

ORIGINAL INVESTIGATION

Hendrik G. de Vries · Martin A. van der Meulen
Rima Rozen · Dickie J. J. Halley · Hans Scheffer
Leo P. ten Kate · Charles H. C. M. Buys
Gerard J. te Meerman

Haplotype identity between individuals who share a CFTR mutation allele “identical by descent”: demonstration of the usefulness of the haplotype-sharing concept for gene mapping in real populations

Received: 30 November 1995 / Revised: 11 April 1996

Abstract Cystic fibrosis (CF) patients with the A455E mutation, in both the French Canadian and the Dutch population, share a common haplotype over distances of up to 25 cM. French Canadian patients with the 621+1G→T mutation share a common haplotype of more than 14 cM. In contrast, haplotypes containing the Δ F508 mutation show haplotype identity over a much shorter genomic distance within and between populations, probably because of the multiple introduction of this most common mutation. Haplotype analysis for specific mutations in CF or in other recessive diseases can be used as a model for studying the occurrence of genetic drift conditional on gene frequencies. Moreover, from our results, it can be inferred that analysis of shared haplotypes is a suitable method for genetic mapping in general.

Introduction

In recent times, it has been repeatedly observed that haplotypes surrounding rare alleles of a gene are large (Rozen et al. 1990; Cannon-Albright et al. 1994; Gruis et al. 1994; Nystrom-Lahti et al. 1994; Houwen et al. 1994; Hastbacka et al. 1994; Sulisalo et al. 1994; Meyers et al. 1994; Heyer and Tremblay 1995). Sharing of large genomic ar-

eas can be used as a method for mapping disease genes: this is termed identity by descent (IBD) mapping (Houwen et al. 1994; te Meerman et al. 1995). An empirical question is whether haplotype sharing can be observed in real populations to an extent where IBD mapping using haplotype sharing is feasible.

As an empirical model for high and low frequency alleles weakly associated with a disease or with dominance and low penetrance, we have chosen to study mutations leading to the recessive disease cystic fibrosis (CF). Only a small fraction of the disease alleles present in the population can be observed in diseased individuals. To date, more than 500 presumed mutations have been identified in the CF transmembrane conductance regulator gene. The most common mutation (Δ F508) has a high frequency in Caucasian populations (up to 1.5%). A less frequent mutation is the A455E missense mutation. According to data from the CF consortium (Cystic Fibrosis Genetic Analysis Consortium 1994), this mutation is mainly detected among French Canadian and Dutch CF patients. The A455E mutation comprises 8% of all CF mutations in the French Canadian population of the Saguenay-Lac St. Jean region of Quebec (Rozen et al. 1992) and 3% in The Netherlands (unpublished data). The overall CF carrier frequency is 1 in 15 in this particular French Canadian population (Rozen et al. 1992) and 1 in 30 in The Netherlands (ten Kate 1977). This results in allele frequencies of 1/400 (8% of 1/30) and 1/2000 (3% of 1/60), respectively. The concentrated geographical distribution indicates that this mutation has been recently introduced.

In models for multifactorial disease, the mutated gene concept is not directly applicable, since alleles act as risk factors in combination with other factors. However, if gene/gene or gene/environment interactions are modelled in founder populations, where the multifactorial background is more in common than in mixed populations, specific alleles of genes may be involved, leading to association at the population level. The implication is that such alleles show increased frequencies in affected individuals. It cannot be ruled out that multifactorial diseases are determined by only a few genes; this would make an increased

H. G. de Vries · M. A. van der Meulen · H. Scheffer ·
C. H. C. M. Buys · G. J. te Meerman (✉)
Department of Medical Genetics, University of Groningen,
Antonius Deusinglaan 4, 9713 AW Groningen, The Netherlands
Tel.: +31-50-3632925; Fax: +31-50-3632947;
email: G.J.te.meerman@med.rug.nl

R. Rozen
Departments of Human Genetics, Pediatrics and Biology,
Montreal, Canada

D. J. J. Halley
Department of Clinical Genetics, Erasmus University,
Rotterdam, The Netherlands

L. P. ten Kate
Department of Human Genetics, Free University of Amsterdam,
The Netherlands

frequency of specific alleles in affected individuals even more likely (van Ommen 1995). By studying haplotype sharing in two populations, one with a very late origin (French Canadians) and one with a longer history (southern part of The Netherlands), we demonstrate the usefulness of the haplotype-sharing concept for gene mapping in real populations.

