The search for the cystic fibrosis gene

MANUEL BUCHWALD, LAP-CHEE TSUI, AND JOHN R. RIORDAN
Research Institute, The Hospital for Sick Children and Departments of Medical Genetics, Medical Biophysics, Biochemistry, and Clinical Biochemistry, University of Toronto, Toronto, Ontario M5G 1X8, Canada

BUCHWALD, MANUEL, LAP-CHEE TSUI, AND JOHN R. RIORDAN. *The search for the cystic fibrosis gene.* Am. J. Physiol 257 (Lung Cell. Mol. Physiol. 1): L47–L52, 1989.—This article reviews progress made to date in the attempt to find the gene causing cystic fibrosis (CF). The search has been focused on a region of the human genome containing no more than 10 genes and has been accomplished in two steps. First, linkage analysis was used to localize the gene to the q31 band of chromosome 7. Subsequently, molecular cloning methods were employed to isolate DNA segments estimated to be within 400,000 base pairs of the CF gene. The current challenge in the research is to find the genes located in this interval and to then prove that one of them is the CF gene. The approaches by which this will be accomplished are discussed.

linkage analysis; molecular genetics; gene cloning; deoxyribonucleic acid diagnosis

CYSTIC FIBROSIS (CF) is the most common severe autosomal recessive disorder of the Caucasian population with an estimated incidence of 1/2,000 and a heterozygote (carrier) frequency of 1/20–1/25 (56, 58, 68). However, its frequency in Oriental and black populations is considerably lower (11, 69). CF usually presents in childhood as failure to thrive in combination with recurrent pulmonary infections and/or maldigestion, but a subset of patients have a milder phenotype and may not be identified until later in life (57, 68). Laboratory confirmation of the diagnosis of CF is obtained by the sweat test, in which patients are made to sweat by the local injection of pilocarpine and the concentration of Cl^- is then measured. All CF patients have sweat chloride values higher than 60 mmol/l (many with values over 100 mmol/l), whereas most normal individuals have values below 50 mmol/l (57).

The mean age of patients has increased dramatically over the past two decades from ~10 yr in the 1960s to ~25 yr today in major centers with large CF clinics and the appropriate medical support infrastructure (12, 62). The improvements in mortality and morbidity have been accomplished by aggressive symptomatic treatment of the major medical problems faced by CF patients: chronic obstructive lung disease and pancreatic insufficiency. However, for the past decade no significant new therapies that prolong the life of the average CF patients have been introduced, and it is possible that we may have reached the limit of medical treatment in the absence of knowledge of the basic defect in the disorder.

The autosomal recessive mode of inheritance of CF was recognized early in its history, and this genetic etiology pointed to the existence of a gene that when defective caused the disease. The idea that specific changes in the genetic makeup of individuals results in clinical and metabolic disorders can be traced back to the work of Garrod in the early part of this century who coined the expression "inborn errors of metabolism." Since then, more than 3,000 such disorders have been catalogued, and the nature of the basic defect in ~300 has been identified through biochemical and physiological studies on patient samples or their cultured cells. The development of techniques of molecular genetics has led to the isolation of the gene coding for the defective proteins in many of these inborn errors of metabolism (34). Cloning of disease gene has allowed a detailed examination of the type and number of mutations prevalent in specific populations and the correlation between the mutations, the function of the mutant protein, and their clinical impact (8, 66).

Attempts to identify the basic defect in CF through biochemical and physiological studies of patients, their tissues, and cells have covered various areas of cellular function (see Refs. 15 57, 68). However, only recently have such studies indicated a route to the CF gene. Electrophysiological studies of isolated sweat glands and of nasal epithelium in situ showed a decreased Cl^- permeability in the epithelium lining these tissues in CF patients (30, 41). Subsequent results confirmed the observations in isolated epithelia and in cells that had been transplanted into heterologous hosts or that had been cultured in vitro, thus suggesting that the defect resided in the epithelial cells themselves and was not due to some humoral factor circulating in CF patients (see Refs. 42, 43, 64). More recently, it has been shown that the abnormality in CF epithelial cells involves the regulation of Cl^- channels in the apical membrane by processes mediated by an adenosine 3′,5′-cyclic monophosphate (cAMP)-dependent protein kinase. The precise details of the regulation of these channels are not known, but evidence is consistent with the defect being due to a membrane or membrane-bound molecule that could be the Cl^- channel itself or a molecule that regulates its function (25, 33, 50, 65). If the current attempts to clone the Cl^- channel gene are successful, it will be possible to test the hypothesis that the channel is the product of the CF gene.

