

J.L. Serre^{a,b}
E. Mornet^{b,c}
B. Simon-Bouy^{b,c}
J. Boué^c
A. Boué^c

^a Laboratoire d'Epidémiologie
Génétique, INSERM U. 155,
Paris,

^b Laboratoire de Cytogénétique et
Génétique Moléculaire Humaine,
Université de Versailles-Saint
Quentin, Versailles, et

^c Centre d'Etude de Biologie
Prénatale, Paris, France

General Cystic Fibrosis Mutations Are Usually Missense Mutations Affecting Two Specific Protein Domains and Associated with a Specific RFLP Marker Haplotype

Key Words

Cystic fibrosis
Population genetics
RFLP markers
Selection
Heterozygote advantage

Abstract

Some 250 different mutations have so far been screened in the cystic fibrosis (CF) gene. The 50 nonsense, 33 splicing and 60 frameshift mutations are randomly distributed within the gene, unlike the 107 missense mutations or amino acid deletions. A large excess of missense mutations affects the exons encoding the first transmembrane (MS1) and first ATP-binding fold (NBF1) domains. Sixty-four of the 107 missense mutations may be classified as private, demic, local and general mutations on the basis of their geographic distribution in Europe. Private and demic mutations are randomly distributed within the gene; local and general mutations are not. It is well known that some RFLP markers are in linkage disequilibrium with some mutations. Private, demic and local mutations are randomly associated with each class of RFLP haplotypes. In contrast, general mutations, frequent and infrequent, are not randomly associated with RFLP markers. General mutations usually affect a specific part of the gene and are more likely to be associated with a specific RFLP marker. This suggests the existence of selective factors favoring these mutations, a hypothesis formerly postulated as a possible cause of the high frequency of the disease.

Introduction

Cystic fibrosis (CF) is an autosomal recessive disease due to a deleterious mutation in a chloride channel gene (CFTR = CF transmembrane conductance regulator), located on chromosome 7. From a biological point of view, molecular cloning of the gene [1], purification of the protein, and subsequent analyses will increase understanding of the molecular and cellular physiopathology of CF (chronic obstructive lung disease and pancreatic enzyme insufficiency).

From a medical point of view, identification of a large number of deleterious mutations and of microsatellite sequences within the CFTR gene provides a highly effective means of prenatal diagnosis or even, in some specific populations, of carrier screening [2–5]. CF is a notorious conundrum in population genetics. Why is this disease so frequent among Caucasians but unknown in other populations? In Caucasians, the mean prevalence at birth is 1 in 2,500. According to the Hardy-Weinberg law, this means that the frequency of the mutation, or more exactly, the frequency of the cluster of deleterious mutations of the CFTR gene, is equal to 2%, and unaffected carriers are numerous – 4%, or 1 in 25. How could a lethal mutation have reached such a frequency?

Since the early sixties, population geneticists have developed various hypotheses and models. A balance between negative selection and recurrent mutations was ruled out long before molecular data provided definitive evidence, and genetic drift [6–8] may be questionable. What frequency of the $\Delta F508$ mutation would genetic drift have produced in remote eras (neolithic or paleolithic) before natural selection lowered it to the present 1.5%? So the only surviving model postulates selective factors favoring CF mutations through heterozygote advantage and/or

meiotic drive [8–13]. But this balanced polymorphism hypothesis cannot be easily tested, either statistically, because the required sample size would be too large [7], or physiologically, because the CFTR protein is still under study.

Between October 1989 and December 1992, 250 mutations were characterized on the CF gene in a worldwide survey conducted by the Cystic Fibrosis Genetic Analysis Consortium (see Appendix). In this study, the cartographic location of the observed CF mutations was analyzed depending on their respective nature (nonsense, splicing, frameshift or missense), using a null hypothesis of random mutation at each potential site. Molecular data were examined to see if they pointed to the possible existence of selective factors favoring CF mutations. The set of CF mutations was thus classified on the basis of geographic distribution. Each class of mutation was then analyzed according to the distribution of mutations within the gene or peptide chain, and according to the associated RFLP markers in linkage disequilibria.

