

Genotype-phenotype relationships in a cohort

D. Hubert*, T. Bienvenu**, N. Desmazes-Dufeu*, I. Fajac+,
J. Lacronique*, R. Matran+, J.C. Kaplan**, D.J. Dusser*

Genotype-phenotype relationships in a cohort of adult cystic fibrosis patients. D. Hubert, T. Bienvenu, N. Desmazes-Dufeu, I. Fajac, J. Lacronique, R. Matran, J.C. Kaplan, D.J. Dusser. ©ERS Journals Ltd 1996.

ABSTRACT: In cystic fibrosis (CF), relationships between genotype and phenotype have been shown for pancreatic status but not for pulmonary disease.

One hundred and ten adult CF patients were classified according to the expected effect of their mutations on cystic fibrosis transmembrane conductance regulator (CFTR) protein: Group 1 (n=48) included $\Delta F508$ homozygotes; Group 2 (n=26), patients with two "severe" mutations and no expected CFTR production; Group 3 (n=17), patients with expected partly functional CFTR corresponding to at least one "mild" mutation; Group 4 (n=19), patients with no mutation identified or only one identified "severe" mutation.

As compared to Groups 1 and 2: patients from Groups 3 and 4 had higher arterial oxygen tension (P_{a,O_2}) (9.5 ± 1.9 and 9.9 ± 1.5 vs 8.8 ± 1.5 and 8.3 ± 1.7 kPa, respectively $p<0.02$); and a slower decline in their pulmonary function, estimated by the mean annual loss in forced vital capacity (FVC) (1.2 ± 1.0 and 1.5 ± 1.1 vs 2.0 ± 0.9 and $2.2\pm 1.0\%$, respectively; $p<0.01$) and in forced expiratory volume in one second (FEV₁) (1.7 ± 1.1 and 1.9 ± 1.3 vs 2.6 ± 1.0 and $2.8\pm 1.0\%$, respectively; $p<0.005$). They had fewer episodes of colonization of the airways by *Pseudomonas aeruginosa*, and colonization occurred at a more advanced age (median age 25 and 19 vs 15 and 17 yrs, respectively; $p<0.01$) and required fewer intravenous antibiotic courses ($p<0.01$). Pancreatic insufficiency was less frequent in Groups 3 (23%) and 4 (63%) than in Groups 1 (100%) and 2 (96%).

This study suggests that the phenotype of adult cystic fibrosis patients, including the severity of the lung disease, is related to the severity of the cystic fibrosis transmembrane conductance regulator mutations.

Eur Respir J., 1996, 9, 2207–2214.

*Service de Pneumologie, **Service de Biochimie-Génétique, and +Service de Physiologie-Explorations Fonctionnelles, UFR Cochin-Port-Royal, Université René Descartes, Paris, France.

Correspondence: D. Dusser
Service de Pneumologie
Hôpital Cochin
27 rue du Fg Saint Jacques
75679 Paris cedex 14
France

Keywords: Cystic fibrosis
cystic fibrosis transmembrane conductance regulator
genotype-phenotype
pulmonary function

Received: March 12 1996
Accepted after revision July 8 1996

Since the discovery of the cystic fibrosis (CF) gene in 1989 [1, 2], more than 550 different CF mutations have been described [3]. The association of two different mutations in compound heterozygotes with CF is responsible for a large heterogeneity of genotypes. Although the knowledge of different mutations of the cystic fibrosis transmembrane conductance regulator (CFTR) gene has progressed during recent years, the mechanisms accounting for the great variability of clinical expression of the disease are still poorly understood [4].

One possible explanation for the great variability of phenotypes of the disease is that the severity of CF might be related to the level of deficiency in the CFTR protein, which may vary depending on the type of mutation of the CFTR gene. Other possible explanations include influence of other genes on the expression of the CF gene and environmental factors.

The first studies of genotype-phenotype relationships in CF focused on the most frequent mutation, the $\Delta F508$ mutation, which is found in about 70% of CF

chromosomes in Western countries. KEREM *et al.* [5] showed that $\Delta F508$ homozygotes almost always have pancreatic insufficiency, whereas 28% of $\Delta F508$ compound heterozygotes and 64% of patients who do not carry a $\Delta F508$ mutation have adequate pancreatic function. In contrast, no correlation was found between the $\Delta F508$ genotype and pulmonary status [5–11]. In the majority of studies [5, 6, 8, 9, 11], the lack of correlation between genotype and pulmonary status might, in part, be explained by selection of young patients (mean age ranging 9–18 yrs), which may give insufficient background to assess the long-term outcome of pulmonary involvement. In two studies which focused on CF adult patients only [7, 10], and in most other studies [5, 6, 8, 9], comparison between homozygotes or compound heterozygotes for the $\Delta F508$ mutation were made without differentiating compound heterozygotes according to the type of their non- $\Delta F508$ mutation, which may influence the severity of the pulmonary outcome. KUBESCH *et al.* [12], who classified $\Delta F508$ compound heterozygotes

according to the type of non- $\Delta F508$ mutation, showed genetic determinants of the colonization of airways by *Pseudomonas aeruginosa*, but not for pulmonary function. Concerning the relationship between the severity of pulmonary involvement and other mutations than $\Delta F508$, a mild pulmonary disease was described in two sisters homozygous for the G551S mutation [13], and in compound heterozygotes for the missense mutation A455E, a mutation commonly found in The Netherlands [14].