FINDING THE CF GENE

Use of Linkage Analysis to Map the CF Gene

For disorders in which the biochemical defect has not been defined, such as CF, molecular genetics allows the

isolation of the gene without any prior knowledge of the nature of the gene product, the so-called "reverse genetics" approach. This strategy has resulted in the identification of the genes coding for the defective product in Duchenne muscular dystrophy, chronic granulomatous disease, and retinoblastoma (reviewed in Ref. 38). Subsequent analysis of the protein coded for by these genes has led to insights not only into the pathology of the diseases but also into new areas of cellular function and regulation that were previously unknown. It is expected that if this approach can be successfully applied in CF, similar novel insights may be gained into the functioning of epithelial cells and may lead to improved diagnosis and therapy.

In deploying the reverse genetics strategy, the search for the unknown gene is performed in two steps. In the first, linkage analysis is used to map the disease gene by showing that it cosegregates with a genetic marker of known location. A more refined localization can be achieved by studying somatic cell hybrids that contain various portions of the particular chromosome to which the gene has been mapped. Molecular cloning techniques are subsequently used to isolate genes in the region and to show that one of them is the disease gene.

The use of linkage analysis to study CF antedates the development of molecular genetics. In the 1950s and 1970s the inheritance of serum and red cell proteins and antigens was studied in CF families, but no association with the disease was found (27, 54, 55). Since these markers covered only a small portion of the human genome, such a result was not surprising. In the early 1980s a seminal paper on the use of DNA markers that could be detected after treatment of genomic DNA with restriction enzymes [restriction fragment length polymorphisms (RFLPs)] (6) revived interest in CF linkage studies. Several groups proceeded to collect families with two or more affected patients and used DNA obtained from the family members to analyze the transmission of DNA markers. The probability that a marker was located near the CF locus was determined by the use of statistical techniques [LOD score method (36)]. If no linkage was found one could also eliminate a portion of the region surrounding the DNA marker as the site of the CF gene. The major theoretical obstacle to the application of the method in CF was the possibility that CF might not be caused by a single defective gene. Instead, it was possible that two (or more) genes in the human genome could each cause CF when an individual was homozygous for a defective version of one of them. Evidence of genetic heterogeneity was weak in the early part of this decade when linkage studies were initiated (13), but since then both population studies and the linkage studies themselves have confirmed that genetic heterogeneity is unlikely in CF (14, 44).

At the Eighth Human Gene Mapping Meeting (Helsinki, August 1985) a workshop on the mapping of the CF gene indicated that approximately one-half of the genome had been excluded by the various linkage studies. At the same time a report indicated that CF was linked to the paraoxonase enzyme activity polymorphism (PON) with a recombination fraction of 10% (20, 49). This result suggested that genetic heterogeneity was not prevalent in CF and that it was possible to use linkage analysis to localize the CF gene. However, since little was known about the molecular basis of the paraoxonase protein or gene, it was not possible to proceed further from these data.

Linkage between CF and an anonymous DNA marker (D0CRI-917) at a distance of 15% recombination as well as linkage between 917 and PON was reported later in 1985 (59). Thus, a linkage group 917-PON-CF was identified and this confirmed the existence of the CF locus. The DNA marker 917 was mapped to the long arm of chromosome 7 by the use of somatic cell hybrids and its locus named D7S15 (31). Within a short time two other DNA segments on chromosome 7 were found closer to CF than either of the above two markers. The oncogene *met* and the anonymous DNA marker pJ3-11 (D7S8) were found to be ~1% recombination from the CF gene (61, 67).

A large collaborative study from seven laboratories confirmed the assignment of CF to chromosome 7 by the analysis of more than 200 families with two or more affected individuals using six different chromosome 7 markers (2). Current evidence places the order of the closest markers as: centromere-COL1A2-D7S15-PON-*met*-CF-D7S8-TCRB-telomere on the long arm of chromosome 7 (32). On the basis of deletions in some of the hybrids used to map D7S15, the known locations of the COL1A2 and TCRB genes, and in situ hybridization with *met* and D7S8, the location of the *met*-CF-D7S8 cluster was identified at 7q22-7q31 (16, 53, 61, 70). However, more recent studies using new markers suggest that the cluster is located at 7q31-7q32 (19, 45).

Cloning the CF Gene

The distance between the closest flanking markers, *met* and D7S8, was estimated to be 1–5 cM (1 cM = 1% recombination) on the basis of recombination data (2, 32). Subsequent physical mapping of the region using pulsed-field gel electrophoresis (7) is consistent with a distance of $1.5 \pm 0.5 \times 10^6$ base pairs (bp) (18, 40), placing an upper limit to the amount of DNA that has to be cloned to include the CF gene. This stretch of DNA is likely to contain between 10 and 40 genes. Three different strategies have been applied to the cloning of the CF gene.