Material and Methods

Since identification of the predominant $\Delta F508$ mutation of the CFTR in 1989 [14] and subsequent study in all populations [15–17], the CF gene has been extensively screened. Our analysis refers to the 250 different mutations (50 nonsense, 33 splicing, 60 frameshift and 107 missense mutations or amino acid deletions) listed by the Cystic Fibrosis Genetic Analysis Consortium in December 1992 (partly confidential data). CF mutations were characterized within samples, ranging from 29,567 CF chromosomes for the predominant $\Delta F508$ mutation to a few hundred CF chromosomes for private mutations. For most of the mutations, a few thousand CF chromosomes were studied.

For each kind of mutation, the observed distribution of mutations between the exons of the CF gene was compared to the expected distribution using a null hypothesis of random mutation at potential sites, depending on the respective natures of the studied muta-

tions. Since the CF gene sequence is known [18], the expected random distribution of mutations was calculated using either the respective and variable exon length for frameshift or missense mutations, or the potential sites within the DNA sequence for splicing or nonsense mutations. Exons 6a and 6b, 14a and 14b, and 17a and 17b were not distinguished. The gene is therefore partitioned in 24 exons and 23 introns. The cartographic location of the missense mutations between the domains of the protein has been studied by grouping the corresponding exons.

Only 155 CF mutations out of the total of 250 could be divided into the following four classes on the basis of geographic distribution because the more recently characterized mutations cannot yet be classified in this way: *private* mutations observed only once in the worldwide survey of the Cystic Fibrosis Consortium; *demic* mutations observed twice or more, but within the same population; *local* mutations observed in two or three closed populations or countries, and *general* mutations observed everywhere, or in most countries. Such classification is provisional because some private mutations may have been misclassified since most of the laboratories did not test all the identified mutations within their patients' DNA. Some private mutations may therefore actually be demic or even local.

A great number of molecular markers has been detected near the CFTR locus, especially RFLPs like *XV2C/Taq1* and *KM19/Pst1* [19]. Depending on the presence or absence of the respective endonuclease sites, there are four kinds of haplotypic or chromosomal combinations: A = (-,-); B = (-,+); C = (+,-); or D = (+,+). Since 1986, molecular analyses of RFLP haplotypes within affected and control individuals have provided evidence for a close association between the B haplotype and the disease. This association was probably due to a high disequilibrium between this marker and one predominant or several deleterious alleles. This hypothesis proved correct after identification of $\Delta F508$ by Kerem et al. [14]. The expected random associations between mutations and each kind of RFLP were calculated according to the mean European frequency of these RFLP haplotypes on normal chromosomes [16].

The significance level (p value) of χ^2 tests was calculated using the tabulated values, except for statistical tests for which sample size were too small. In these cases, the exact p values were computed, using a turbo-Pascal program [20] which generates the exact probability distribution of χ^2 .

Results and Discussion

Cartographic Distribution of the 250 CF Mutations

The cartographic distribution is shown in table 1. As there is no disparity between observed and expected distributions, splicing and frameshift mutations may be considered to be randomly distributed. This conclusion is still valid when exons are grouped in order to obtain expected numbers higher than 5. Four years ago molecular geneticists started their hunt for CF mutations other than $\Delta F508$, nonrandomly with regard to the domains or exons. The random distribution of splicing or frameshift mutations suggests that the whole gene has now been screened so there is no more census bias in the cartographic distribution of some kinds of mutations. The hypothesis of the existence of a mutation hot spot [17, 21] must therefore be questioned, at least for this kind of mutation (splicing, frameshift).

The test value for the distribution of nonsense mutations is borderline, even when grouping exons. If significant, the low number of nonsense mutations in the protein C-terminal would be in agreement with the observation that deletions of this domain may not affect protein activity [22].