WELSH and SMITH [15] described four mechanisms by which mutations disrupt CFTR function: 1) defective protein production due to splicing mutations, frameshift or nonsense mutations (class I mutations); 2) defective protein processing (class II mutations, including the $\Delta F508$ mutation); 3) defective regulation, mainly due to missense mutations in the two nucleotide binding domains (class III mutations); and 4) defective conduction associated with missense mutations in the membrane-spanning domains (class IV mutations). This classification makes it possible to predict the likely effect of a known mutation on the CFTR function, although the effect of a given mutation on cell function is not yet known. Using this approach, we have classified a cohort of 110 CF adult patients into four groups according to the predicted severity of CFTR dysfunction, and we have examined whether the different groups of genotype had any influence on the phenotype of the disease, in particular for pulmonary and pancreatic involvement.

Study of the relationship between genotype and phenotype in a group of CF adults provides the opportunity to examine long-term outcome of the disease, in particular for pulmonary involvement, and to identify possible discrepancies in outcome between the groups of genotype as well as within each group of genotype. It also provides the opportunity to include some milder phenotypes of patients whose diagnosis is established later in life.

Patients and methods

Patients

One hundred and ten adult patients with cystic fibrosis were included in this study (70 males and 40 females; mean age 26 yrs). All patients regularly attended the adult CF centre at Cochin Hospital, Paris (France), between January 1, 1990, and March 31, 1995. All but one had bronchiectasis confirmed by computed tomographic (CT) scan. The only patient without bronchiectasis presented with congenital bilateral absence of the vas deferens, pansinusitis with nasal polyposis and a familial history of CF. Diagnosis was confirmed by two abnormal chloride concentrations (≥ 70 mmol·L⁻¹) in pilocarpine iontophoresis sweat tests and/or the identification of two mutations in the CFTR gene. Only three patients had a normal sweat test: two related patients, who were compound heterozygotes for the G542X and 3849+10 kb cytosine (C)→thymine (T) mutations, and one patient who was compound heterozygote for the R1070Q and D1152H mutations.

Clinical assessment

Data concerning clinical presentation included age, symptoms at diagnosis, and history of meconium ileus. Sweat chloride concentrations were recorded.

Pancreatic status was determined by the fat content in stool samples collected over 3 days: patients showing normal results (faecal fat less than 5 g·day⁻¹) and currently not treated by pancreatic enzyme replacement were defined as pancreatic sufficient, and the remaining as pancreatic insufficient. Body weight was expressed as a percentage of ideal weight for height.

Lung function tests were performed regularly in all patients using a spirometer Autospiro AS500, and the best values for forced vital capacity (FVC) and forced expiratory volume in one second (FEV₁) over the previous year were used, expressed as the percentage of predicted values (% pred). Because the study population comprised patients of different ages, the decline in their pulmonary function was evaluated in relation to age. Therefore, we estimated the percentage of mean annual loss in FVC and FEV₁, defined as (100 - % pred FVC)/age and (100 - % pred FEV₁)/age, respectively. Arterial oxygen tension (P_{a,O_2}) and arterial carbon dioxide tension (P_{a,CO_2}) were measured at rest.

Patients with *P. aeruginosa* growth for at least 6 months were considered to be colonized with this microorganism. The number of intravenous (*i.v.*) antibiotic courses per year and the number of days of *i.v.* antibiotics per year were recorded. Finally, information was obtained on history of pneumothorax, nasal polyps, hepatic cirrhosis, gallstones, diabetes mellitus, joint pains and pregnancy.

Deoxyribonucleic acid (DNA) analysis

Peripheral blood samples were collected from the CF patients, and genomic DNA was extracted by standard methods. The 27 exons of the CFTR gene and the surrounding intronic sequences were analysed by denaturing gradient gel electrophoresis, as described previously [16, 17]. Computer analysis was performed using Melt 87, generously provided by L. Lerman (Massachusetts Institute of Technology). Polymerase chain reaction (PCR) products that displayed an altered behaviour in the gel were subsequently sequenced. Direct sequencing of PCR products was carried out by the Sanger dideoxy-mediated chain-termination method with Sequenase version 2.0 (United States Biochemical (USB)) [18].

The intron mutations 3849+10 kb C→T and 1811+1.6 kb adenine (A)→guanine (G) were separately screened, according to the instructions of HIGHSMITH *et al.* [19] and CHILLON *et al.* [20].