Gene transfer. Williamson and colleagues attempted to clone the CF gene by utilizing the cellular transforming properties of the *met* oncogene. They reasoned that if DNA from a cell line containing the activated form of *met* were introduced into an untransformed cell, subsequent transformed clones that were isolated would have to contain the *met* gene and surrounding DNA. If some of the clones contained both *met* and D7S8, then they should also contain the CF gene if no DNA rearrangements had occurred. Scambler et al. (48) isolated such cell lines and found that they contained between 3 and 13 Mb (1 Mb = 10^6 bp) of human DNA. Clones containing human DNA were identified from cosmid libraries prepared from the hybrid cell lines, and genes within those clones were identified through the detection

of regions rich in G and C residues (23). Such regions [CpG islands (5)] have been found near genes and are characterized by the presence of restriction sites for enzymes that cut G/C-rich regions (e.g., NotI, BssHII).

One clone containing a CpG island was found to code for a gene that was thought to be a candidate CF gene (21). The gene was located between *met* and D7S8 by pulsed-field gel electrophoresis mapping and by an increase in linkage disequilibrium. The latter is a measure of the association of the alleles of the linked markers to normal and CF alleles. It is expected that the linkage disequilibrium value will reach a maximum at the location of the CF gene. Although the identification of this gene aroused hope that the CF gene had been identified, subsequent analyses did not support that view. The tissue distribution of the transcripts from the gene, the absence of any differences in the sequence of the normal and CF genes, and the observation of recombination between the candidate gene and CF were evidence against this gene (now called *irp*-mouse int-1 related protein) being the CF gene (4, 24, 62). Notwithstanding this disappointment, the DNA probes generated from the gene and its surrounding region (PT-3, CS7, XV2c, and KM19, collectively known as the D7S23 locus) have found great utility for DNA diagnosis of CF, since they are much closer to CF than *met* and D7S8 and provide more accurate information (see below).

Saturation mapping. Tsui and colleagues (46, 60) identified more than 250 single copy clones isolated from a chromosome 7 library and mapped them to various regions of chromosome 7 by the use of somatic cell lines containing portions of the chromosome. The rationale of this probe hunt was that if a large enough number were isolated, some of these should map between *met* and D7S8 and thus lie closer to the CF gene. Two clones were identified in this manner by the absence of recombination in informative families. The two DNA segments were found to be 10 kb apart and were located ~500 kb from the *met* locus by the use of pulsed-field mapping (46). They could thus be used as a starting point for a search for genes located between *met* and D7S8.

Walking and jumping. Walking is the term used for the procedure employed to clone contiguous segments of DNA. Single-copy segments at the ends of clones are used to identify other clones that overlap the original one. An end from an overlapping clone showing the least overlap is again chosen and used in the same manner. The progress of the walk can be monitored by detecting the presence of restriction sites used in pulsed-field gel electrophoresis mapping in the cloned segments and by measuring the degree of linkage disequilibrium between polymorphic markers and the disease gene.

Chromosome jumping (or hopping) allows the cloning of a DNA segment that is separated from the starting point by 50–200 kb (39), thus bypassing unclonable regions of the genome that can interrupt walking. In this procedure, however, the intervening DNA is not cloned. Thus, after a jump has been completed, it may be necessary to walk back toward the starting point to identify all genes present in that stretch of DNA. Probes near the CF gene have been used for chromosome jumping (or hopping). Collins and co-workers (9) used *met* as a starting point and were able to isolate a clone (CF63) that is located ~50 kb in the direction of D7S8. Further jumps from this point were abandoned after the identification of *irp,* since this DNA segment is ~600 kb from *met* in the direction of D7S8. Jumps have also been performed using D7S8 as a starting point (28).

The two DNA segments between *met* and D7S8 isolated by the saturation mapping technique (D7S122 and D7S140) are currently being used for a combined walking and jumping approach (45). Approximately 250 kb of genomic DNA have been cloned that include the region of the *irp* gene. All of the above results in combination suggest that, at most, 300–400 kb of genomic DNA between *met* and D7S8 remain to be cloned at this point. A total of 3–10 genes may be situated in this interval.

Identification of the CF Gene

The current challenge is to analyze the cloned DNA to detect genes and to subsequently determine which one is the CF gene. Detection of genes in mammalian genomic DNA is complicated by two factors. Eukaryotic genes are broken into coding (exons) and noncoding (introns) segments, and some exons can be small, making them difficult to detect. Second, mammalian genomic DNA contains a variety of repetitive DNA that is represented to varying degrees throughout the genome (52). Thus to identify a gene within a segment of genomic DNA requires that it be free of repetitive elements and that it include an exon large enough to yield a visible and reproducible signal. In practice genomic fragments free of repetitive elements are used to analyze RNA to detect a transcript (in Northern blots). Alternatively, DNA from other species can be analyzed (Zoo blots) (35). This latter strategy is based on the hypothesis that exon sequences are more likely to have been conserved throughout evolution than those of introns. A positive signal detected in a Zoo blot would be suggestive evidence that a coding portion of a gene is contained within the DNA segment being tested.