Materials and methods

Recruitment of CF patients

The A455E mutation, which is associated with a less severe CF phenotype, has mainly been detected in Canada and in The Netherlands (Cystic Fibrosis Genetic Analysis Consortium 1994). In Canada, this mutation was introduced by French immigrants during the period 1650–1900 (Rozen et al. 1990, 1992). The Dutch patients with an A455E mutation all come from southern parts of The Netherlands (unpublished results). Blood samples from 15 independent Dutch CF patients with the A455E mutation were collected in Groningen and Rotterdam. Samples from 10 French Canadian CF patients with the A455E mutation, from a subpopulation in north-eastern Quebec, viz., the Saguenay-Lac St. Jean region (Rozen et al. 1990, 1992), were collected in Montreal. The father of proband 3 and the mother of proband 6 were carriers of the 621+1G→T mutation and were sibs. Because all patients were compound heterozygotes, haplotype sharing could also be determined around two other CF mutations ($\Delta F508$, 621+1G→T). Clinical diagnosis in all patients was confirmed by demonstrating the $\Delta F508$ (Scheffer et al. 1989), A455E (Kerem et al. 1990), and 621+1G→T (Zielenski et al. 1991a) mutations. In all cases, DNA from one of the parents was used for phase determination.

Microsatellite analysis

Three intragenic microsatellites, IVS8CA (intron 8), IVS17BTA, and IVS17BCA (intron 17b; Zielenski et al. 1991b) and 7 extragenic microsatellites, D7S518, D7S501, D7S523, D7S486, D7S480, D7S490, and D7S635 (Gyupay et al. 1994), were analyzed by single amplification. The primer sequences are shown in Table 1. The

polymerase reaction (PCR) was performed with 400 ng DNA, 50 mM KCl, 10 mM TRIS (pH 9.0), 1.5 mM $MgCl_2$, 0.01% gelatin, 0.1% Triton X-100, 0.5 μM of each microsatellite primer (biotinylated), and 0.25 U Super *Taq* (SphearoQ), in a total volume of 50 μl . Part of the PCR product was diluted to 1:20 and mixed with a formamide solution. Absolute PCR product sizes were determined by adding standard size markers (generated from Pharmacia m13mp18) to each sample. The reaction products were detected in a 6% denaturing polyacrylamide gel with an Automated Laser Fluorescent system (Pharmacia LKB). In total, a region of 25 cM flanking the CF gene was genotyped. For certain microsatellite alleles, it was not possible to establish phase. In these cases, haplotypes were assigned by comparing alleles of flanking microsatellites between the CF patients.

Results

The results of the haplotype analysis of the Canadian and Dutch CF patients with a A455E mutation are shown in Table 2. The haplotype of the intragenic markers on the A455E chromosome of the 10 French Canadian patients is identical in all cases, namely 22–35–13, in which the numbers represent the numbers of repeats of IVS8CA, IVS17BTA, and IVS17BCA, respectively. The extragenic alleles are numbered according to the relative length of the PCR product. Of the 10 Canadian patients, 9 share a region of at least 5 cM surrounding the A455E mutation (Fig. 1). Patient nos. 1, 2, and 3 share a region of more than 25 cM. This is also the case, with a different haplotype, for patient nos. 4, 5, and 6. The haplotype of the intragenic markers of 14 of the 15 Dutch patients are identical to the Canadian patients. They share a region of more than 4 cM surrounding A455E (Fig. 2). The remaining patient shows a different intragenic haplotype (22–37–13).

The haplotypes of the non-A455E chromosomes of the Canadian and Dutch patients are shown in Table 3. Four Canadian patients have a 621+1G→T mutation on the other CF chromosome, and six have a $\Delta F508$ mutation.