Classification of mutations

Using the approach to the classification of CFTR mutations proposed by WELSH and SMITH [15], the patients were classified into four groups according to the probable effect of their mutations on CFTR function, regardless of clinical features. Group 1 included $\Delta F508$ homozygote patients, in which the $\Delta F508$ CFTR

protein is not correctly processed. Group 2 included patients with no CFTR production or expected non-functional CFTR, which corresponded either to the existence of two "severe" mutations (nonsense mutations, frameshift or splicing mutations) or to compound heterozygotes for one "severe" mutation and one ΔF508 mutation. The 1811+1.6 kb A→G mutation in intron 11 of the CFTR gene, according to previous reports [20], was also considered as a "severe" mutation. Group 3 included patients with partly functional CFTR, with at least one "mild" mutation, such as a missense mutation or splicing mutation, which creates cryptic splice sites. The 3849+10 kb C→T mutation [19, 21] and the 3272-26 G→A mutation [22] found in the patients studied are examples of such "mild" mutations. Group 4 comprised patients with only one "severe" mutation or no identified mutation, after screening of the whole coding region of the CFTR gene.

Statistical analysis

Group data were expressed as the median (range) for: the age at diagnosis; the age at Pseudomonas colonization; the number of antibiotic courses; and the number of days of antibiotics, since normal distribution of these variables cannot be demonstrated. All other variables were expressed as mean±sd. Comparisons between groups of genotypes were made by a one-way analysis of variance (ANOVA) or a Kruskal Wallis test for parametric and nonparametric variables, respectively. Parametric and nonparametric two-by-two comparisons between groups were performed using either a multiple contrasts method (Scheffe's test) or a Tukey-type multiple comparisons test, respectively [23]. Comparisons between the pancreatic sufficient and the pancreatic insufficient patients were performed using either the unpaired Student's t-test or the unpaired Mann-Whitney U-test for the parametric and the non-parametric variables, respectively. Comparisons of frequency were made with the Chi-squared method. Values of p equal to or less than 0.05 were considered to indicate statistical significance.

Results

Groups of genotypes

Mutations were identified in 191 chromosomes out of 220 (87%) (table 1). The ΔF508 mutation was found in 134 chromosomes (61%), the mutation G542X in nine (4%), whereas all the other mutations were rare, since they were found in no more than 1% of the cohort. Both mutations were identified in 89 of the 110 patients (81%), only one mutation could be detected in 13 patients (12%) and no mutation was detected in eight patients (7%). The classification of patients into four groups of genotype is reported in table 1.

Clinical characteristics at time of diagnosis (table 2)

The median age attained and age at diagnosis were different in the four groups of genotype (p<0.002 and p<0.001, respectively). Patients from Group 3 were older

Table 1. – Genotype of the 110 CF patients: details of the CF mutations and classification into four groups

Genotype groups	Genotype	Pts n
1	ΔF508/ΔF508	48*
2	ΔF508/G542X	6
	ΔF508/E827X	3†
	ΔF508/R553X	2
	ΔF508/W1282X	2
	ΔF508/E595X	1
	ΔF508/E60X	1
	ΔF508/W846X	1
	ΔF508/1078delT	1
	ΔF508/2143delT	1
	ΔF508/2347delG	1
	ΔF508/3659delC	1
	ΔF508/4382delA	1
	ΔF508/2183 AA→G	1
	ΔF508/1717-1 G→A	1
	ΔF508/1811+1.6 kb A→G	1
	E595X/Y1092X	1
	1717-1 G→A/1078delT	1
3	ΔF508/I336K	1
	ΔF508/G27E	1
	ΔF508/D192N	1
	ΔF508/I980K	1
	ΔF508/P205S	1
	ΔF508/2789+5 G→A	1
	ΔF508/3272-26 A→G	1
	G542X/3849+10 kb C→T	2‡
	G542X/2789+5 G→A	1
	W361R/297-3 C→T	1
	G551D/1717-1 G→A	1
	N1303H/2183 AA→G	1
	2789+5 G→A/2183 AA→G	1
R1070Q/D1152H	1	
R1070Q/unidentified	1	
S1251N/unidentified	1	
4	ΔF508/unidentified	7
	ΔI507/unidentified	2
	1811+1.6 kb A→G/unidentified	1
	1161delC/unidentified	1
	unidentified/unidentified	8

*: two patients are brothers; †: three brothers; ‡: two sisters. CF: cystic fibrosis; Pts: patients; A: adenine; G: guanine; C: cytosine; T: thymine.

than patients from Groups 1 and 2. The median age at diagnosis was greater in Groups 3 and 4 compared to Groups 1 and 2. Symptoms at diagnosis were different between the four groups (p<0.002): they were mainly digestive in Group 1 and respiratory in Group 3. Only eight meconium ileus were observed in this cohort: five in Group 1, two in Group 2, and one in Group 4.

Sweat chloride concentrations were lower in Group 3 compared to all the other groups (p<0.001).

Pancreatic and nutritional status (table 2)

Pancreatic insufficiency was present in all patients in Group 1 and in all but one in Group 2, but was less frequent in Group 3 (24%) and Group 4 (63%) than in Groups 1 and 2 (p<0.001). No difference appeared between the four groups either for diabetes or for the nutritional status estimated by the percentage of predicted weight.