If positive signals are obtained with either of the two methods the fragment is used to screen cDNA libraries to isolate a portion of the transcript. This complementary DNA (cDNA) clone is then used as a starting point to isolate the complete gene. As genes in the region between *met* and D7S8 are isolated, each is analyzed to determine whether it is the CF gene. Evidence in favor of a candidate gene being the CF gene may be obtained from comparisons of the sequence of the cDNA derived from a control and a CF sample. A single base pair difference could be due to the presence of an RFLP rather than the mutation, so this evidence need not be conclusive. Once a specific change in sequence is observed, oligonucleotides surrounding the altered site can be synthesized and used to analyze a large set of genomic DNAs or RNAs from CF patients using the polymerase chain-reaction method (47). These results may provide stronger evidence that the CF gene has been identified.

Supportive evidence may be obtained from detection of the corresponding mRNA in tissues that are expected to be affected in the disease, e.g., pancreas, lung, sweat

gland. Direct analysis of the amounts and size of the mRNA in the affected tissues obtained from patients may show specific differences and these would also support the identification. Antibodies to a polypeptide produced by fusing the cDNA to an expression vector or to a polypeptide derived from translation of the cDNA sequence can be used to analyze control and CF extracts derived from the appropriate tissues. Differences in the location or amount of the corresponding protein can also be taken as evidence in favor of the identification of the CF gene. Finally, one can attempt to study the function of the candidate cDNA to determine whether the normal version can correct the defect in a CF cell. In this latter experiment a candidate cDNA (or the whole gene if it is small enough) is introduced into CF cells by any one of several DNA-mediated gene-transfer methods, and the physiological defect characteristic of CF cells is measured. Given appropriate expression of the introduced gene and sufficient gene product production to overcome the effect of the resident CF gene product, one may be able to observe correction of the CF phenotype.

IMPACT OF THE RESULTS TO DATE

Although the CF gene has not yet been identified, DNA-based diagnosis is now possible in affected families with a much higher precision than previously available methods. In addition, some insight has been gained into the possible number of CF mutant alleles, the relationship of these alleles to different clinical forms of CF, and their ethnic distribution.

Before the identification of linked DNA markers, genetic diagnosis in CF was not possible, and counseling to couples with an affected child and their immediate family was based on probability estimates inferred from the autosomal recessive inheritance of CF (58). The availability of closely linked DNA markers now allows the unequivocal determination of the genetic status of members of nuclear families in which one affected individual has been examined. The observation that the distribution of alleles of closely linked DNA markers is different in chromosomes bearing the CF gene and those bearing the normal gene (linkage disequilibrium) also has implications for DNA-based diagnosis. For example, individuals marrying into a family where the CF gene is segregating can be provided with an estimate of their chances of being a heterozygote that is more precise than the population risk (1/20–1/25 in Caucasians) (1, 3).

Although the diagnostic criteria for CF are definitive, significant clinical heterogeneity does exist (51). One of the questions that can be answered, either by the use of the closely linked markers or the gene itself, is whether such heterogeneity is genetically based. For example, continued presence of sufficient pancreatic function in CF patients to not require the use of pancreatic supplements appears to be a heritable trait (10). This hypothesis has received confirmation from studies of allele association with closely linked DNA markers. Thus CF patients with pancreatic sufficiency show a different pattern of allele distribution than those with pancreatic insufficiency, and this difference is best explained by postulating the presence of at least two different mutant alleles in the Caucasian population (17, 29).

Similarly, different patterns of allele association in various ethnic populations suggest that there may be several mutant alleles of the CF gene (11, 22, 26, 37). Both of the above sets of observations will be more precisely delineated once the CF gene is identified. It should then be possible to correlate particular sets of alleles of the closely linked DNA markers with specific mutations in the CF gene. If this turns out to be the case it may be possible to understand the evolutionary history of the various CF mutations and their common or independent origin in various ethnic groups. In addition, it may be possible to show that a given mutant allele leads to a specific clinical outcome, thus improving our ability to provide a prognosis to newly identified patients and perhaps also to indicate the most appropriate therapy.

We thank our various colleagues for their continued dedication to the search for the CF gene.

The research from our laboratories has been supported by a Research Development Programme grant from the Canadian and US Cystic Fibrosis Foundations and by grants from the National Institutes of Health, the Sellers Fund of the Hospital for Sick Children, and other private donations. L.-C. Tsui is a Scholar of the Canadian Cystic Fibrosis Foundation.

Address for correspondence: M. Buchwald, Dept. of Genetics, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada.

