PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Progress Towards Cloning the Cystic Fibrosis Gene [and Discussion]

L.-C. Tsui, J. M. Rommens, J. Burns, S. Zengerling, J. R. Riordan, L. R. Carlock, K.-H. Grzeschik, M. Buchwald, S. V. Hodgson, J.-J. Cassiman and H. Sharma

Phil. Trans. R. Soc. Lond. B 1988 **319**, 263-273
doi: 10.1098/rstb.1988.0048

Phil. Trans. R. Soc. Lond. B **319**, 263–273 (1988)
Printed in Great Britain

Progress towards cloning the cystic fibrosis gene

By L.-C. Tsui[1,2], J. M. Rommens[1], J. Burns[1], S. Zengerling[1], J. R. Riordan[3], L. R. Carlock[4], K.-H. Grzeschik[5] and M. Buchwald[1,2]

[1] *Department of Genetics, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8*

[2] *Departments of Medical Genetics and Medical Biophysics, University of Toronto, Ontario, Canada M5S 1A1*

[3] *Department of Biochemistry, The Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8; Departments of Biochemistry and Clinical Biochemistry, University of Toronto, Ontario, Canada M5S 1A1*

[4] *Department of Molecular Biology and Genetics, Wayne State University, Detroit, Michigan 48202, U.S.A.*

[5] *Institut für Humangenetik der Universität, Vesaliusweg 12–14, D-4400, Münster, F.R.G.*

Genetic linkage analysis with polymorphic DNA markers (restriction fragment length polymorphisms: RFLPs) has allowed the assignment of the cystic fibrosis (CF) locus to the long arm of chromosome 7, within the region of band q31. Two of these markers, *MET* and *D7S8*, are tightly linked to the disease locus. Although recent data suggest that they are located on opposite sides of CF, the two can be separated by as much as 5 centimorgans. To obtain a better description of the CF locus and, eventually, to identify the affected gene, additional DNA markers are required to connect *MET* and *D7S8*, physically. We have screened the flow-sorted chromosome-7-specific library and thus far isolated 28 new probes from the 7q31 region by DNA hybridization analysis that uses a series of somatic cell hybrids containing various portions of human chromosome 7. Together with the previously identified markers, *MET*, *D7S8*, *D7S13* and *D7S16*, these new markers should provide a fine genetic and physical map for the chromosomal region surrounding CF. DNA segments can then be sequentially cloned by chromosome walking from points closest to the CF locus and examined for genes that are preferentially expressed in tissues known to be affected in the disease.

Introduction

With a disease frequency of 1 in 2000 live-births and a carrier frequency close to 5% in the Caucasian population, cystic fibrosis (CF) has often been regarded as the most common, severe, autosomal recessive disorder (for review see Talamo *et al.* (1983)). The major clinical symptoms of CF include chronic obstructive pulmonary disease, pancreatic enzyme insufficiency and elevated sweat electrolyte levels. If untreated, affected children usually die at an early age because of severe lung infection. As a result of recent advances in clinical management, however, the lifespan of patients has increased markedly and many of them now live to adulthood.

Biochemistry of CF cells

The basic biochemical defect in CF is unknown. A number of recent studies have provided convincing evidence that the regulation of ion transport is altered in CF epithelia, but it is not yet clear whether that is the primary defect of the disease. Abnormally high electrical potential differences were first observed across epithelial surfaces of CF nasal polyps and bronchi (Knowles *et al.* 1983). Similar observations were also reported for the sweat gland secretory and reabsorptive ducts (Quinton 1983; Quinton & Bijman 1983; Sato 1984). Studies using the latter tissue actually revealed a decreased chloride ion conductance across the ductal wall (Quinton 1983; Quinton & Bijman 1983). The electrolyte transport defect in CF has also been shown to persist in cultured epithelial cells derived from a variety of tissues, including nasal polyps (Yankaskas *et al.* 1985), tracheae (Widdicombe *et al.* 1985) and sweat-gland ducts (J. R. Riordan, unpublished observations). More recent studies showed that, although chloride-channel activities were absent in intact cultured CF airway epithelial cells, there is no difference between the normal and CF chloride-channel conductivities in membranes excised from these cells (Welsh & Liedtke 1986; Frizzell *et al.* 1986). The difference in channel conductivity is apparently elicited by cyclic AMP (cAMP) in response to β-adrenergic stimulation. However, no difference in the intracellular levels of cAMP was observed between normal and CF. It is also of interest to note that the activation of chloride and potassium channels by calcium appeared to be normal in CF (Frizzell *et al.* 1986; Welsh & Liedtke 1986).

These findings, taken together, suggest that the defect observed in CF chloride transport is probably due to an altered regulation at a site distal to cAMP accumulation (Welsh & Liedtke 1986; Frizzell *et al.* 1986). The persistence of the CF phenotype in the cultured epithelial cells strongly suggests that the CF gene is expressed in these cells. Unfortunately, very little is known about the biochemistry of chloride-ion transport and it is therefore difficult to pinpoint the regulatory defect in CF.