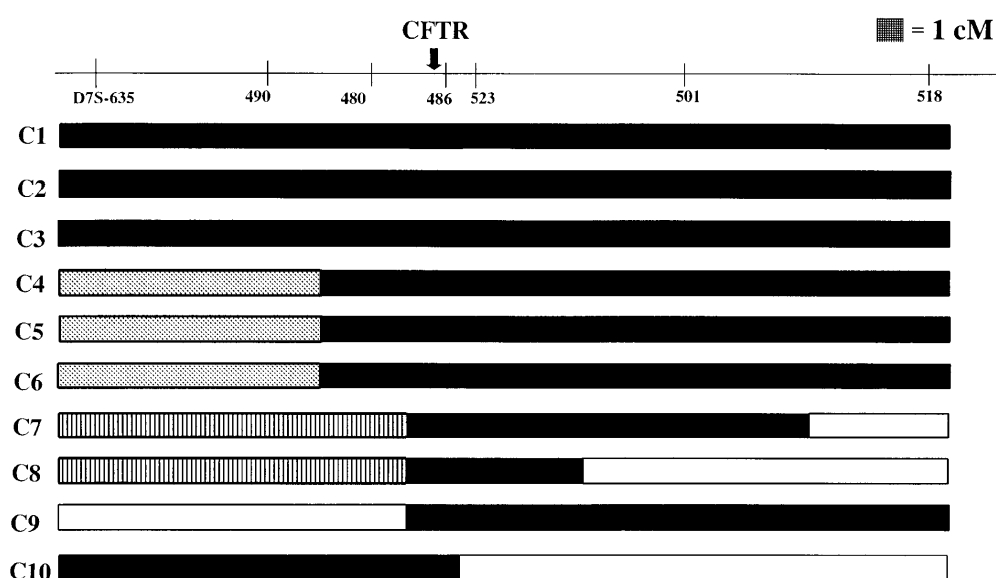
Table 1 PCR primers of microsatellites at and flanking the CF locus. Data have been derived from Zielenski et al. (1991b) and Gyupay et al. (1994)

Marker	Primer sequences 5'→3'	Heterozygosity	Distance to CFTR gene (cM)
D7S518	CAGTAGGCAGGGGTGG GGGTGTGTCTGTGTGACAAC	0.87	15
D7S501	CACCGTTGTGATGGCAGAG ATTTCTTACCAGGCAGACTGCT	0.81	7
D7S523	CTGATTCATAGCAGCACTTG AAAACATTTCCATTACCACTG	0.80	1
D7S486	AAAGGCCAATGGTATATCCC GCCAGGTGATTGATAGTGC	0.81	0.2
IVS8BTA	TCTATCTCATGTAAATGCTG GTTTCTAGAGGACATGATC	0.41	0
IVS17BTA	GACAATCTGTGTGCATCG GCTCGATTCTATAGGTTATC	0.89	0
IVS17BCA	AAACTTACCGACAAGAGGA TGTCACCTCTTCATACTCAT	0.38	0
D7S480	CTTGGGGACTGAACCATCTT AGCTACCATAGGGCTGGAGG	0.86	2
D7S490	CCTTGGGCCAATAAGGTAAG AGCTACTTGCAAGTGAACAGCATTT	0.78	5
D7S635	CCAGGCCATGTGGAAC AGTTCTTGGCTTGCGTCAGT	0.81	10

Table 2 Microsatellite haplotypes for independent A455E chromosomes from a French Canadian and a Dutch population (! phase unknown)

	5 ^	3 ^	2 ^	Intragenic			1 ^	6 ^	8 ^	(cM)
	D7S-635	490	480	8CA	17BTA	17BCA	486	523	501	518
<i>French Canadian patients</i>										
1	7!	1	3	22!	35!	13	8	4	5	10
2	7	1	3	22	35!	13	8	4!	5	10
3	7	1	3	22	35	13	8	4	5!	10
4	3!	4	3	22!	35	13	8	4	5	10
5	3	4	3	22	35	13	8	4	5	10
6	3	4	3!	22	35	13	8	4	5!	10
7	1!	1	2	22!	35	13	8	4	5	2
8	1	1	2	22	35	13	8	4!	4	10
9	4!	2	2	22	35	13	8	4	5	10
10	7	1!	3	22	35	13	8	3	3	2
<i>Dutch patients</i>										
1	7	1	3	22	35!	13	8	4	3	2
2	1	1	3	22!	35	13	8	4	1	12
3	1	4	6	22	35	13	8	4	1!	4!
4	6	1	3	22	35	13	8	5	7	9
5	6	1	3	22	35	13	8	5	7	9
6	6	1	3	22!	35	13	8	5	5	2
7	6	1!	3	22	35	13	10	5	5	2
8	1	1	3!	22	35	13	8	5!	3	11
9	4	1	3	22	37	13	8	5	3	10
10	3	1	3	22!	35	13	8	3	3	12!
11	5	6	3	22	35	13	8	5	6	2
12	8	2	3	22!	35	13	8	5	7	2
13	5!	6	3	22	35	13	8	3	7!	10
14	10	2	3	22	35	13	8!	2	3	3
15	1	4	5	22	35!	13	8	5	7	3