There is a highly significant disparity between the observed and expected distribution of missense mutations or amino acid deletions (table 1, last column). As previously noted [21], there is an excess of mutations in NBF1 as well as in MS1. An alternative explanation of the existence of a mutation hot spot in this part of the gene is that amino acid substitutions in MS1 or in NBF1 are far more critical for protein folding than missense mutations affecting MS2 or NBF2.

Sixty-four missense mutations out a total of 107 could be classified according to their geographic dispersion. Private and demic

Table 1. Disparity between observed and expected distributions of CF mutations depending on their nature

Exon ^a	Mutations							
	nonsense		splicing		frameshift		missense and aa deletions	
	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.
1	1	0.48	1	1.27	0	0.73	0	1.29
2	3	1.83	0	1.27	1	1.50	2	2.69
3	4	1.64	2	1.27	2	1.45	6	2.59
4	4	2.02	3	1.27	8	2.91	8	5.19
5	0	0.77	1	1.27	1	1.23	3	2.17
6 a+b	1	3.56	3	2.54	1	3.94	5	7.02
7	0	2.22	0	1.27	6	3.33	13	4.45
8	1	1.54	1	1.27	0	1.25	0	2.23
9	0	2.02	1	1.27	0	2.48	2	4.40
10	3	2.31	2	1.27	2	2.60	7	4.63
11	4	1.16	1	1.27	1	1.30	13	2.31
12	1	1.25	2	1.27	1	1.18	3	2.11
13	8	9.06	1	1.27	12	9.76	9	17.41
14 a+b	4	1.45	1	2.54	1	2.28	2	4.06
15	1	2.02	0	1.27	4	3.36	3	6.01
16	0	0.39	0	1.27	1	1.09	1	1.95
17 a+b	4	3.56	2	2.54	3	5.10	17	9.11
18	0	0.96	3	1.27	0	1.38	1	2.45
19	5	2.89	4	1.27	8	3.38	2	6.01
20	2	2.41	3	1.27	3	2.10	7	3.76
21	3	1.35	0	1.27	2	1.23	2	2.17
22	1	1.54	0	1.27	2	2.35	1	4.19
23	0	1.35	2	1.27	0	1.41	0	2.52
24	0	2.22			1	2.68	0	4.76
Total	50		33		60		107	
χ^2	34.8		17.59		31.6		109.4	
p	0.056		0.72		0.11		<10 ⁻⁴	

The expected distribution was calculated using the potential nonsense mutation sites in the coding sequence, or the potential number of splicing mutation sites (introns 6a and 6b have been unified as have 14a and 14b, and 17a and 17b), or the respective length of various exons for frameshift and missense mutations. aa = Amino acid; obs. = observed; exp. = expected.

^a For splicing mutations, the number in this column refer to the intron.

Table 2. Disparity between the observed and random expected distributions between the domains of the CFTR protein for missense mutations and amino acid deletions

Domains of the protein	Missense mutations									
	total		private		demic		local		general	
	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.
N terminal	2	3.98	1	1.08	0	0.45	0	0.45	0	0.41
MS1	35	23.65	9	6.82	3	2.83	4	2.83	4	2.59
NBF1	25	13.45	5	3.64	4	1.51	5	1.51	6	1.38
R	9	17.41	3	4.72	1	1.95	0	1.95	0	1.79
MS2	24	23.58	5	6.39	1	2.64	3	2.64	0	2.42
NBF2	12	16.13	6	4.37	3	1.81	0	1.81	1	1.66
C terminal	0	7.28	0	1.98	0	0.82	0	0.82	0	0.75
Total	107		29		12		12		11	
χ^2	28.8		4.73		7.6		13.6		21.9	
p	$<5 \cdot 10^{-4}$		0.4		0.11		0.035		$<10^{-3}$	

Only 64 missense mutations out of a total of 107 could be assigned as private, demic, local or general. obs. = Observed; exp. = expected.

mutations are randomly distributed within the CF gene, whereas local, and especially general, mutations are not (table 2). Both of these classes are almost always mutations within MS1 and NBF1 domains. The fact that private and demic mutations are randomly distributed while local and general mutations, the so-called successful mutations, are mostly confined to specific locations in the peptide chain, is in agreement with the hypothesis of selective factors favoring the expansion of these mutations. It is hardly likely that migration, founder effect or genetic drift would have only favored the spread of 19 MS1 or NBF1 missense mutations out a total of 23.