Mapping of the CF locus

As a first step in our attempt to identify the primary lesion in CF, we have initiated a search for the disease locus (*CF*) by using a genetic linkage approach. The simple autosomal recessive mode of inheritance of the disease suggests that it is possible to detect linkage to *CF* by studying segregation of genetic markers in two-generation families with two or more affected children. In addition, results of extensive population analyses are consistent with the assumption that CF is due to mutation(s) in a single gene (Danks *et al.* 1984; Romeo *et al.* 1985). Although several early attempts using conventional protein markers were not successful (Steinberg & Morton 1956; Steinberg *et al.* 1956; Goodchild *et al.* 1976), a linkage was demonstrated between *CF* and a yet unassigned marker, *PON*, a genetic determinant for serum paraoxonase activity (Eiberg *et al.* 1985).

Our effort has been focused on the use of DNA markers defined by restriction fragment length polymorphisms (RFLPs) (Botstein *et al.* 1980). After screening a large number of markers, covering a significant portion of the human genome (Tsui *et al.* 1985*a*, *b*), we discovered a linkage between *CF* and a randomly isolated DNA marker, *D7S15* (formerly *D0CRI-917*) (Tsui *et al.* 1985*c*). Subsequent localization of this DNA marker to chromosome 7 (Knowlton *et al.* 1985) and demonstration of linkage between CF and a number of other DNA

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

markers known to be on chromosome 7 (White *et al.* 1985; Wainwright *et al.* 1985; Scambler *et al.* 1985, 1986; Buchwald *et al.* 1986) established that *CF* is on chromosome 7.

Figure 1 illustrates our current understanding of the CF region. The regional assignments of the listed markers are shown on the left and the corresponding genetic map on the right, with distance estimates in centimorgans† (cM). Two of these marker loci, *MET* (White *et al.* 1985) and *D7S8* (Wainwright *et al.* 1985), appear to be less than 1 cM from *CF*, providing two close reference points for the CF gene as well as serving as markers for genetic diagnosis.

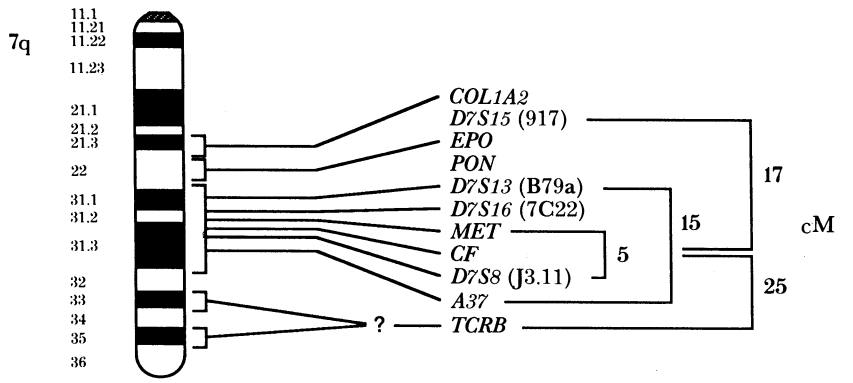

FIGURE 1. Genetic markers on the long arm of human chromosome 7 surrounding the CF locus. The markers (from top) are *COL1A2* (proα2(I) collagen), DNA segment *D7S15* (*D0CRI-917*), *EPO* (erythropoietin), *PON* (paraoxonase), *D7S13* (B79a), *D7S16* (7C22), *MET* (the met proto-oncogene), *D7S8* (J3.11), *A37* (this work), *TCRB* (The T-cell receptor β chain genes).

To obtain a more accurate description of the CF locus, we and others have combined and analysed the data from a total of 211 families, each with two or more affected children (Beaudet *et al.* 1986). The maximal likelihood estimate for the recombination fraction (θ) between *CF* and *MET* was 0.004 (with a lod score of 91.0), that between *CF* and *D7S8* was 0.003 (lod = 71.3) and that between *MET* and *D7S8* was 0.018 (lod‡ = 69.3). The results of multi-point analysis (Beaudet *et al.* 1986) and linkage disequilibrium calculations (Buetow *et al.* 1986) favour the interpretation that *MET* and *D7S8* are on opposite sides of the disease locus. Subsequent linkage analysis based on a larger number of non-CF pedigrees suggested that *MET* and *D7S8* could be 5 cM apart (Lathrop *et al.* 1988). Assuming a general estimate of 1 million base pairs (b.p.) per cM, the region between *MET* and *D7S8* could represent 5×10^6 b.p. of DNA sequence (which may contain as many as 100–200 genes).

On the basis of two-point and multi-point linkage analysis, we have derived the most likely order for a number of markers surrounding *CF*: *COL1A2–D7S15–PON–*[*MET–CF–D7S8*]*–TCRB* (Tsui *et al.* 1986*a*). However, to obtain a more precise location for *CF*, we have re-examined the chromosomal location for some of the CF-linked DNA markers by using a panel of human cell lines carrying deletions, and mouse–human somatic cell hybrids containing portions of chromosome 7 (Zengerling *et al.* 1987). The most informative cell line in this analysis is GM1059, a human fibroblast cell line which has a single interstitial deletion spanning band q31, 46, XX, del7 (pter → q22::q32 → qter). Hybridization analyses and careful

† The morgan is the unit of relative distance between genes on a chromosome. One centimorgan represents a crossover value of 1%.