Fig. 1 DNA sharing around the CF gene of 10 unrelated French Canadian CF patients (C1–10) with the A455E mutation. DNA regions with the same grade of shading represent identical haplotypes. White regions represent unique haplotypes within this population



Out of the 15 Dutch CF patients, 13 have a $\Delta F508$ mutation on the other chromosome. The two remaining Dutch patients, nos. 2 and 10, have the 1717–1G→A and the R553X mutation, respectively, on their other chromosome. The French Canadian patients with the 621+1G→T mutation have identical intragenic haplotypes (21–31–13)

and share a DNA region of more than 14 cM. Little or no sharing can be observed when the $\Delta F508$ chromosomes in and between the two populations are compared.

Fig. 2 DNA sharing around the CF gene of 15 unrelated Dutch CF patients (D1–15) with the A455E mutation. DNA regions with the same grade of *shading* represent identical haplotypes. *White regions* represent unique haplotypes within this population

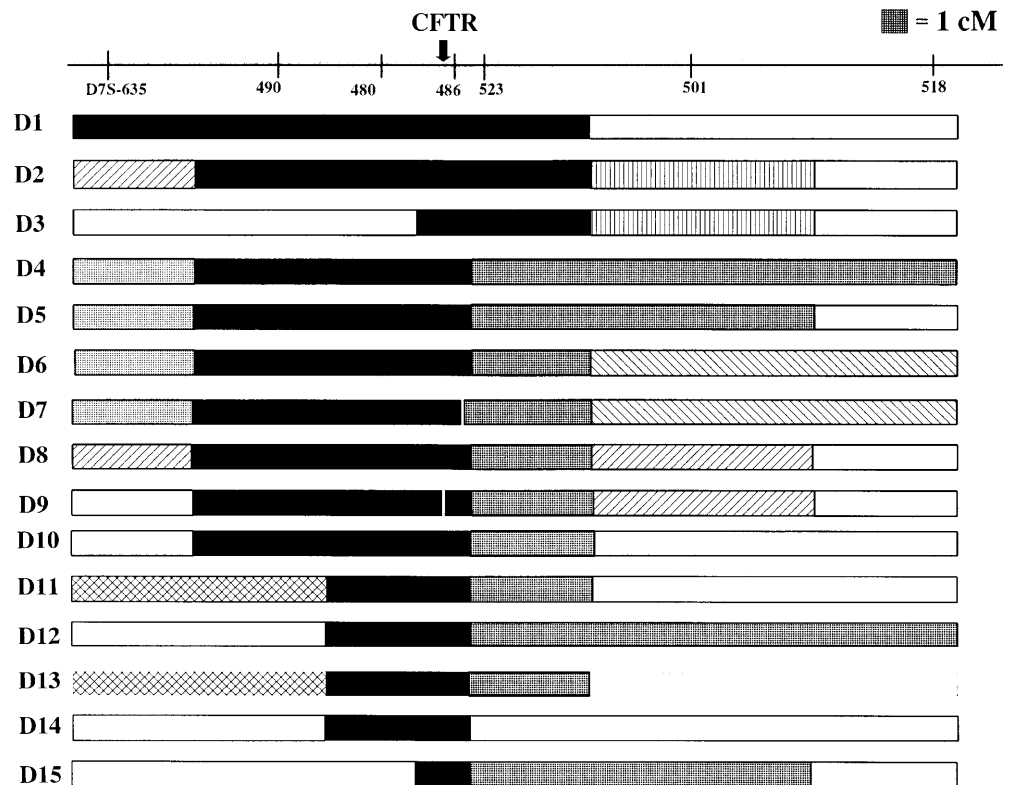


Table 3 Microsatellite haplotypes for independent $\Delta F508$ and 621+1G→T chromosomes from a French Canadian and a Dutch population (! phase unknown)