Geographic Dispersion

The classification of the 155 mutations of the cluster is reported in table 3. Private mu-

Table 3. Classification of the set of CF mutations according to their geographic pattern

Class of mutation	Number of mutations	Percent of the total number	Percent of the total CF chromosomes
Private	73	47	0.25
Demic	35	23	5
Local	30	19	10
General	17	11	84.75

tations, though numerous, only account for 0.25% of CF chromosomes, due, of course, to their very low relative frequency. General mutations account for 11% of the cluster but for nearly 85% of all CF chromosomes (the prevalent $\Delta F508$ mutation accounts for 67% of them).

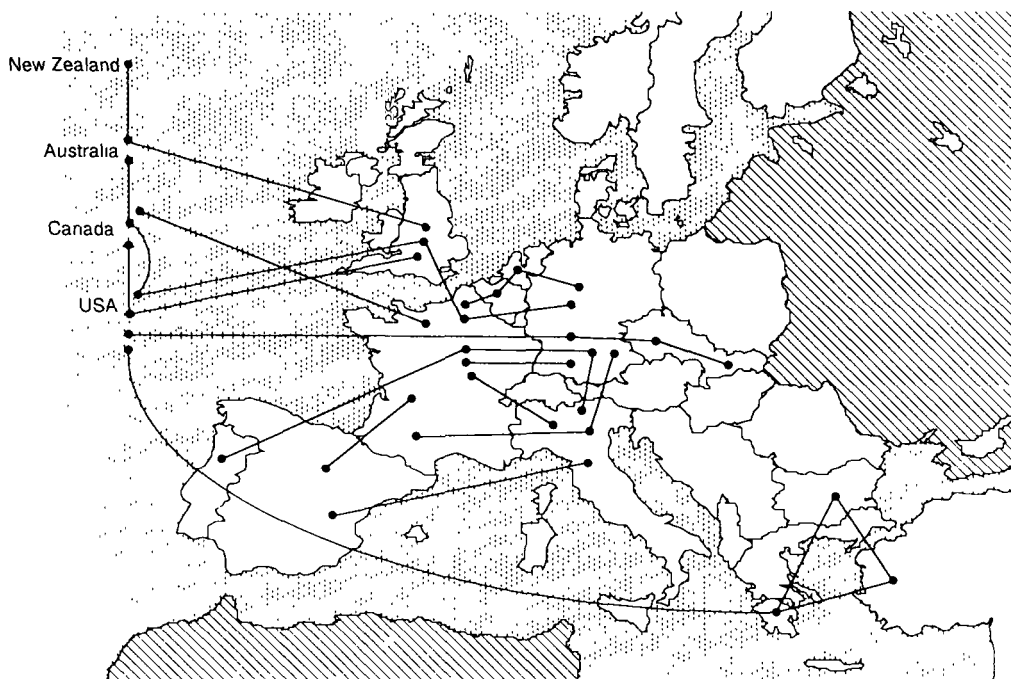


Fig. 1. Geographic distribution of a few local cystic fibrosis mutations (points refer to countries where a mutation has been screened, not to the location of the mutation within each country).

Without any general mutations, CF would be a very rare disease (one affected newborn in more than 100,000), which is probably the case in non-Caucasian populations. Even without $\Delta F508$, CF would be a common recessive disease (one in 23,000). The fact that CF is so frequent among Caucasians is only due to general mutations, especially $\Delta F508$, which from a population genetics standpoint are 'successful mutations' because they have diffused in most populations. Each of these mutations has reached a frequency which is not in agreement with mutation-selection balance, or even with genetic drift for $\Delta F508$.