‡ A lod score represents the $\log_{10}$ of the likelihood (L) ratio L(recombination fraction)/L(0.5).

densitometer tracing of the radioautographs revealed that *D7S8*, *MET*, *D7S13* and *D7S16* were all deleted from the del7 chromosome, resulting in half hybridizing band intensity when compared with diploid human fibroblast DNA (Estivill *et al.* 1986; Zengerling *et al.* 1987). In contrast, *COL1A2*, *D7S15* and *TCRB* were all present on the del7 chromosome (Zengerling *et al.* 1987). Taking together the chromosomal assignments for these markers and their genetic linkage relations with the disease locus, it is evident that *CF* is located in the middle third of the long arm of chromosome 7, most probably within the proximal half of band q31.

Additional information on the gene order has been derived from linkage disequilibrium analysis. Pairwise allelic association values derived from the combined data set suggested that *CF* is located between *MET* and *D7S8* (Beaudet *et al.* 1986; Buetow *et al.* 1986). Because the MET locus can be separated into two regions defined by the two probes metD and metH, the analysis could be further extended to determine the orientation of the met proto-oncogene and *D7S8*. The result of the latter analysis suggested that the order of these three loci is *MET*$_D$–*MET*$_H$–*D7S8* (Buetow *et al.* 1986; K. H. Buetow & L.-C. Tsui, unpublished observation). Based on the known association values in the human leucocyte antigen (HLA) region, the physical distance between *MET* and *D7S8* is estimated to be between 160 and 3000 kilobases. The upper limit of this estimate is thus close to that predicted from linkage analysis. In addition, because metD lies approximately 15 kilobases 5′ of metH (Park *et al.* 1986), this result would suggest that *CF* is located 3′ to the met proto-oncogene.

Analysis of the linkage data has also revealed that the majority, if not all, of the CF patients have inherited mutation(s) at a single locus (Tsui *et al.* 1986*a*, *b*; Beaudet *et al.* 1986). Although it is possible that some families might have segregated CF mutation(s) at a different locus, a formal proof for genetic heterogeneity in CF must await the molecular characterization of the affected gene(s). From the linkage disequilibrium detected between *CF* and the two tightly linked markers, it seems probable that the disease is the result of a single (or a small number of) mutaion(s).

For the rest of our discussion, we assume that there is only one single gene involved in CF. This gene, if mutated, confers the CF trait. We refer to this gene as the 'CF gene' and we make an effort to distinguish its normal and mutant counterparts (alleles) in our discussion.

Isolation of additional DNA markers

Despite the availability of a large number of molecular cloning techniques, the task of identifying the CF mutation is formidable. Although some of the approaches involve the use of CF epithelial cells, the majority exploit the inferred location of the CF gene based on the linked DNA markers. We have initiated a major effort towards the physical mapping of the 7q31 region as our next step in the search for the CF mutation. This approach has been facilitated by the invention of pulsed-field gel electrophoresis (Schwartz & Cantor 1984; Carle & Olson 1984) capable of resolving DNA fragments of up to a few million base pairs in length, and the discovery of restriction endonucleases capable of generating human DNA fragments of such sizes (reviewed in Van Ommen & Verkerk (1986)). Because *MET* and *D7S8* do not share any common restriction fragments (H. Lehrach & F. Collins, personal communication; our unpublished data), additional DNA markers are required to bridge the gap.

To isolate additional chromosome-7 probes within the CF region, we have screened almost 4000 DNA segments from a flow-sorted chromosome-7-specific genomic library (Deaven *et al.*

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 1. Results of human chromosome 7 marker screening

experiment	I[a]	II[b]	total
clones isolated	2268	1724	3992
clones with human or hamster repetitive DNA	1446	789	2235
probes tested on somatic cell hybrids	255	266	521
probes mapped on 7	118	81	199
probes within 7q31	10	15	25

[a] DNA segments were individually purified from the flow-sorted chromosome-7-specific phage library (Deaven *et al.* 1986) and used as probes in hybridization analysis; the majority of the 7q31-specific probes were identified by half-intensity hybridization signal with GM1059.

[b] DNA segments were purified in a pool from the above phage library and subcloned in plasmid vector pUC13; individual inserts from plasmids were used in hybridization analysis; all 7q31-specific probes were identified with the use of the somatic cell hybrid panel shown in figure 2.