REFERENCES

1. BEAUDET, A. L., G. L. FELDMAN, S. D. FERNBACH, G. J. BUFFONE, AND W. E. O'BRIEN. Linkage disequilibrium, cystic fibrosis and genetic counselling. *Am. J. Hum. Genet.* 44: 319–326, 1989.
2. BEAUDET, A., A. BOWCOCK, M. BUCHWALD, L. CAVALLI-SFORZA, M. FARRALL, M.-C. KING, K. KLINGER, J.-M. LALOUEL, G. LATHROP, S. NAYLOR, J. OTT, L.-C. TSUI, B. WAINWRIGHT, P. WATKINS, R. WHITE, AND R. WILLIAMSON. Linkage of cystic fibrosis to two tightly linked DNA markers: joint report from a collaborative study. *Am. J. Hum. Genet.* 39: 681–693, 1986.
3. BEAUDET, A. L., J. E. SPENCE, M. MONTES, W. E. O'BRIEN, X. ESTIVILL, M. FARRALL, AND R. WILLIAMSON. Experience with new DNA markers for the diagnosis of cystic fibrosis. *N. Engl. J. Med.* 318: 50–51, 1988.
4. BERGER, W., J. HEIN, H. GEDSCHOLD, I. BAUER, A. SPEER, M. FARRALL, R. WILLIAMSON, AND C. COUTELLE. Crossovers in two German cystic fibrosi families determine probe order for MET, 7C22 and XV-2c/CS.7. *Hum. Genet.* 77: 197–199, 1987.
5. BIRD, A. P. CpG-rich islands and the function of DNA methylation. *Nature Lond.* 321: 209–213, 1986.
6. BOTSTEIN, D., R. L. WHITE, M. SKOLNICK, AND R. W. DAVIES. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314–331, 1980.
7. CANTOR, C. R., C. L. SMITH, AND M. K. MATHEW. Pulsed-field gel electrophoresis of very large DNA molecules. *Annu. Rev. Biophys. Biophys. Chem.* 17: 287–304, 1988.
8. CHILDS, B., N. A. HOLTZMAN, H. H. KAZAZIAN, AND D. L. VALLE (Editors). *Molecular Genetics in Medicine.* New York: Elsevier, 1988.
9. COLLINS, F. S., L. M. L. DRUMM, J. L. COLE, W. K. LOCKWOOD, G. F. VANDE WOUDE, AND M. C. IANNUZZI. Construction of a general human chromosome jumping library, with application to cystic fibrosis. *Science Wash. DC* 235: 1046–1049, 1987.
10. COREY, M., P. DURIE, D. MOORE, G. FORSTNER, AND H. LEVISON. Familial concordance of pancreatic function in cystic fibrosis. *J. Pediatr.* In press.
11. CUTTING, G. R., S. E. ANTONORAKIS, K. H. BUETOW, L. M. KASCH, B. J. ROSENSTEIN, AND H. H. KAZAZIAN. Analysis of DNA polymorphism haplotypes linked to the cystic fibrosis locus in North American black and Caucasian families supports the exist-