The geographic pattern of local mutations (fig. 1) reflects the common origin and history of population migrations, for instance between Germany, Bohemia and Slovakia, be-

tween Germany and France, France and England, France and Canada, and particularly between Europe as a whole and North America.

Association of CF Mutations with RFLP Markers

To date, 47 mutations (15 private, 9 demic, 6 local and 17 general) have been reported together with their associated RFLP haplotypes. Table 4 shows the observed numbers of mutations for each class of mutation and each kind of associated RFLP haplotype. Two general mutations (S549N and R553X) were associated with two different haplotypes and were entered as two halves for each haplotype.

Table 4. Numbers of observed mutations associated with each kind of RFLP haplotype, for each geographic class of CF mutation, and expected numbers of mutations assuming that RFLP haplotypes and mutations are randomly associated

Class of RFLP		Class of mutation			
		private	demic	local	general
RFLP A	obs.	5	5	2	3
	exp.	5.17	3.11	2.07	5.87
RFLP B	obs.	2	1	2	8.5
	exp.	2.47	1.47	0.99	2.81
RFLP C	obs.	5	1	2	3.5
	exp.	5.49	3.29	2.19	6.21
RFLP D	obs.	3	2	0	2
	exp.	1.87	1.13	0.75	2.11
Total		15	9	6	17
χ^2		0.8	3.6	1.8	14.1
p		0.84	0.3	0.6	$<5 \cdot 10^{-3}$

obs. = Observed; exp. = expected.

There is clearly no disparity between the random expected and the observed distributions of associated RFLPs within private, demic and local mutations. In contrast, there is a large excess of B-associated haplotypes within the general mutations. Not only is $\Delta F508$ largely associated with the B haplotype, but so too are some of the most frequent secondary mutations, namely 621 + 1G \rightarrow T, A455E, 1717-1 G \rightarrow T, G542X, S549N, G551D, W1282X, and N1303K.

Overspread or so-called successful mutations seem to be more often associated with a B haplotype, although this marker is the least frequent among normal chromosomes. This fact may be consistent with the existence of selective factors postulated by advocates of meiotic drive or heterozygote advantage in

order to explain the high disease frequency. Such selective factors, if they exist, could have been connected with a specific kind of mutation, thus leading to their geographic spread. Two kinds of selective factors may exist: those acting according to whether or not a mutation affects the MS1 or NBF1 domain, and those acting according to whether or not a mutation occurs on a B chromosome.

The B sequence is probably not responsible for the selection, but could be a marker in linkage disequilibrium both with the CFTR locus and another gene or DNA sequence responsible for selective effects (or meiotic drive). In this case, CF mutations could have been driven by hitchhiking, as previously suggested [23].

CF is a very peculiar disease in terms of population genetics analysis. The severity of the disease, and its frequency, have resulted in the rapid accumulation of much data, since well over one hundred laboratories perform RFLP analyses in prenatal diagnosis and identify mutations for biological and medical purposes. Within less than 5 years, since polymorphism inside and around the gene has been better elucidated than in most diseases. It is therefore to be hoped that analysis of the cellular biology and physiology of the CFTR protein, as well as the search for other genes acting on the variable expressivity of the disease, will provide answers to the questions that have long puzzled population geneticists.

Acknowledgement

This study was supported by the Association Française de Lutte contre la Mucoviscidose (AFLM).