1986). The results are shown in table 1. A total of 1757 clones were isolated as probes apparently free of highly repetitive DNA sequences and therefore suitable for hybridization analysis. Over 500 of these secondary isolates were radioactively labelled and hybridized to human–rodent somatic cell hybrids containing a whole, or portions of, human chromosome 7 (figure 2). Nearly 200 (38%) were found to be *bona fide* 'single-copy' human chromosome-7 probes. The remainder of the clones were found to contain either moderately repetitive human-specific sequences or sequences originating from hamster chromosomes co-purified during flow-sorting (Deaven *et al.* 1986). Based on the results of hybridization analyses with the somatic cell hybrids, 25 (13%) of the single-copy probes were identified as being located in the 7q31 region and therefore potentially useful for mapping the CF gene (figure 3). Restriction enzyme digestion showed that these probes are all different from each other.

In a parallel approach, we have attempted to isolate the CF gene directly by searching for genes that are preferentially expressed in cultured sweat-gland epithelial cells. Using radioactively labelled copy DNA (cDNA) probes made to total poly(A)$^+$ mRNA isolated from these cells, we have selected genomic DNA clones from the flow-sorted human chromosome-7-specific library. Three of these clones (5-21, 89 and 117) were found to be located within the q31 region. These three clones (presumably representing genes or fragments of genes) are currently being characterized.

A subtraction hybridization procedure has also been established to increase the chance of recovering clones corresponding to epithelial cell-specific genes. In a preliminary experiment, ^{32}P-labelled cDNA probes were first allowed to anneal with excess poly(A)$^+$ RNA from fibroblasts to a high Rot† value (> 2000 mol s l^{-1}) and the unhybridized fraction (subtracted probe) was incubated with a gel-blot containing RNA samples from cultured sweat-gland cells and fibroblasts to a Rot value of *ca.* 1. More than ten different sweat-gland-specific transcripts, ranging from several hundred to 15000–20000 nucleotides in length, were readily detectable on a radioautograph exposed for only 2 h (figure 4). Similarly, subtracted probes have been used to screen a small portion of the chromosome-7-specific library, and several clones have been isolated. These clones do not appear to contain any repetitive DNA sequences; their chromosomal locations are being determined by hybridization analysis with the somatic cell hybrid panel as described above (figure 2). Standard genetic linkage analysis will also be done.

† 'Rot' is defined here as the initial concentration of single-stranded cDNA multiplied by the total time of hybridization.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

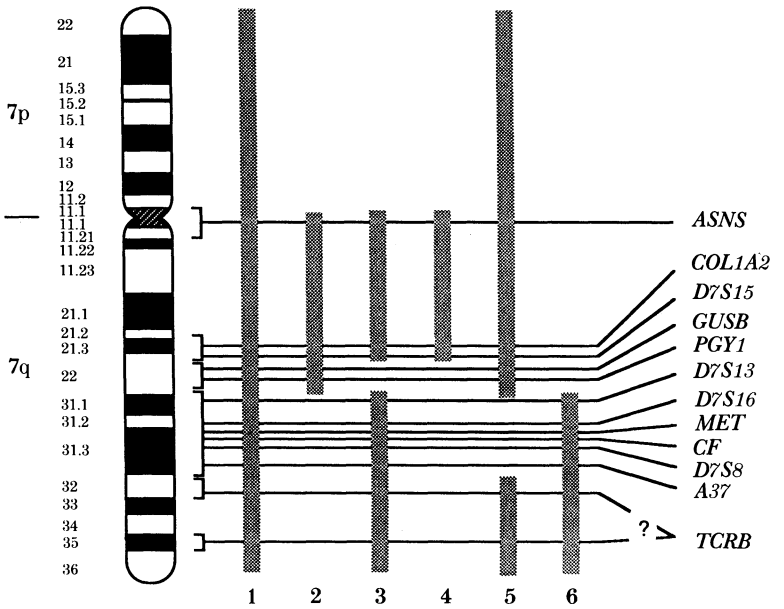

FIGURE 2. Extent of chromosome 7 retained in the six somatic cell hybrids used in the present analysis. Vertical bars indicate the portions of chromosome 7 present in these cell lines. Lines 1–3 correspond to 4AF1/106, 1CF2/5 and 1EF2/3, which are human–hamster hybrids containing human chromosome 7 as the only human material (Arfin *et al.* 1983). Line 4 is a derivative of line 3. Line 5 is a mouse–human hybrid derived from fusion between Rag (mouse cell) and GM1059; it contains the del7 chromosome but not the normal counterpart from GM1059. Line 6 contains only a portion of human chromosome 7 as indicated. Details of the two latter cell lines will be presented elsewhere. The marker abbreviations are the same as in figure 1, with the addition of *ASNS* (asparagine synthetase), *GUSB* (β-glucuronidase) and *PGY1* (the multidrug resistance genes).