Table 2. – Comparison of clinical data between the four groups of genotype, excluding respiratory data*

Variable	Group 1 (n=48)	Group 2 (n=26)	Group 3 (n=17)	Group 4 (n=19)	p-value ⁺
Age yrs	24 (15–40) [†]	24 (19–33) [†]	32 (18–45)	26 (18–47)	<0.002
Male sex n	34 (71)	18 (69)	9 (53)	9 (47)	NS
Age at diagnosis yrs	0.8 (0–34) ^{†‡}	2 (0–28) ^{†‡}	25 (0–37)	9 (0–45)	<0.001
Sweat chloride mmol·L ⁻¹	127±22 [†]	127±16 [†]	90±31	116±21 [†]	<0.001
Weight % pred	82±11	83±12	86±11	83±10	NS
Symptoms at diagnosis					
Respiratory n	22 [†]	14 [†]	16	13	<0.002
Digestive n	37 ^{†‡}	14 [†]	1	9	<0.001
Meconium ileus n	5	2	0	1	NS
Pancreatic insufficiency n	48 (100) ^{†‡}	25 (96) ^{†‡}	4 (24)	12 (63) [†]	<0.001
Diabetes n	6	3	2	1	NS
Hepatic cirrhosis n	5	6	1	0	NS
Gallstones n	7	4	4	5	NS
Nasal polyps n	13	7	8	2	NS
Joint pains n	10	6	2	6	NS
Pregnancy n	0	0	4	4	<0.01

*: results are expressed as median and range in parenthesis, as mean±SD, or as absolute value and percentage in parenthesis. NS: nonsignificant; % pred: percentage of predicted value. +: for comparison of all four groups; †: p<0.05, compared to Group 3; ‡: p<0.05, compared to Group 4.

As compared to the pancreatic insufficient patients, the pancreatic sufficient subjects: were older (31 (18–47) vs 24 (15–40) yrs; p<0.001); were diagnosed later (24 (1–45) vs 2 (0–34) yrs; p<0.001); had lower sweat chloride concentrations (94±24 vs 127±20 mmol·L⁻¹; p<0.001); had later colonization by *P. aeruginosa* (26 (16–45) vs 15 (1–34) yrs; p<0.001); had better FVC (64±31 vs 52±21% pred; p<0.05) and FEV₁ (51±33 vs 38±22% pred; p<0.05); and slower estimated mean annual loss in FVC (1.2±1.1 vs 2.0±0.9%; p<0.001) and in FEV₁ (1.6±1.2 vs 2.5±1.0%; p<0.001).

Pulmonary status (table 3)

There was a trend towards lower FVC and FEV₁ values in Group 2 compared to Groups 3 and 4 (p=0.15 and p=0.09, respectively). The estimated mean annual loss in FVC and in FEV₁ was greater in Groups 1 and 2

than in Groups 3 and 4 (p<0.01). P_aO₂ was significantly lower in patients from Group 2 than in patients from Groups 3 and 4, and in patients from Group 1 than in patients from Group 4 (p<0.02).

Pseudomonas aeruginosa colonization was more frequent in Group 1 than in Groups 3 and 4, and the age at colonization was lower in Group 1 than in Group 3 (p<0.01). The number of *i.v.* antibiotic courses and the number of days of *i.v.* antibiotics per year differed significantly between the four groups of genotype (p<0.01). Patients from Group 3 received fewer *i.v.* antibiotic courses than patients from Group 2, and the duration of their *i.v.* antibiotic infusions was shorter than in patients from Groups 1 and 2.

The comparison of the pulmonary status between the combined Groups 1 and 2 and the combined Groups 3 and 4 reinforced the results obtained with the four groups analysed separately. Thus, the condition of the combined Groups 1 and 2 appeared more severe than the

Table 3. – Comparison of pulmonary data between the four groups of genotype*

Variable	Group 1 (n=48)	Group 2 (n=26)	Group 3 (n=17)	Group 4 (n=19)	p-value ⁺
FVC % pred	54±20	49±22	62±30	62±24	NS
FEV ₁ % pred	40±21	34±22	48±30	50±28	NS
Estimated annual loss in FVC %	2.0±0.9 [†]	2.2±1.0 ^{†‡}	1.2±1.0	1.5±1.1	<0.01
Estimated annual loss in FEV ₁ %	2.6±1.0 ^{†‡}	2.8±1.0 ^{†‡}	1.7±1.1	1.9±1.3	<0.005
P _a O ₂ kPa	8.8±1.5 [‡]	8.3±1.7 ^{†‡}	9.5±1.9	9.9±1.5	<0.01
P _a CO ₂ kPa	5.7±0.9	5.9±1.1	5.5±1.3	5.2±0.5	NS
<i>P. aeruginosa</i> colonization n	45 ^{†‡}	22	11	13	<0.02
Age at <i>P. aeruginosa</i> colonization yrs	15 (1–33) [†]	17 (4–30)	25 (8–34)	19 (7–45)	<0.01
Number of <i>i.v.</i> antibiotic courses·yr ⁻¹	3 (0–12)	4 (0–12) [†]	0 (0–10)	1 (0–7)	<0.01
Number of days of <i>i.v.</i> antibiotics·yr ⁻¹	42 (0–365) ^{†‡}	64 (0–365) [†]	0 (0–300)	15 (0–142)	<0.01
Pneumothorax n	6	2	3	1	NS