Appendix

List of the members of the Cystic Fibrosis Genetic Analysis Consortium

Amos, Boston U, USA	Goossens, Créteil, France	Meitinger, U München, FRG
Anvret, Stockholm, Sweden	Graham, Belfast, UK	Molano, Madrid, Spain
Baranov, Leningrad, Russia	Halley, Rotterdam, The Netherlands	Morel, Lyon, France
Barton, Cambridge, UK	Harris, Oxford, UK	Morgan, McGill, Canada
Beaudet, Baylor, USA	Higgins, Birmingham, UK	Nukiwa, Tokyo, Japan
Boué, Paris, France	Highsmith, NC Memorial Hospital, USA	Ober, U Chicago, USA
Cao, U Cagliari, Italy	Hood, California Institute Technology, USA	Olek, U Bonn, FRG
Carbonara, Torino, Italy	Hortst, Münster, FRG	Orr, U Minnesota, USA
Cassiman, U Leuven, Belgium	Jaume-Roig, Son Dureta, Spain	Pignatti, U Verona, Italy
Cheadle, U Wales, UK	Jones, WGH Edinburgh, UK	Pivetta, Buenos Aires, Argentina
Claustres, Montpellier, France	Kalaydjieva, Sofia, Bulgaria	Ramsay, SAIMR, South Africa
Cochaux, Brussels, Belgium	Kant, U Pennsylvania, USA	Richards, GeneScreen, USA
Collin, U Michigan, USA	Kerem, Jerusalem, Israel	Romeo, Gaslini, Genoa, Italy
Coskun, Hacettepe U, Turkey	Kitzis, CHU Paris, France	Rowley, Rochester, USA
Coutelle, Berlin, FRG	Klinger, Integrated Genetics, USA	Rozen, Montreal, Canada
Cutting, Johns Hopkins, USA	Knight, London, UK	Scheffer, U Groningen, The Netherlands
Dallapiccola, Rome, Italy	Komel, Ljubljana, Yugoslavia	Schmidtke, Hannover, FRG
Dean, NCI Frederick, USA	Krueger, Hahnemann, USA	Schwartz, U Copenhagen, Denmark
De Arce, Dublin, Ireland	Kulozik, U Ulm, FRG	Sebastio, Naples, Italy
de la Chapelle, Helsinki, Finland	Lavinha, Lisbon, Portugal	Seltzer, U Colorado, USA
Desnick, Mount Sinai, New York, USA	Le Gall, Rennes, France	Super, Manchester, UK
Edkins, Perth, Australia	Lissens, Vrije U, Brussels, Belgium	Thibodeau, Rochester, USA
Efremov, Skopje, Yugoslavia	Loukopoulos, Athens, Greece	Traystman, U Nebraska, USA
Elles, St Mary's, Manchester, UK	Lucotte, Collège de France, Paris, France	Trembath, ICH, London, UK
Erlich, Cetus, USA	Macek, Free U, Berlin, FRG	Tümmier, Hannover, FRG
Estivill, Barcelona, Spain	Malik, Basel, Switzerland	Verellen-Dumoulin, Brussels, Belgium
Ferec, Brest, France	Mao, Collaborative Research, USA	Willems, Antwerp, Belgium
Ferrari, Milano, Italy	Mathew, Guy's, London, UK	Williamson, St Mary's, London, UK
George, Christchurch, New Zealand	Mazurczak, Warsaw, Poland	
Gerard, Harvard, USA		
Gilbert, Cornell, New York, USA		
Godet, Villeurbanne, France		