ence of multiple mutations of the cystic fibrosis gene. *Am. J. Hum. Genet.* 44: 307–318, 1989.
12. *Cystic Fibrosis Foundation Report 1978.* Rockville, MD: Cystic Fibrosis Found., 1980.
13. DANKS, D. M., J. ALLAN, P. D. PHELAN, AND C. CHAPMAN. Mutations at more than one locus may be involved in cystic fibrosis-evidence based on first-cousin data and direct counting of cases. *Am. J. Hum. Genet.* 35: 838–844, 1983.
14. DANKS, D. M., P. D. PHELAN, AND C. CHAPMAN. Retraction: no evidence for more than one locus in cystic fibrosis. *Am. J. Hum. Genet.* 36: 1398–1402, 1984.
15. DAVIS, P. B., AND P. A. DI SANT'AGNESE. A review, cystic fibrosis at forty-quo vadis? *Pediatr. Res.* 14: 83–87, 1980.
16. DEAN, M., M. PARK, M. M. LE BEAU, T. S. ROBINS, M. O. DIAZ, J. D. ROWLEY, D. G. BLAIR, AND G. F. VANDE WOUDE. The human *met* oncogene is related to the tyrosine kinase oncogenes. *Nature Lond.* 318: 385–388, 1985.
17. DEVOTO, M., M. ANTONELLI, F. BELLINI, G. BORGO, O. CASTIGLIONE, L. CURCIO, B. DALLAPICOLLA, M. FERRERI, P. GASPARINI, A. GIUNTA, L. MARIANELLI, G. MASTELLA, G. NOVELLI, P. PIGNATTI, L. ROMANO, G. ROMEO, AND M. SEIA. Rare alleles of DNA RFLP's closely linked to the CF gene are significantly more frequent in Italian CF patients without pancreatic insufficiency. *Am. J. Hum. Genet.* 43: A82, 1988.
18. DRUM, M. L., C. L. SMITH, M. DEAN, J. L. COLE, M. C. IANNUZZI, AND F. S. COLLINS. Physical mapping of the cystic fibrosis region by pulsed-field gel electrophoresis. *Genomics* 2: 346–354, 1988.
19. DUNCAN, A., M. BUCHWALD, AND L.-C. TSUI. In situ hybridization of two cloned chromosome 7 sequences tightly linked to the cystic fibrosis locus. *Cytogenet. Cell Genet.* In press.
20. EIBERG, H., K. SCHMIEGELOW, L.-C. TSUI, M. BUCHWALD, E. NIEBUHR, P. D. PHELAN, R. WILLIAMSON, W. WARWICK, C. KOCH, AND J. MOHR. Cystic fibrosis, linkage with PON (HGM 8). *Cytogenet. Cell Genet.* 40: 623, 1985.
21. ESTIVILL, X., M. FARRALL, P. J. SCAMBLER, G. M. BELL, K. M. F. HAWLEY, N. J. LENCH, G. P. BATES, H. C. KRUYER, P. A. FREDERICK, P. STAINER, E. K. WATSON, R. WILLIAMSON, AND B. J. WAINWRIGHT. A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. *Nature Lond.* 326: 840–845, 1987.
22. ESTIVILL, X., M. FARRALL, R. WILLIAMSON, M. FERRARI, M. SEIA, A. M. GIUNTA, G. NOVELLI, L. POTENZA, D. DALLAPICOLLA, G. BORGO, P. GASPARINI, P. F. PIGNATTI, L. DE BENEDETTI, E. VITALE, M. DEVOTO, AND G. ROMEO. Linkage disequilibrium between cystic fibrosis and linked DNA polymorphisms in Italian families: a collaborative study. *Am. J. Hum. Genet.* 43: 23–28, 1988.
23. ESTIVILL, X., AND R. WILLIAMSON. A rapid method to identify cosmids containing rare restriction sites. *Nucleic Acids Res.* 15: 1415–1425, 1987.
24. FARRALL, M., B. J. WAINWRIGHT, G. L. FELDMAN, A. BEAUDET, Z. SRETENOVIC, D. HALLEY, M. SIMON, L. DICKERMAN, M. DEVOTO, G. ROMEO, J.-C. KAPLAN, A. KITZIS, AND R. WILLIAMSON. Recombinations between IRP and cystic fibrosis. *Am. J. Hum. Genet.* 43: 471–475, 1988.
25. FRIZZELL, R. A., G. RECHKEMMER, AND R. L. SHOEMAKER. Altered regulation of airway epithelial cell chloride channels in cystic fibrosis. *Science Wash. DC* 233: 558–560, 1986.
26. FUJIWARA, T. M., K. MORGAN, R. H. SCHWARTZ, R. A. DOHERTY, S. R. MILLER, K. KLINGER, P. STANISLOVITIS, N. STUART, AND P. C. WATKINS. Genealogical analysis of cystic fibrosis families and chromosome 7q RFLP haplotypes in the Hutterite Brethren. *Am. J. Hum. Genet.* 44: 327–337, 1989.
27. GOODCHILD, M. C., J. H. EDWARDS, K. P. GLENN, C. GRINDEY, R. HARRIS, P. MACKINTOSH, AND J. WENTZEL. A search for linkage in cystic fibrosis. *J. Med. Genet.* 7: 417–419, 1976.
28. IANUZZI, M. C., M. DEAN, M. L. DRUMM, N. HIDAKA, J. L. COLE, A. PERRY, C. STEWART, B. GERRARD, AND F. S. COLLINS. Isolation of additional polymorphic clones from the cystic fibrosis region, using chromosome jumping from D758. *Am. J. Hum. Genet.* 44: 695–703, 1989.
29. KEREM, B.-S., J. A. BUCHANAN, P. DURIE, M. L. COREY, H. LEVINSON, J. M. ROMMENS, M. BUCHWALD, AND L.-C. TSUI. DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. *Am. J. Hum. Genet.* 44: 827–834, 1989.