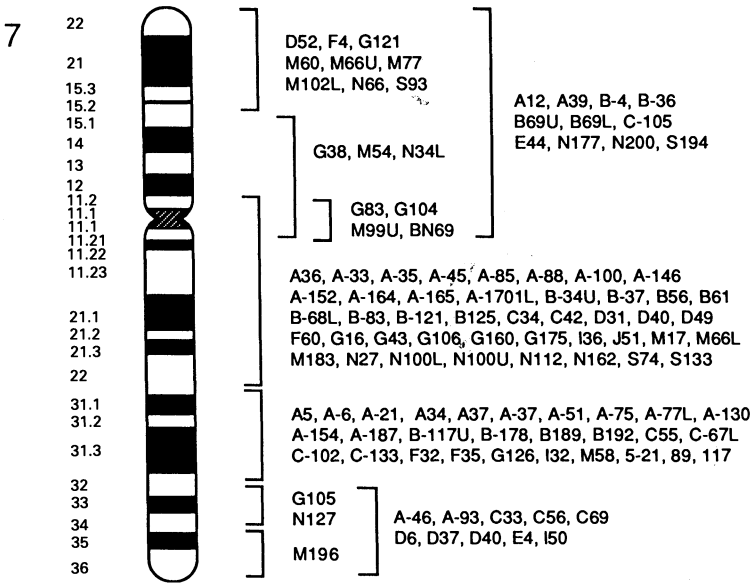

FIGURE 3. Regional localization of 111 chromosome 7 markers. Markers designated A5, A12, etc. are from experiment I (Table 1); A-6, A-21, etc. from experiment II; 5-21, 89 and 117 from the sweat-gland cDNA hybridization screening (see text).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

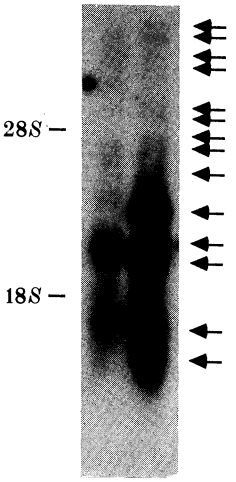

FIGURE 4. Subtraction hybridization analysis. Radioautograph shows the result of gel-blot hybridization of poly(A)$^+$ RNAs from cultured fibroblasts (left lane) and sweat-gland cells (right lane) with sweat-gland-specific ^{32}P-labelled cDNA probe pre-annealed with excess poly(A)$^+$ RNA from fibroblasts.

Including the previously identified probes for *MET*, *D7S8* (J3.11), *D7S13* (B79a) and *D7S16* (7C22), a total of 32 markers are available from the 7q31 region (figure 3). Because the somatic hybrid cell lines used to define the physical location of these probes localize them to a chromosome segment of approximately 20 cM (or 20000 kilobases) in size, we estimate that 40 evenly distributed probes would be sufficient to cover the q31 region entirely when analysed by pulsed-field gel electrophoresis. Additional somatic-cell hybrids and patient cell lines carrying deletions and rearrangements near the CF locus are being sought to narrow down further the regional localization of these probes. Because *MET* and *D7S8* have been estimated to be as much as 5 cM apart from each other, it is possible that several of the newly isolated probes may reside between these two markers.

While additional 7q31 DNA markers are being isolated from the library, we have begun to examine the newly available probes for associated RFLPs useful in genetic linkage analysis. One of them, designated as A37, was found to detect a *Pst* I RFLP, and was studied with the CF families. This study showed that *A37* is approximately 1 cM from *CF*, most likely on the opposite side of *MET* and on the same side of, but distal to, *D7S8* (see figures 1 and 2). Several families, which revealed recombination events between *CF* and *MET* and *A37*, are particularly useful for rapid identification of markers that are located between these two markers (no recombination was detected between *CF* and *D7S8* in our families). However, extensive family studies will be required to confirm the localization and to obtain information for linkage disequilibrium analysis.

The physical relation among the newly isolated and the previously reported DNA markers are being examined by pulsed-field gel electrophoresis. The orthogonal (Carle & Olson 1984) and reversed field (Carle *et al.* 1986) techniques have been used in our laboratory. Some of the above markers have been analysed, but none of them has yet revealed any overlapping restriction fragments. However, we expect to obtain a practical map through the analysis of additional probes. Once the *MET–D7S8* region is defined, chromosome 'walking' or 'jumping' (Collins & Weissman 1984; Poustka *et al.* 1987; Collins *et al.* 1987) strategies can

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

be used with other molecular techniques to search for genes that are expressed in the tissues affected in CF.

Because at present a candidate CF gene can only be tested by its chromosomal location, additional assays are required to provide a more definitive proof of its involvement, if any, in CF. Expression in secretory epithelial tissue is a necessary, but not sufficient, requirement for the putative CF gene. A more important criterion, however, would be its ability to correct the CF phenotype *in vitro* with epithelial cells from patients, as the properties of ion transport in the cultured CF cells are clearly distinguishable from those of the normals.

It is hoped that the knowledge about the physical and genetic location of the CF locus, the availability of cultured cells carrying the disease phenotype and the application of molecular biology will eventually lead to the discovery of the basic defect in CF.