	5 ^	3 ^	2 ^	Intragenic			1 ^	6 ^	8 (cM)	
	D7S-635	490	480	8CA	17BTA	17BCA	486	523	501	518
<i>French Canadian patients</i>										
2 621+1G→T	1	4	5	21	31!	13	8	5!	7	3
3 621+1G→T	1	4	5	21	31	13	8	5	1!	6
6 621+1G→T	1	4	5!	21	31	13	8	5	7!	5
9 621+1G→T	1!	4	5	21	31	13	8	5	5	2
1 $\Delta F508$	9!	3	1	17!	32!	13	8	7	1	3
7 $\Delta F508$	9!	3	1	17!	32	13	8	1	5	2
4 $\Delta F508$	7!	4	1	17!	32	13	8	7	1	4
5 $\Delta F508$	9	6	7	17!	32	13	5	5	7	5
8 $\Delta F508$	2	4	3	17	35	17	3	5!	8	7
10 $\Delta F508$	8	4!	2	23	31	13	11	5	3	4
<i>Dutch patients</i>										
5 $\Delta F508$	6	1	6	17	31	13	10	4	3	2
7 $\Delta F508$	1	6!	8	17	31	13	10	6	5	9
8 $\Delta F508$	4	2	7!	17	31	13	10	3!	6	11
15 $\Delta F508$	1	6	5	17	31!	13	10	2	7	4
1 $\Delta F508$	4	1	3	17	31!	13	3	3	4	2
9 $\Delta F508$	8	6	3	17	31	13	8	4	1	7
6 $\Delta F508$	7	3	3	17!	32	13	8	2	1	2
11 $\Delta F508$	1	4	7	17	32	13	8	6	5	9
14 $\Delta F508$	1	2	1	17	32	13	5!	3	7	7
4 $\Delta F508$	3	3	2	17	35	17	10	5	4	2
12 $\Delta F508$	8	1	2	17!	35	17	10	2	7	4
13 $\Delta F508$	4!	1	5	17	47	13	10	4	8!	3
3 $\Delta F508$	1	2	1	23	31	13	8	5	5!	10!
10 1717 1G→A	1	1	3	17!	35	16!	8	3	7	6!
2 R553X	1	1	6	11	55	17!	10	5	8	7

Discussion

The Saguenay-Lac St.Jean region was settled by approximately 5000 immigrant families during 1838–1911 (Rozzen et al. 1990). The identical haplotype of the intragenic markers and the sharing of large DNA regions on the A455E chromosomes indicate that the A455E mutations that we now observe result from a single introduction into this population. The three intragenic markers in the Dutch population show the same haplotype. Furthermore, some of the Canadian and Dutch patients share a region of more than 4 cM around the CF gene. This indicates that the A455E mutation has been inherited from a not-too-distant predecessor. In the same French Canadian population, extensive sharing of haplotypes surrounding the mutated myotonic dystrophin and rickettsia genes has been demonstrated (Bétard et al. 1995). Dutch patient no. 9 shows the only aberrant haplotype (22–37–13). This can best be explained by a mutation in the original haplotype, because the neighboring markers in this patient are comparable with almost all the other Dutch patients.

The $\Delta F508$ mutation is a very old mutation that was introduced at least 52 000 years ago into Europe (Morral et al. 1994). Much variation is observed in intragenic markers; this is probably because of repeat length mutations subsequent to the $\Delta F508$ mutation. This most frequent CF mutation has been widely distributed, especially in Caucasian populations, over a long period of time. In the Saguenay-Lac St.Jean region, we have detected 3 different intragenic haplotypes in only 6 patients with a $\Delta F508$ mutation. This means that there were probably multiple independent introductions of the $\Delta F508$ mutations into this population. Sharing of DNA regions surrounding $\Delta F508$ mutations is limited to individuals with the same haplotype of intragenic markers. The observed intragenic haplotypes are also the most common haplotypes in Europe (Morral et al. 1994). To detect haplotype sharing surrounding $\Delta F508$ mutations, more observations with identical intragenic haplotypes are necessary.