*: results are expressed as median and range in parenthesis or as mean±SD. FVC: forced vital capacity; % pred: percentage of predicted value; FEV₁: forced expiratory volume in one second; P_aO₂: arterial oxygen tension; P_aCO₂: arterial carbon dioxide tension; NS: nonsignificant. +: for comparison of all four groups; †: p<0.05, compared to Group 3; ‡: p<0.05, compared to Group 4.

combined Groups 3 and 4, not only for the mean annual loss in FVC (2.0 ± 1.0 vs $1.4 \pm 1.0\%$, respectively; $p < 0.001$), the mean annual loss in FEV₁ (2.6 ± 1.0 vs $1.9 \pm 1.0\%$, respectively; $p < 0.001$), and the P_{a,O_2} (8.5 ± 1.6 vs 9.6 ± 1.6 kPa respectively; $p < 0.002$), but also for FVC (52 ± 21 vs $61 \pm 25\%$ pred, respectively; $p < 0.05$), FEV₁ (38 ± 22 vs $48 \pm 25\%$ pred, respectively; $p < 0.05$) and P_{a,CO_2} (5.7 ± 0.9 vs 5.3 ± 1.1 kPa, respectively; $p < 0.05$). *P. aeruginosa* colonization was more frequent (88 vs 71%; $p < 0.05$), the age at colonization was earlier (median (range) 15 (1–33) vs 20 (8–45) yrs, respectively; $p < 0.005$), and the antibiotic courses were more frequent (median (range) 3 (0–12) vs 1 (0–10), respectively; $p < 0.001$) in the combined Groups 1 and 2 than in the combined Groups 3 and 4.

Comparison and discrepancies between pulmonary and pancreatic status

We have analysed the genotype characteristics of the adult CF patients with mild respiratory disease, defined as FEV₁ greater than 70% of the predicted values. Among the 15 patients with mild respiratory insufficiency, 10 had pancreatic insufficiency and two had hepatic cirrhosis (table 4). Mild respiratory insufficiency was

observed in the four genotype groups, even in Group 2 which is believed to carry the most severe dysfunctions of the CFTR protein.

The distribution of the five patients with pancreatic sufficiency but with severe lung disease (FEV₁ <30% pred) among the groups of genotypes is shown in table 5. These patients comprised the only pancreatic sufficient patient from Group 2, but also three patients from Group 3 and one from Group 4.

Other clinical data

No difference was observed between the four groups for history of nasal polyps, hepatic cirrhosis, gallstones, joint pains (table 2) or frequency of pneumothorax (table 3). Among the 40 females, eight had successful pregnancies, all in Groups 3 and 4 ($p < 0.01$).

Discussion

Respiratory insufficiency is the major cause of mortality in cystic fibrosis. Although pulmonary involvement is observed in the large majority of patients, the rate of decline in pulmonary function as well as the severity

Table 4. – Characteristics of patients with FEV₁ >70% of predicted value

Genotype	Age yrs	Age at diagnosis yrs	FVC % pred	FEV ₁ % pred	<i>P. aeruginosa</i> colonization	PI	Hepatic cirrhosis
Group 1							
ΔF508/ΔF508	18	<1	83	75	Yes	Yes	No
ΔF508/ΔF508	19	8	88	72	Yes	Yes	Yes
ΔF508/ΔF508	24	<1	87	84	Yes	Yes	No
ΔF508/ΔF508	25	13	85	82	Yes	Yes	No
ΔF508/ΔF508	37	34	90	83	No	Yes	No
Group 2							
ΔF508/E827X	18	<1	82	76	Yes	Yes	Yes
ΔF508/W846X	29	27	101	95	No	Yes	No
ΔF508/W1282X	31	28	91	77	No	Yes	No
Group 3							
2789+5 G→A/G542X	18	2	107	103	No	No	No
2789+5 G→A/2183 AA→G	36	34	93	87	No	No	No
ΔF508/G27E	39	28	115	78	No	No	No
Group 4							
ΔI507/unid	18	<1	103	103	No	Yes	No
unid/unid	26	5	89	77	No	Yes	No
unid/unid	39	38	96	87	No	No	No
unid/unid	40	38	110	106	No	No	No

PI: pancreatic insufficiency; unid: unidentified. For further definitions see legend to tables 1 and 3.