The authors thank Lidija Covic, Carol Geller, Zahi Hajjar, Dara Kennedy, Shaul Kuper, Douglas Johnson, Danuta Markiewicz, Helenka Panak, Natasa Plavsic, Richard Rozmahel and Martha Zsiga for their assistance in probe isolation and hybridization analysis. This work was supported by grants from the Canadian Cystic Fibrosis Foundation (CCFF), the Cystic Fibrosis Foundation (U.S.A.) and the National Institutes of Health (U.S.A.), and donations from the Cystic Fibrosis Association of North Dakota and Parents for Cystic Fibrosis Research, Plandome, New York. L.-C.T. is a Research Scholar of CCFF. S.Z. is a predoctoral student of the Deutscher Akademischer Austauschdienst.

References

Arfin, S. M., Cirullo, R. E., Arredondo-Vega, F. X. & Smith, M. 1983 Assignment of structural gene for asparagine synthetase to human chromosome 7. *Somatic Cell Genet.* **9**, 517–531.

Beaudet, A., Bowcock, A., Buchwald, M., Cavalli-Sforza, L., Farrall, M., King, M.-C., Klinger, K., Lalouel, J.-M., Lathrop, G., Naylor, S., Ott, J., Tsui, L.-C., Wainwright, B., Watkins, P., White, R. & Williamson, R. 1986 Linkage of cystic fibrosis to two tightly linked DNA markers: joint report from a collaborative study. *Am. J. hum. Genet.* **39**, 681–693.

Botstein, D., White, R., Skolnick, M. & Davis, R. 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. hum. Genet.* **32**, 314–331.

Buchwald, M., Zsiga, M., Markiewicz, D., Plavsic, N., Kennedy, D., Zengerling, S., Willard, H. F., Tsipouras, P., Schmiegelow, K., Schwartz, M., Eiberg, H., Mohr, J., Barker, D., Donis-Keller, H. & Tsui, L.-C. 1986 Linkage of cystic fibrosis to the proα2(1) collagen gene, COL1A2, on chromosome 7. *Cytogenet. Cell Genet.* **41**, 234–239.

Buetow, K. H., Buchwald, M. & Tsui, L.-C. 1986 Linkage disequilibrium between the cystic fibrosis locus (CF) and MET-D7S8 haplotypes. *Am. J. hum. Genet.* **38**, A 91.

Carle, G. F., Frank, M. & Olson, M. V. 1986 Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. *Science, Wash.* **232**, 65–68.

Carle, G. F. & Olson, M. V. 1984 Separation of chromosomal DNA molecules from yeast by orthogonal field alteration electrophoresis. *Nucl. Acids Res.* **12**, 5647–5664.

Collins, F. S., Drumm, M. L., Cole, J. L., Lockwood, W. K., Vande Woude, G. F. & Iannuzzi, M. C. 1987 Construction of a general human chromosome jumping library, with application to cystic fibrosis. *Science, Wash.* **235**, 1046–1049.

Collins, F. S. & Weissman, S. M. 1984 Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6812–6816.

Danks, D. M., Phelan, P. D. & Chapman, C. 1984 Retraction: no evidence for more than one locus in cystic fibrosis. *Am. J. hum. Genet.* **36**, 1398–1402.

Deaven, L. L., Van Dilla, M. A., Bartholdi, M. F., Carrano, A. V., Cram, L. S., Fuscoe, J. C., Gray, J. W., Hilderbrand, C. E., Moyzis, R. K. & Perlman, J. 1986 Construction of human chromosome-specific DNA libraries from flow-sorted chromosomes. In *Molecular biology of Homo sapiens* (*Cold Spring Harb. Symp. quant. Biol.* **51**, 159–167).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Eiberg, H., Schmiegelow, K., Tsui, L.-C., Buchwald, M., Niebuhr, E., Phelan, P. D., Williamson, R., Warwick, W., Koch, C. & Mohr, J. 1985 Cystic fibrosis, linkage with PON. *Cytogenet. Cell Genet.*, **40**, 623.

Estivill, X., Schmidtke, J., Williamson, R. & Wainwright, B. 1986 Chromosomal assignment and restriction fragment length polymorphism analysis of the anonymous DNA probe B79a at 7q22. *Hum. Genet.* **74**, 320–322.

Frizzell, R. A., Rechkemmer, G. & Shoemaker, R. L. 1986 Altered regulation of airway epithelial cell chloride channels in cystic fibrosis. *Science, Wash.* **233**, 558–560.

Goodchild, M. C., Edwards, J. H., Glenn, K. P., Grindey, C., Harris, R., Mackintoch, P. & Wentzel, J. 1976 A search for linkage in cystic fibrosis. *J. med. Genet.* **7**, 417–419.

Knowles, M. R., Stutts, M. J., Spock, A., Fisher, N., Gatzy, J. T. & Boucher, R. C. 1983 Abnormal ion permeation through cystic fibrosis respiratory epithelium. *Science, Wash.* **221**, 1067–1070.

Knowlton, R. G., Cohen-Haguenauer, O., Nguyen, V. C., Frézal, J., Brown, V., Barker, D., Braman, J. C., Schumm, J. W., Tsui, L.-C., Buchwald, M. & Donis-Keller, H. 1985 A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature, Lond.* **318**, 380–382.