If the size of a shared haplotype is large, the common predecessor must be recent. The difference in the extent of sharing found for the $\Delta F508$ mutation and for the other mutations shows that haplotype sharing can be expected to occur only if a limited number of independent introductions of a gene has been made into the population. Haplotype sharing and other association studies can therefore be expected to be successful in founder populations of a specific size relative to the gene frequency. Because of genetic drift, the independent introduction of 10–20 copies will lead to perhaps 1–4 copies remaining in widely varying numbers of individuals after 8 or more generations. This is in agreement with the theoretical expectation according to Fisher (Vogel and Motulski 1986), who computed that the probability of survival of a single gene in a stable population after 10 generations is about 10%. It appears from the French Canadian genealogical data that gene flow cannot be traced back to unique predecessors, even if all genealogical data are complete. The

informativity of genealogical data therefore seems limited relative to the information being learned from direct haplotype comparison.

The similarity in intragenic $\Delta F508$ haplotypes in the two populations suggests a multiple introduction of the same intragenic haplotypes, of which some have survived. In Finland, genetic drift has led, by this process, to a low CF frequency and to a distribution of mutations different from that of the rest of Europe (de la Chapelle 1993).

The suitability of a population for haplotype-sharing studies will increase if there is a larger population with which the founder population is connected. Once a genomic region has been identified that possibly contains an allele predisposing to a disease, it is easy to perform a genome screening in affected individuals at the 1 cM level. This situation seems to exist both for the French Canadians and in The Netherlands, where more isolated areas are connected with larger populations. At genomic distances of 1 cM, association and linkage disequilibrium can be detected in populations much larger than the one originally investigated. At the same time, a large number of meioses is being observed implicitly, thus narrowing the region in which the gene must be located. This is well illustrated from the consistent overlap shown in and between Figs. 1 and 2. In The Netherlands, the A455E data come from an unselected population, which shows that, for rare genes such as the A455E mutation, founder effects are present at the population level. A complication in haplotype comparison is marker allele mutation. In this dataset, we probably observe two such mutations (Dutch patients nos. 7 and 9 show an allele mutation in the microsatellites D7S486 and IVS17BTA, respectively) where the haplotype for surrounding markers seems to be conserved.

The present study shows that the suitability of populations for gene mapping through direct haplotype comparison can be investigated using some of the well-known recessive disease mutations that can easily be detected at the level of carriers. These mutations generally have a high frequency, as would be expected for alleles of genes involved in multifactorial diseases. By relating the gene frequency to the size of the shared haplotypes, an impression can be obtained regarding whether haplotype-sharing analysis can be used to find the map location of other genes.

Acknowledgements The authors thank Dr. J. P. Heyerman (Leyenburg Hospital, The Hague) for sending DNA samples of the parents of Dutch CF patients. This study was made possible by a grant from The Netherlands Organization for Scientific Research (NWO).

References

- Bétard C, Raeymaekers P, Ouellette G, Jomphe M, Labuda M, Glorieux F, Mathieu J, Laberge C, Cassiman J-J, Sandkuijl LA, Gauvreau D (1995) The validation of a novel linkage disequilibrium mapping technique on Steinert myotonic dystrophy and pseudovitamin D deficient rickets in a founder population (abstract). *Med Genet* 2:238