30. KNOWLES, M. R., M. J. STUTTS, A. SPOCK, N. FISHER, J. T. GATZY, AND R. C. BOUCHER. Abnormal ion permeation through cystic fibrosis respiratory epithelium. *Science Wash. DC* 221: 1067–1070, 1983.
31. KNOWLTON, R. G., O. COHEN-HAGUENAUER, V. C. NGUYEN, J. FREZAL, V. BROWN, D. BARKER, J. C. BRAMAN, J. W. SCHUMM, L.-C. TSUI, M. BUCHWALD, AND H. DONIS-KELLER. A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature Lond.* 318: 380–382, 1985.
32. LATHROP, G. M., M. FARRALL, P. O'CONNELL, B. WAINWRIGHT, M. LEPPERT, Y. NAKAMURA, N. LENCH, H. KRUYER, M. DEAN, M. PARK, G. VANDE WOUDE, J.-M. LALOUEL, R. WILLIAMSON, AND R. WHITE. Refined linkage map of chromosome 7 in the region of the cystic fibrosis gene. *Am. J. Hum. Genet.* 42: 38–44, 1988.
33. LI, M., J. D. MCCANN, C. M. LIEDKE, A. C. NAIRN, P. GREENGARD, AND M. J. WELSH. Cyclic AMP-dependent protein kinase opens chloride channels in normal but not cystic fibrosis airway epithelium. *Nature Lond.* 331: 358–360, 1988.
34. MCKUSICK, V. A. *Mendelian Inheritance in Man* (8th ed.). Baltimore, MD: Johns Hopkins Univ. Press, 1988.
35. MONACO, A. P., R. L. NEVE, C. COLLETTI-FEENER, C. J. BERTELSON, D. M. KURNIT, AND L. M. KUNKEL. Isolation of candidate cDNAs for portions of the Duchenne muscular gene. *Nature Lond.* 233: 646–650, 1986.
36. MORTON, N. E. Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* 7: 277–318, 1955.
37. OBER, C., A. BOMBARD, R. DHALIWAL, S. ELIAS, J. FAGAN, T. G. LAFFLER, A. O. MARTIN, AND B. ROSINKY. Studies of cystic fibrosis in Hutterite families by using linked DNA probes. *Am. J. Hum. Genet.* 41: 1145–1151, 1987.
38. ORKIN, S. H. Reverse genetics and human disease (Review). *Cell* 47: 845–850, 1986.
39. POUSTKA, A., AND H. LEHRACH. Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. *Trends Genet.* 2: 174–179, 1986.
40. POUSTKA, A. M., H. LEHRACH, R. WILLIAMSON, AND G. BATES. A long-range restriction map encompassing the cystic fibrosis locus and its closely linked genetic markers. *Genomics* 2: 337–345, 1988.
41. QUINTON, P. M. Chloride impermeability in cystic fibrosis. *Nature Lond.* 301: 421–422, 1983.
42. QUINTON, P., AND G. MASTELLA (Editors). *Cellular and Molecular Basis of Cystic Fibrosis.* San Francisco, CA: San Francisco Press, 1988.
43. RIORDAN, J. R., AND M. BUCHWALD (Editors). *Genetics and Epithelial Cell Dysfunction in Cystic Fibrosis.* New York: Liss, 1987.
44. ROMEO, G., M. BIANCO, M. DEVOTO, P. MENOZZI, G. MASTELLA, A. M. GIUNTA, C. MICALIZZI, M. ANTONELLI, A. BATTISTINI, F. SANTAMARIA, D. CASTELLO, A. MARIANELLI, A. G. MARCHI, A. MANCA, AND A. MIANO. Incidence in Italy, genetic heterogeneity, and segregation analysis of cystic fibrosis. *Am. J. Hum. Genet.* 37: 338–349, 1985.
45. ROMMENS, J. M., B.-S. KEREM, G. MELMER, D. KENNEDY, N. PLAVSIC, R. ROZHAHEL, D. MARKIEWICZ, M. ZSIGA, J. R. RIORDAN, M. BUCHWALD, AND L.-C. TSUI. Genetic and physical mapping of the chromosomal region containing the cystic fibrosis locus (Abstract). *Am. J. Hum. Genet.* 43: A199, 1988.
46. ROMMENS, J. M., S. ZENGERLING, J. BURNS, G. MELMER, B.-S. KEREM, N. PLAVSIC, M. ZSIGA, D. KENNEDY, D. MARKIEWICZ, R. ROZHAMEL, J. R. RIORDAN, M. BUCHWALD, AND L.-C. TSUI. Identification and regional localization of DNA markers on chromosome 7 for the cloning of the cystic fibrosis gene. *Am. J. Hum. Genet.* 43: 645–663, 1988.
47. SAIKI, R. K., S. SCHARF, F. FALOONA, K. B. MULLIS, G. T. HORN, H. A. ERLICH, AND N. ARNHEIM. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science Wash. DC* 230: 1350–1354, 1985.
48. SCAMBLER, P. J., H.-Y. LAW, R. WILLIAMSON, AND C. S. COOPER. Chromosome mediated gene transfer of six DNA markers linked to the cystic fibrosis locus human chromosome seven. *Nucleic Acids Res.* 14: 7159–7174, 1986.
49. SCHMIEGELOW, K., H. EIBERG, L.-C. TSUI, M. BUCHWALD, P. D. PHELAN, R. WILLIAMSON, W. WARWICK, E. NIEBUHR, J. MOHR, M. SCHWARTZ, AND C. KOCH. Linkage between the loci for cystic fibrosis and paraoxonase. *Clin. Genet.* 29: 374–377, 1986.
50. SCHOUMACHER, R. A., R. L. SHOEMAKER, D. R. HALM, E. A. TALLANT, R. W. WALLACE, AND R. A. FRIZZELL. Phosphorylation