Lathrop, G. M., Farrall, M., O'Connell, P., Wainwright, B. J., Leppert, M., Nakamura, Y., Lench, N., Kruyer, H., Dean, M., Park, M., Vande Woude, G., Lalouel, J.-M., Williamson, R. & White, R. 1988 Refined linkage map of chromosome 7 in the region of the cystic fibrosis gene. *Am. J. hum. Genet.* (In the press.)

Park, M., Dean, M., Cooper, C. S., Schmidt, M., O'Brien, S. J., Blair, D. G. & Vande Woude, G. F. 1986 Mechanism of *met* oncogene activation. *Cell* **45**, 895–904.

Poustka, A., Pohl, T. M., Barlów, D. P., Frischauf, A.-M. & Lehrach, H. 1987 Construction and use of human chromosome jumping libraries from *Not*I-digested DNA. *Nature, Lond.* **325**, 353–355.

Quinton, P. M. 1983 Chloride impermeability in cystic fibrosis. *Nature, Lond.* **301**, 421–422.

Quinton, P. M. & Bijman, J. 1983 Higher bioelectric potentials due to decreased chloride absorption in the sweat glands of patients with cystic fibrosis. *New Engl. J. Med.* **308**, 1185–1189.

Romeo, G., Bianco, M., Devoto, M., Menozzi, P., Mastella, G., Giunta, A. M., Micalizzi, C., Antonelli, M., Battistini, A., Santamaria, F., Castello, D., Marianelli, A., Marchi, A. G., Manca, A. & Maino, A. 1985 Incidence in Italy, genetic heterogeneity, and segregation analysis of cystic fibrosis. *Am. J. Hum. Genet.* **37**, 338–349.

Sato, K. 1984 Differing luminal protential difference of cystic fibrosis and control sweat secretory coils *in vitro*. *Am. J. Physiol.* **247**, R646–649.

Scambler, P. J., Wainwright, B. J., Farrall, M., Bell, J., Stanier, P., Lench, N. J., Bell, G., Kruyer, H., Ramirez, F. & Williamson, R. 1985 Linkage on COL1A2 collagen gene to cystic fibrosis and its clinical implications. *Lancet* ii, 1241–1242.

Scambler, P. J., Wainwright, B. J., Watson, E., Bates, G., Bell, G., Williamson, R. & Farrall, M. 1986 Isolation of a further anonymous informative DNA sequence from chromosome seven closely linked to cystic fibrosis. *Nucl. Acids Res.* **14**, 1951–1961.

Schwartz, D. C. & Cantor, C. R. 1984 Separation of yeast chromosome-sized DNAs by pulsed field gradient electrophoresis. *Cell* **37**, 67–75.

Steinberg, A. G. & Morton, N. E. 1956 Sequential test for linkage between cystic fibrosis of the pancreas and the MNS locus. *Am. J. hum. Genet.* **8**, 177–189.

Steinberg, A. G., Schwachman, H., Allen, F. H. & Dooley, R. R. 1956 Linkage studies with cystic fibrosis of the pancreas. *Am. J. hum. Genet.* **8**, 162–176.

Talamo, R. C., Rosenstein, B. J. & Berninger, R. W. 1983 Cystic fibrosis. In *The metabolic basis of inherited disease* (ed. J. B. Stanbury, J. B. Wyngaarden, D. S. Fredrickson, J. L. Goldstein & M. S. Brown), 5th edn, pp. 1889–1971. New York: McGraw-Hill.

Tsui, L.-C., Cox, D. W., McAlpine, P. J. & Buchwald, M. 1985*a* Cystic fibrosis: analysis of linkage of the disease locus to red cell and plasma protein markers. *Cytogenet. Cell Genet.*, **39**, 238–239.

Tsui, L.-C., Zsiga, M., Kennedy, D., Plavsic, N., Markiewicz, D. & Buchwald, M. 1985*b* Cystic fibrosis: progress in mapping the disease locus using polymorphic DNA markers. I. *Cytogenet. Cell Genet.* **39**, 299–301.

Tsui, L.-C., Buchwald, M., Barker, D., Braman, J. C., Knowlton, R. G., Schumm, J., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Brown, V., Helms, C., Gravius, T., Parker, C., Rediker, K. & Donis-Keller, H. 1985*c* Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science, Wash.* **230**, 1054–1057.

Tsui, L.-C., Zengerling, S., Willard, H. F. & Buchwald, M. 1986*a* Mapping of the cystic fibrosis locus on chromosome 7. In *Molecular biology of Homo sapiens*, (*Cold Spring Harb. Symp. quant. Biol.* **51**, 325–335).

Tsui, L.-C., Buetow, K. & Buchwald, M. 1986*b* Genetic analysis of cystic fibrosis using linked DNA markers. *Am. J. hum. Genet.* **39**, 720–728.

Van Ommen, G. J. B. & Verkerk, J. M. H. 1986 Restriction analysis of chromosomal DNA in a size range up to two million base pairs by pulsed field gradient electrophoresis. In *Human genetic disease, a practical approach* (ed. K. E. Davies), pp. 113–133. Oxford: IRL Press.