Table 5. – Characteristics of patients with pancreatic sufficiency and FEV₁ <30% of predicted value

Genotype	Age yrs	Age at diagnosis yrs	FVC % pred	FEV ₁ % pred	<i>P. aeruginosa</i> colonization
Group 2					
ΔF508/4382del A	27	18	43	21	Yes
Group 3					
ΔF508/3272-26 G→A	21	8	32	28	Yes
G542X/3849+10kb C→T	35	32	53	25	Yes
W361R/297-3 C→T	45	25	25	12	Yes
Group 4					
unid/unid	30	15	38	27	Yes

unid: unidentified. For further definitions see legends to tables 1 and 3.

and the nature of pulmonary infection are very different among patients. In the present study of 110 adult CF patients it was possible to demonstrate a relationship between the groups of genotype and the severity of respiratory involvement, which was expressed both for pulmonary function and severity of airway infection by *P. aeruginosa*.

Patients from Group 3 (with at least one mild mutation) and Group 4 (with at least one unidentified mutation) exhibited a milder lung disease than patients from Group 1 ($\Delta F508$ homozygotes) and Group 2 (with two severe mutations). In the former groups, whereas patients were older, their P_{a,O_2} was higher and the decline in pulmonary function estimated by the mean annual loss in FVC and FEV₁ was slower than in the two groups corresponding to the severe genotypes (Groups 1 and 2). Very few reports have related pulmonary severity to genotype. In the studies by GAN *et al.* [14], and STRONG *et al.* [13] patients with the A455E mutation or with the G551S mutation, respectively, who would be classified as Group 3 in the present study, had mild pulmonary disease. Another study in eight adult CF patients suggested that a less severe phenotype, including pulmonary phenotype, was genetically determined [24]. The latter finding is in accordance with the present results, since these eight patients were compound heterozygotes who corresponded to Groups 3 and 4.

In a study including a large number of patients carrying the $\Delta F508$ mutation, no difference was found in the severity of pulmonary involvement between a group of $\Delta F508$ homozygotes and groups of compound heterozygotes for $\Delta F508$ [11]. In the latter study, most of the compound heterozygotes for $\Delta F508$ were associated with another mutation (G542X, R553X, W1282X, 1717-1G→A, 621+1G→T) that corresponded to Group 2. Therefore, the lack of difference in the severity of pulmonary involvement between homozygotes for $\Delta F508$ and these compound heterozygotes is in accordance with the present results. For the two genotypes, R117H/ $\Delta F508$ and N1303K/ $\Delta F508$, which would be classified as Group 3, the lack of difference in pulmonary involvement might be due to a selection bias created by mortality in this cross-sectional study and/or by the young age of the patients in most groups of this study, since the mean age of all groups but one was less than 15 yrs.

As compared to the two groups with severe mutations (Groups 1 and 2), we found that patients from Group 3 (mild mutation) had less severe bacterial airway infections. These results are in agreement with the study by KUBESCH *et al.* [12], who showed that the risk of chronic airway colonization by *P. aeruginosa* was related to the CFTR mutation genotype. It was found that the $\Delta F508$ /nonsense compound heterozygotes had the highest risk for early colonization by *P. aeruginosa*, whereas the missense and splicing mutations were low-risk alleles for the acquisition of the most prevalent opportunistic pathogen. The severity of the respiratory involvement of patients from Group 4 (with only one severe mutation or with unidentified mutations) appeared to be intermediate between subjects from Group 3 and subjects with severe mutations (Groups 1 and 2). These characteristics may be explained by the fact that patients in Group 4 had heterogenous genotypes, among which were unidentified mutations. The entire coding sequence

of the gene was analysed, so that most of the unidentified mutations in Group 4 are probably located in non-coding regions of the CFTR gene, locations that are likely to induce less severe phenotypes. However, mutations in noncoding regions may also induce severe dysfunction of the CFTR protein.

Concerning pancreatic status, our study is in accordance with previous studies which showed that patients carrying two copies of "severe" mutations (nonsense mutations, splicing mutations due to change in the consensus splicing sequences and frameshift mutations) exhibit pancreatic insufficiency, whereas patients with a "mild" mutation on one or both chromosomes are pancreatic sufficient [5, 25]. The milder severity both of pulmonary and pancreatic involvement is likely to explain why patients from Group 3 and 4 had more advanced age at diagnosis than patients from the two groups with severe mutations (Groups 1 and 2). An association between the age at diagnosis and the type and location of the CFTR mutation has already been observed, with earlier diagnosis in $\Delta F508$ homozygotes and compound heterozygotes with a second mutation being a nonsense or a frameshift mutation and/or a mutation located in the nucleotide binding folds [12].

The finding of a milder severity of disease in groups corresponding to milder mutations (Groups 3 and 4) as compared to groups with severe mutations (Groups 1 and 2) is unlikely to be explained by a biased selection of subjects. By studying adult patients, very severe phenotypes of the disease responsible for death early in life were excluded. These very severe phenotypes are likely to belong to the severe mutations of Groups 1 and 2, which probably explains why the frequency of $\Delta F508$ homozygote patients in this population (44%) was lower than in the general population of caucasian CF patients. Therefore, we believe that our finding of a milder phenotype in Groups 3 and 4 would have been reinforced if we had included the most severe forms of the disease responsible for early death.