References

- 1 Riordan JR, Rommens JM, Kerem BS, Alon N, Rozmahel R, Grzelczak Z, Zilienski J, Lok S, Plasvic N, Chou JL, Drumm ML, Iannuzzi MC, Collins FS, Tsui LC: Identification of the cystic fibrosis: Cloning and characterization of complementary DNA. *Science* 1989;245:1066-1072.
- 2 Schwartz M, Johansen HK, Koch C, Brandt NJ: Frequency of the $\Delta F508$ mutation on cystic fibrosis chromosomes in Denmark. *Hum Genet* 1990;85: 427-428.
- 3 Beaudet A: Invited editorial: Carrier screening for cystic fibrosis. *Am J Hum Genet* 1990;47:603-605.
- 4 Ferec C, Audrezet MP, Mercier B, Guillermit H, Moulrier P, Quere I, Verlingue C: Detection of over 98% cystic fibrosis mutations in a Celtic population. *Nature Genet* 1992;1: 188-191.
- 5 Cutting GR, Curristin SM, Nash E, Rosenstein BJ, Lerer I, Abeliovich D, Hill A, Graham C: Analysis of four diverse population groups indicates that a subset of cystic fibrosis mutations occur in common among Caucasians. *Am J Hum Genet* 1992; 50:1185-1194.
- 6 Wright SW, Morton NE: Genetic studies on cystic fibrosis in Hawaii. *Am J Hum Genet* 1968;20:157-169.
- 7 Wagener D, Cavalli-Sforza LL, Barakat R: Ethnic variation of genetic disease: Roles of drift for recessive lethal genes. *Am J Hum Genet* 1978; 30:262-270.
- 8 Jorde LB, Lathrop GM: A test of the heterozygote advantage hypothesis in cystic fibrosis. *Am J Hum Genet* 1988;42:808-815.
- 9 Danks DM, Allan J, Anderson CM: A genetic study of fibrocystic disease of the pancreas. *Ann Hum Genet* 1965;28:323-356.
- 10 Anderson CM, Allan J, Johansen PG: Comments on the possible existence and nature of a heterozygote advantage in cystic fibrosis; in Hottinger A, Berger H (eds): *Cystic Fibrosis*. Basel, Karger, 1967, vol 10, pp 381-387.
- 11 Knudson AG, Wayne I, Hallett Y: On the selective advantage of cystic fibrosis heterozygotes. *Am J Hum Genet* 1967;19:388-392.
- 12 Romeo G, Devoto M, Galieta LJV: Why is the cystic fibrosis gene so frequent? *Hum Genet* 1989;84:1-5.
- 13 Serre JL, Simon-Bouy B, Mornet E, Jaume-Roig B, Balassopoulou A, Schwartz M, Taillandier A, Boué J, Boué A: Studies of RFLP closely linked to the cystic fibrosis locus throughout Europe lead to new considerations in population genetics. *Hum Genet* 1990;84:449-454.
- 14 Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC: Identification of the cystic fibrosis gene: Genetic analysis. *Science* 1989;245:1073-1080.
- 15 Cystic Fibrosis Genetic Analysis Consortium: Worldwide Survey of the $\Delta F508$ mutation. *Am J Hum Genet* 1990;47:354-359.
- 16 EWCG: Gradient of distribution in Europe of the major CF mutation and of its associated haplotype. *Hum Genet* 1990;85:436-446.
- 17 Tsui LC: The spectrum of cystic fibrosis mutations. *Trends Genet* 1992;8:392-398.
- 18 Zielenski J, Rozmahel R, Bozon D, Kerem B, Grzelczak Z, Riordan JR, Rommens J, Tsui LC: Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics* 1991; 10:214-228.
- 19 Estivill X, Farrall M, Scambler PJ, Bell GM, Hawley KMF, Lench NJ, Bates GP, Kruyer HC, Frederick PA, Stanier P, Watson EK, Williamson R, Wainwright B: A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. *Nature* 1987;326:840-845.
- 20 Muller B, Clerget-Darpoux F: A test based on the exact probability distribution of the χ^2 statistic incorporation into the MASC method. *Ann Hum Genet* 1991;55:69-75.
- 21 Kerem B, Zielenski J, Markiewicz D, Bozon D, Gazit E, Yahav J, Kennedy D, Riordan JR, Collins FS, Rommens J, Tsui LC: Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds in the cystic fibrosis gene. *Proc Natl Acad Sci USA* 1990;87:8447-8451.
- 22 Rich Devra P: Studies on the structure and function of CFTR. Philippe Laudat Conference (INSERM), Strasbourg-Le Bichenberg, Sep 1992.
- 23 Wagener D, Cavalli-Sforza LL: Ethnic variation in genetic disease: Possible roles of hitchhiking and epistasis. *Am J Hum Genet* 1975;27:348-364.