fails to activate chloride channels from cystic fibrosis airway cells. *Nature Lond.* 330: 752–754, 1987.
51. SING, C. F., D. R. RISSER, W. F. HOWATT, AND R. P. ERICKSON. Phenotypic heterogeneity in cystic fibrosis. *Am. J. Med. Genet.* 13: 179–195, 1982.
52. SINGER, M. F. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* 28: 433–434, 1982.
53. SPENCE, M. A., AND L.-C. TSUI. Report of the committee on the genetic constitution of chromosomes 7, 8 and 9. *Cytogenet. Cell Genet.* 46: 170–189, 1987.
54. STEINBERG, A. G., AND N. E. MORTON. Sequential test for linkage between cystic fibrosis of the pancreas and the MNS locus. *Am. J. Hum. Genet.* 8: 177–189, 1956.
55. STEINBERG, A. G., H. SCHWACHMAN, F. H. ALLEN, AND R. R. DOOLEY. Linkage studies with cystic fibrosis of the pancreas. *Am. J. Hum. Genet.* 8: 162–176, 1956.
56. STURGESS, J. M., E. CZEGLEDY-NAGY, M. COREY, AND M. W. THOMPSON. Cystic fibrosis in Ontario. *Am. J. Med. Genet.* 22: 383–393, 1985.
57. TALAMO, R. C., B. J. ROSENSTEIN, AND R. W. BENINGER. Cystic Fibrosis, In: *The Metabolic Basis of Inherited Disease* (5th ed.), edited by J. B. Stanbury, J. B. Wyngaarden, D. S. Frederickson, J. L. Goldstein, and M. S. Brown. New York: McGraw-Hill, 1983, p. 1889–1971.
58. THOMPSON, M. W. Genetics of cystic fibrosis. In: *Perspectives in Cystic Fibrosis*, edited by J. M. Sturgess. Toronto, Canada: Canadian Cystic Fibrosis Foundation, 1980, p. 281–291.
59. TSUI, L.-C., M. BUCHWALD, D. BARKER, J. C. BRAMAN, R. KNOWLTON, J. W. SCHUMM, H. EIBERG, J. MOHR, D. KENNEDY, N. PLAVSIC, M. ZSIGA, D. MARKIEWICZ, G. AKOTS, V. BROWN, C. HELMS, T. GRAVIUS, C. PARKER, K. REDIKER, AND H. DONIS-KELLER. Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science Wash. DC* 230: 1054–1057, 1985.
60. TSUI, L.-C., J. M. ROMMENS, J. BURNS, S. ZENGERLING, J. R. RIORDAN, L. R. CARLOCK, K.-H. GRZESCHIK, AND M. BUCHWALD. Progress towards cloning the cystic fibrosis gene. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 319: 263–273, 1988.
61. WAINWRIGHT, B. J., P. J. SCAMBLER, J. SCHMIDTKE, E. A. WATSON, H.-Y. LAW, M. FARRALL, H. J. COOKE, H. EIBERG, AND R. WILLIAMSON. Localization of cystic fibrosis locus to human chromosome 7 cen-q22. *Nature Lond.* 318: 384–385, 1985.
62. WAINWRIGHT, B. J., P. J. SCAMBLER, P. STANIER, E. K. WATSON, G. BELL, C. WICKING, X. ESTIVILL, M. COURTNEY, A. BOUE, P. S. PEDERSEN, R. WILLIAMSON, AND M. FARRALL. Isolation of a human gene with protein sequence similarity to human and murine int-1 and *Drosophila* segment polarity mutant *wingless*. *EMBO J.* 7: 1743–1748, 1988.
63. WARWICK, W. J., R. E. POGUE, H. U. GERBER, AND C. J. NESBITT. Survival patterns in cystic fibrosis. *J. Chron. Dis.* 28: 609–622, 1975.
64. WELSCH, M. J., AND R. B. FICK. Cystic fibrosis. *J. Clin. Invest.* 80: 1523–1526, 1987.
65. WELSCH, M. J., AND C. M. LIEDKE. Chloride and potassium channels in cystic fibrosis airway epithelium. *Nature Lond.* 322: 467–470, 1986.
66. WHITE, R., AND C. T. CASKEY. The human as an experimental system in molecular genetics. *Science Wash. DC* 240: 1483–1488, 1988.
67. WHITE, R., S. WOODWARD, M. LEPPERT, P. O'CONNELL, Y. NAKAMURA, M. HOFF, J. HERBST, J.-M. LALOUEL, M. DEAN, AND G. VANDE WOUDE. A closely linked genetic marker for cystic fibrosis. *Nature Lond.* 318: 382–384, 1985.
68. WOOD, R. E., T. F. BOAT, AND C. F. DOERSHUK. State of the art: cystic fibrosis. *Am. Rev. Respir. Dis.* 113: 833–878, 1976.
69. WRIGHT, S. W., AND N. E. MORTON. Genetic studies on cystic fibrosis in Hawaii. *Am. J. Hum. Genet.* 20: 157–169, 1968.
70. ZENGERLING, S., L.-C. TSUI, K.-H. GRZESCHIK, K. OLEK, J. R. RIORDAN, AND M. BUCHWALD. Mapping of DNA markers linked to the cystic fibrosis locus on the long arm of chromosome 7. *Am. J. Hum. Genet.* 40: 228–236, 1987.