- Cannon-Albright LA, Goldgar DE, Gruis NA, Neuhasen S, Anderson DE, Lewis CM, Jost M, Tran TD, Nyguen K (1994) Localization of the 9p melanoma susceptibility locus to a 2 cM region between D9S736 and D9S171. *Genomics* 23:265
- Chapelle A de la (1993) Disease gene mapping in isolated human populations: the example of Finland. *J Med Genet* 30:857–865
- Cystic Fibrosis Genetic Analysis Consortium (1994) Population variation of common cystic fibrosis mutations. *Hum Mutat* 4:167–177
- Gruis N, Sandkuijl LA, Bergman W, Frants RR (1994) Common 9p haplotype in Dutch FAMMM families. Genetics of the familial atypical multiple mole-melanoma syndrome. PhD thesis, Leiden
- Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994) The 1993–1994 Génethon human genetic linkage map. *Nat Genet* 7:246
- Hastbacka J, Chapelle A de la, Mahtani MM, Clines G, Reeve-Daly MP, Daly M, Hamilton BA, Kusumi K, Triveldi B, Weaver A, Coloma A, Lovett M, Buckler A, Kaitila I, Lander ES (1994) The diastrophic dysplasia gene encodes a novel sulphate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell* 78:1073–1087
- Heyer E, Tremblay M (1995) Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 56:970–978
- Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet* 8:380–386
- Kate LP ten (1977) Cystic fibrosis in The Netherlands. *Int J Epidemiol* 6:23–35
- Kerem B, Zielenski J, Markiewicz D, Bozon D, Gazit E, Yahav J, Kennedy D, Riordan JR, Collins FS, Rommens JM, Tsui L-C (1990) Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. *Proc Natl Acad Sci USA* 87:8447–8451
- Meerman GJ te, Meulen MA van der, Sandkuijl LA (1995) Perspectives of identity by descent (IBD) mapping in founder populations. *Clin Exp Allergy* 25 (Suppl 2):97–102
- Meyers DA, Postma DS, Panhuysen CIM, Xu J, Amelung PJ, Levitt RC, Bleeker ER (1994) Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. *Genomics* 23:464
- Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Giménez J, Reis A, Varon-Mateeva R, Macek Jr M, Kalaydjieva L, Angelicheva D, Dancheva R, Romeo G, Russo MP, Garnerone S, Restagno G, Ferrari M, Magnani C, Claustres M, Desgeorges M, Schwartz M, Schwartz M, Novelli G, Ferec C, Arce M de, Nemeti M, Kere J, Anvret M, Dahl N, Kadasi L (1994) The origin of the major cystic fibrosis mutation ($\Delta F508$) in European populations. *Nat Genet* 7:169–175
- Nystrom-Lahti M, Sistonen P, Mecklin J-P, Pylkkanen L, Aaltonen LA, Jarvinen H, Weissenbach J, Chapelle A de la, Peltonen P (1994) Close linkage to chromosome 3p and conservation of ancestral founding haplotypes in hereditary nonpolyposis colorectal cancer families. *Proc Natl Acad Sci USA* 91:6054–6058
- Ommen GJ van (1995) A foundation for limb-girdle muscular dystrophy. *Nat Med* 1:412–414
- Rozen R, Schwartz RH, Hilman BC, Stanislovitis P, Horn GT, Klinger K, Daigneault J, Braekeleer M de, Kerem B-S, Tsui L-C, Fujiwara TM, Morgan K (1990) Cystic fibrosis mutations in North American populations of French ancestry: analysis of Quebec French Canadian and Louisiana Acadian families. *Am J Hum Genet* 47:606–610
- Rozen R, De Braekeleer M, Daigneault J, Ferreira-Rajabi L, Gerdes M, Lamoureux L, Aubin G, Simard F, Fujiwara TM, Morgan K (1992) Cystic fibrosis mutations in French Canadians: three CFTR mutations are relatively frequent in a Quebec population with an elevated incidence of cystic fibrosis. *Am J Med Genet* 42:360–364
- Scheffer H, Verlind E, Penninga D, Meerman G te, Kate LP ten, Buys CHCM (1989) Rapid screening for $\Delta F508$ deletion in cystic fibrosis. *Lancet* II:1345–1346
- Sulisalo T, Francomano CA, Sistonen P, Maher JF, McKusik VA, Chapelle A de la, Kaitila I (1994) High-resolution genetic mapping of the cartilage-hair hypoplasia (CHH) gene in Amish and Finnish families. *Genomics* 20:347–353
- Vogel F, Motulski AG (1986) Human genetics, 3rd edn. Springer, Berlin Heidelberg New York
- Zielenski J, Bozon D, Kerem B, Markiewicz D, Durie P, Rommens JM, Tsui L-C (1991a) Identification of mutations in exons 1 through 8 of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics* 10:229–235
- Zielenski J, Markiewicz D, Rininsland F, Rommens JM, Tsui L-C (1991b) A cluster of highly polymorphic dinucleotide repeats in intron 17b of the CFTR gene. *Am J Hum Genet* 49:1256–1262