Although we have usually observed a good relationship between the severity of pancreatic and pulmonary involvement, it is of interest to point out that in some individuals there were discrepancies between the severity of the pulmonary and pancreatic status. A few patients elicited pancreatic sufficiency but severely impaired respiratory function. All but one of these patients were classified in genotype Groups 3 and 4. Conversely, 10 patients showed a good respiratory function associated with pancreatic insufficiency. These cases were mostly observed in Groups 1 and 2, despite a $\Delta F508$ mutation either homozygote or associated with a nonsense mutation (E827X, W846X and W1282X). Such atypical, mild pulmonary illness with severe pancreatic disease has already been described in CF patients with two nonsense mutations [26]. Several hypotheses have been proposed to explain the unexpected mild pulmonary or pancreatic involvement despite a severe mutation. A first hypothesis is that the truncated protein could be partially functional in certain tissues. Another explanation comes from alternative splicing of messenger ribonucleic acid (mRNA), which produces different forms of a protein and could minimize the effect of a gene mutation if the mutation occurs in an exon that is not included in a mature transcript encoding a functional

isoform. This process can occur in a tissue-specific manner, and expression of this CFTR isoform only in airway epithelia could explain the difference in severity of pulmonary and pancreatic disease.

These results show that, as a whole, the type of mutations responsible for CF makes it possible to predict the outcome and the severity of the disease in most adult patients and, in particular, the severity both of the pancreatic and respiratory involvement. The classification proposed by WELSH and SMITH [15] was based on the expected severity of dysfunction of the CFTR protein depending on the expected consequence of the mutation on the processing and/or function of the protein. Interestingly, this hypothesis was indirectly confirmed by measurement of the function of the chloride channel in the sweat gland, since we observed that in addition to an overall less severe disease, patients from Group 3 also elicited a chloride concentration in sweat that was lower than patients from all the other groups. However, although we found a relationship between the genotype and the phenotype in our four groups of genotypes as a whole, some individuals within each group did not fit the general behaviour of the group, despite the fact that they were carrying the same mutation.

These findings were reported in several studies and are of great interest in the understanding of the disease. Several hypotheses were proposed to explain such discrepancies in phenotype within patients of the same genotype, although none of them are fully convincing. One possible hypothesis is that variants in the CFTR gene that alter splicing efficiency of an exon can affect the phenotype produced by a mutation. This possibility has been described recently for exon 9 [27]. Another possible hypothesis is the existence of genetic factors outside the CF locus [4, 11, 28]. DÖRING *et al.* [28] have shown that the genotype for α_1 -antitrypsin was correlated to the phenotype of CF, with carriers of deficient α_1 -antitrypsin variants having an earlier onset of *P. aeruginosa* lung infection, higher total immunoglobulin G (IgG) and specific *P. aeruginosa* serum antibodies. Finally, it has been suggested that nongenetic environmental factors, such as exposure to passive smoking [29], may also influence the course of pulmonary disease in patients with cystic fibrosis [30].

This classification of 110 adult CF patients into four groups of genotype, according to the probable remaining function of their cystic fibrosis transmembrane conductance receptor protein, showed a relationship between the presumed severity of the mutations and the observed phenotype, not only for pancreatic status but also for the severity of respiratory involvement. Patients with the expected partially functional cystic fibrosis transmembrane conductance receptor (Group 3 and, to a lesser extent, Group 4), appeared to have less severe disease, in particular for pancreatic status, respiratory function and airway infection by *P. aeruginosa*. Our results show that the attempt to identify and to classify the mutations in cystic fibrosis is of interest for the prediction of the outcome of the disease. However, like others, we found that some individuals do not fit the outcome of their general group of genotype. Our findings reinforce the conviction that other factors besides the cystic fibrosis transmembrane conductance receptor mutation are involved in the severity and outcome of the disease.

Acknowledgements: The authors thank S. Bousquet and C. Herbulot for their technical assistance in the identification of the CFTR mutations. They are grateful to A. Baculard, J. Just, A. Sardet, G. Tournier (Hôpital Trousseau, Paris), A. Munck, J. Navarro (Hôpital Robert Debré, Paris), J. de Blic, M. Le Bourgeois, G. Lenoir, M. Sorin (Hôpital Necker-Enfants-Malades, Paris) and P. Foucaud (Hôpital Louis Mignot, Versailles) for their close collaboration in the collection of clinical data of the patients and in gathering the cohorts of adult patients.

