmembrane conductance regulator intracellular processing, trafficking, and opportunities for mutation-specific treatment. *Chest.* 139:1480–90.
15. Borowitz D, *et al.* (2009) Cystic Fibrosis Foundation evidence-based guidelines for management of infants with cystic fibrosis. *J. Pediatr.* 155:S73–93.
16. Canton R, *et al.* (2005) Antimicrobial therapy for pulmonary pathogenic colonisation and infection by Pseudomonas aeruginosa in cystic fibrosis patients. *Clin. Microbiol. Infect.* 11:690–703.
17. Farrell PM, *et al.* (2008) Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *J. Pediatr.* 153:S4–14.
18. Miller MB, Gilligan PH. (2003) Laboratory aspects of management of chronic pulmonary infections in patients with cystic fibrosis. *J. Clin. Microbiol.* 41:4009–15.
19. Zhou J, Garber E, Desai M, Saiman L. (2006) Compliance of clinical microbiology laboratories in the United States with current recommendations for processing respiratory tract specimens from patients with cystic fibrosis. *J. Clin. Microbiol.* 44:1547–9.
20. Castellani C, *et al.* (2009) European best practice guidelines for cystic fibrosis neonatal screening. *J. Cyst. Fibros.* 8:153–73.
21. Gibson LE, Cooke RE. (1959) A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. *Pediatrics.* 23:545–9.
22. Gullo L, *et al.* (1997) Faecal elastase 1 in children with cystic fibrosis. *Eur. J. Pediatr.* 156:770–2.
23. Lucarelli M, *et al.* (2002) Simultaneous cycle sequencing assessment of (TG)m and Tn tract length in CFTR gene. *Biotechniques.* 32:540–7.
24. Lucarelli M, *et al.* (2006) A 96-well formatted method for exon and exon/intron boundary full sequencing of the CFTR gene. *Anal. Biochem.* 353:226–35.
25. Ferraguti G, *et al.* (2011) A template for mutational data analysis of the CFTR gene. *Clin. Chem. Lab Med.* 49:1447–51.
26. Narzi L, *et al.* (2002) Comparison of two different protocols of neonatal screening for cystic fibrosis. *Clin. Genet.* 62:245–9.
27. Narzi L, *et al.* (2007) Does cystic fibrosis neonatal screening detect atypical CF forms? Extended genetic characterization and 4-year clinical follow-up. *Clin. Genet.* 72:39–46.
28. Lucarelli M, *et al.* (2010) A new complex allele of the CFTR gene partially explains the variable phenotype of the L997F mutation. *Genet. Med.* 12:548–55.
29. Teixeira S, *et al.* (2013) Immunohistochemical analysis of CFTR in normal and disrupted spermatogenesis. *Syst. Biol. Reprod. Med.* 59:53–9.
30. Dube E, Hermo L, Chan PT, Cyr DG. (2008) Al-

terations in gene expression in the caput epididymides of nonobstructive azoospermic men. *Biol. Reprod.* 78:342–51.

31. Patrizio P, Salameh WA. (1998) Expression of the cystic fibrosis transmembrane conductance regulator (CFTR) mRNA in normal and pathological adult human epididymis. *J. Reprod. Fertil. Suppl.* 53:261–70.

32. Xu WM, *et al.* (2011) Defective CFTR-dependent CREB activation results in impaired spermatogenesis and azoospermia. *PLoS One.* 6:e19120.

33. Chen MH, *et al.* (2010) Involvement of CFTR in oviductal. *Hum. Reprod.* 25:1744–54.

34. Muchekehu RW, Quinton PM. (2010) A new role for bicarbonate secretion in cervico-uterine mucus release. *J. Physiol.* 588:2329–42.

35. Hodges CA, Palmert MR, Drumm ML. (2008) Infertility in females with cystic fibrosis is multifactorial: evidence from mouse models. *Endocrinology.* 149:2790–7.

36. Trezise AE, *et al.* (1993) CFTR expression is regulated during both the cycle of the seminiferous epithelium and the oestrous cycle of rodents. *Nat. Genet.* 3:157–64.

37. Yu J, *et al.* (2011) Association of genetic variants in CFTR gene, IVS8 c.1210–12T[5_9] and c.1210–35_1210–12GT[8_12], with spermatogenetic failure: case-control study and meta-analysis. *Mol. Hum. Reprod.* 17:594–603.

38. Claustres M. (2005) Molecular pathology of the CFTR locus in male infertility. *Reprod. Biomed. Online.* 10:14–41.

39. Dohle GR, *et al.* (2002) Genetic risk factors in infertile men with severe oligozoospermia and azoospermia. *Hum. Reprod.* 17:13–6.

40. Elia J, *et al.* (2009) Human semen hyperviscosity: prevalence, pathogenesis and therapeutic aspects. *Asian J. Androl.* 11:609–15.

41. Rossi T, *et al.* (2004) High frequency of (TG)mTn variant tracts in the cystic fibrosis transmembrane conductance regulator gene in men with high semen viscosity. *Fertil. Steril.* 82:1316–22.

42. Ahmad A, Ahmed A, Patrizio P. (2013) Cystic fibrosis and fertility. *Curr. Opin. Obstet. Gynecol.* 25:167–72.

43. Gervais R, *et al.* (1996) Hypofertility with thick cervical mucus: another mild form of cystic fibrosis? *JAMA.* 276:1638.

44. Schoyer KD, Gilbert F, Rosenwaks Z. (2008) Infertility and abnormal cervical mucus in two sisters who are compound heterozygotes for the cystic fibrosis (CF) DeltaF508 and R117H/7T mutations. *Fertil. Steril.* 90:1201–22.

45. Leoni GB, *et al.* (1995) A specific cystic fibrosis mutation (T338I) associated with the phenotype of isolated hypotonic dehydration. *J. Pediatr.* 127:281–3.

46. Padoan R, Seia M, Cremonesi L, Giunta A. (1996) Mild CF phenotype associated with T338I missense mutation. *J. Pediatr.* 128:721–2.

47. Cutting GR. (2010) Modifier genes in Mendelian disorders: the example of cystic fibrosis. *Ann. N. Y. Acad. Sci.* 1214:57–69.

48. Gisler FM, von KT, Kraemer R, Schaller A, Gallati S. (2013) Identification of SNPs in the cystic fibrosis interactome influencing pulmonary progression in cystic fibrosis. *Eur. J. Hum. Genet.* 21:397–403.

49. Wang X, *et al.* (2006) Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. *Cell.* 127:803–15.