References

1. Riordan JR, Rommens JM, Kerem BS, *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989; 245: 1066–1073.
2. Kerem B, Rommens JM, Buchanan JA, *et al.* Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989; 245: 1073–1080.
3. Tsui LC. The spectrum of cystic fibrosis mutations. *Trends Genet* 1992; 8: 392–398.
4. Rosenstein BJ. Genotype-phenotype correlations in cystic fibrosis. *Lancet* 1994; 343: 746–747.
5. Kerem E, Corey MD, Kerem BS, *et al.* The relation between genotype and phenotype in cystic fibrosis: analysis of the most common mutation ($\Delta F508$). *N Engl J Med* 1990; 323: 1517–1522.
6. Borgo G, Gasparini P, Bonizzato A, Cabrini G, Mastella G, Pignatti PF. Cystic fibrosis: the $\Delta F508$ mutation does not lead to an exceptionally severe phenotype. A cohort study. *Eur J Pediatr* 1993; 152: 1006–1011.
7. Santis G, Osborne L, Knight RA, Hodson ME. Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. *Lancet* 1990; 336: 1081–1084.
8. Johansen HK, Nir M, Høiby N, Koch C, Schwartz M. Severity of cystic fibrosis in patients homozygous and heterozygous for $\Delta F508$ mutation. *Lancet* 1991; 337: 631–634.
9. Campbell PW, Phillips JA, Krishnamani MRS, Maness KL, Hazinki TA. Cystic fibrosis: relationship between clinical status and $\Delta F508$ deletion. *J Pediatr* 1991; 118: 239–241.
10. Burke W, Aitken ML, Chen SH, Scott CR. Variable severity of pulmonary disease in adults with identical cystic fibrosis mutations. *Thorax* 1992; 102: 506–509.
11. The Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between genotype and phenotype in patients with cystic fibrosis. *N Engl J Med* 1993; 329: 1308–1313.
12. Kubesch P, Dörk T, Wulbrand U, *et al.* Genetic determinants of airways' colonisation with *Pseudomonas aeruginosa* in cystic fibrosis. *Lancet* 1993; 341: 189–193.
13. Strong TV, Smit LS, Turpin SV, *et al.* Cystic fibrosis gene mutation in two sisters with mild disease and normal sweat electrolytes levels. *N Engl J Med* 1990; 325: 1630–1634.
14. Gan KH, Veeze HJ, Van den Ouweland AMW, *et al.* A cystic fibrosis mutation associated with mild lung disease. *N Engl J Med* 1995; 333: 95–99.
15. Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* 1993; 73: 1251–1254.
16. Bienvenu T, Cazeneuve C, Kaplan JC, Beldjord C. Mutation heterogeneity of cystic fibrosis in France: screening by denaturing gradient gel electrophoresis using psoralen-modified oligonucleotide. *Hum Genet* 1995; 6: 23–29.

17. Fanen P, Ghanem N, Vidaud M, *et al.* Molecular characterization of cystic fibrosis: 16 novel mutations identified by analysis of the whole cystic fibrosis conductance transmembrane regulator (CFTR) coding regions and splice site junctions. *Genomics* 1992; 13: 770–776.
18. Sanger F, Milken S, Coulson AR. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci* 1977; 74: 5463–5467.
19. Highsmith WE, Burch LH, Zhou Z, *et al.* A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. *N Engl J Med* 1994; 331: 974–980.
20. Chillón M, Dörk T, Casals T, *et al.* A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811+1.6 kb A→G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. *Am J Hum Genet* 1995; 56: 623–629.
21. Augarten A, Kerem BS, Yahav Y, *et al.* Mild cystic fibrosis and normal or borderline sweat test in patients with the 3849+10kb C→T mutation. *Lancet* 1993; 342: 25–26.
22. Kanakis E, Tzetzis M, Antonialdi T, *et al.* Mild cystic fibrosis phenotype in patients with the 3272-26 A/G mutation. *J Med Genet* 1995; 32: 406–407.
23. Zar JH. Biostatistical analysis. Englewood Cliffs, NJ, USA, Prentice-Hall, 1984; pp. 718.
24. Férec C, Verlingue C, Guillermit H, *et al.* Genotype analysis of adult cystic fibrosis patients. *Hum Mol Genet* 1993; 2: 1557–1560.
25. Kristidis P, Bozon D, Corey M, *et al.* Genetic determination of exocrine pancreatic function in cystic fibrosis. *Am J Hum Genet* 1992; 50: 1178–1184.
26. Cutting GR, Kasch LM, Rosenstein BJ, Tsui LC, Kazazian HH, Antonarakis SE. Two patients with cystic fibrosis, nonsense mutations in each cystic fibrosis gene, and mild pulmonary disease. *N Engl J Med* 1990; 323: 1685–1689.
27. Kiesewetter S, Macek MJ, Davis C, *et al.* A mutation in CFTR produces different phenotypes depending on chromosomal background. *Nature Genet* 1993; 5: 274–278.
28. Döring G, Krogh-Johansen H, Weldinger S, Høiby N. Genotypes of α_1 -antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for $\Delta F508$. *Pediatr Pulmonol* 1994; 18: 3–7.
29. Campbell PW, Parker RA, Roberts BT, Krishnamani MR, Phillips JA. Association of poor clinical status and heavy exposure to tobacco smoke in patients with cystic fibrosis who are homozygous for the $\Delta F508$ deletion. *J Pediatr* 1992; 120: 261–264.
30. Dean M, Santis G. Heterogeneity in the severity of cystic fibrosis and the role of CFTR gene mutations. *Hum Genet* 1994; 93: 364–368.