Wainwright, B. J., Scambler, P. J., Schmidtke, J., Watson, E. A., Law, H.-Y., Farrall, M., Cooke, H. J., Eiberg, H. & Williamson, R. 1985 Localization of cystic fibrosis locus to human chromosome 7cen-q22. *Nature, Lond.* **318**, 384–385.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Welsh, M. J. & Liedtke, C. M. 1986 Chloride and potassium channels in cystic fibrosis airway epithelium. *Nature, Lond.* **322**, 467–470.

White, R., Woodward, S., Leppert, M., O'Connell, P., Nakamura, Y., Hoff, M., Herbst, J., Lalouel, J.-M., Dean, M. & Vande Woude, G. 1985 A closely linked genetic marker for cystic fibrosis. *Nature, Lond.* **318**, 382–384.

Widdicombe, J. H., Welsh, M. J. & Finkbeiner, W. E. 1985 Cystic fibrosis decreases the apical membrane chloride permeability of monolayers cultured from cells of tracheal epithelium. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6167–6171.

Yankaskas, J. R., Knowles, M. R., Gatsky, T. & Boucher, R. C. 1985 Persistence of abnormal chloride ion permeability in cystic fibrosis nasal epithelial cells in heterologous culture. *Lancet* i, 954–956.

Zengerling, S., Tsui, L.-C., Grzeschik, K.-H., Olek, K., Riordan, J. R. & Buchwald, M. 1987 Mapping of DNA markers linked to the cystic fibrosis locus on the long arm of chromosome 7. *Am. J. hum. Genet.* **40**, 228–236.

Discussion

S. V. Hodgson (*Guy's Hospital, London, U.K.*). From the point of view of using these linked probes for genetic counselling, I ask Dr Tsui on what grounds the diagnosis of CF had to be changed in the cases he mentioned, and how difficult he feels it is to be sure of the clinical diagnosis.

L.-C. Tsui. Before I answer Dr Hodgson's question, I have to state that I am not a clinician, so I do not know how sure one can be in CF diagnosis. From what I gathered from my colleagues, CF patients are usually very typical; the uncertainty comes only when they have very mild symptoms. This is what happened in two of our studied cases where, upon genetic analysis, the patients were found to be non-CF. In both cases, the patients showed very mild symptoms when they were diagnosed. For example, they did not appear to have any malabsorption problem. From the results of our genetic analyses with a series of DNA markers on both sides of the CF locus, these patients would have to have two cross-overs in one case and a double cross-over in another, within a small distance from CF. The odds were highly against these events. Therefore the children were asked to be re-examined and they were subsequently found to be non-affected. Although these errors had caused us a lot of trouble in our data interpretation and led us to speculate genetic heterogeneity in CF, I am glad they have a happy ending.

J.-J. Cassiman (*Centre for Human Genetics, University of Leuven, Belgium*). We examined 21 CF families in Belgium (133 individuals) by using the classical four to six probes. Our results show that in these families 50% of the children have CF (bias of ascertainment); there are two cross-overs, one in a sib and one in a CF child. The allele frequencies fit those published, except for 7C22 ($\chi^2 > 5$). The homozygote frequencies for the different haplotypes are as expected, but one-third of the mothers are homozygous. This phenomenon could be examined in large series.

L.-C. Tsui. We do not see any correlation between mothers of CF patients and homozygosity for markers spanning the CF locus in our 50 Canadian families, each with two or more affected children.

H. Sharma (71 *Barrack Road, Hounslow, U.K.*). Does Dr Tsui have any information on the gene for CF, such as putative protein?

L.-C. Tsui. No. As far as I know, the gene for CF has not been identified yet.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Note added in proof (18 *March* 1988). Estivill *et al.* (1987) reported the isolation of a candidate CF locus by selection for methylation-free islands. Subsequent analysis revealed that this gene encodes a protein that is related to the mouse *int-1* oncogene and the *Drosophila wingless* gene, hence given the name 'int-1 related protein' or IRP (R. Williamson, personal communication). Several DNA segments have been cloned from the IRP locus and some of them revealed frequent RFLPs in the caucasian population. Family studies showed that IRP is tightly linked to CF but several recombination events have been noted between the two, arguing against IRP being the CF gene. Strong linkage disequilibrium was observed for three of the RFLPs (detected by KM19, CS.7 and XV-2c), suggesting that the majority of chromosomes carrying the CF mutation are from a single mutational event. A long-range restriction map has also been established for a 4.5 megabasepair-region containing *D7S18* (formerly *D7S16*), *MET*, the IRP locus (*D7S23*) and *D7S8* (Poustka & Bates 1987).

More recently, we have identified a DNA segment that is closely linked to CF from our collection of DNA markers in the 7q31 region. Genetic and long-range physical mapping data indicated that this marker, *D7S340*, is located between *MET* and *D7S8*. The relation between *D7S340* and CF is currently being investigated.

References

Estivill, X., Farrall, M., Scambler, P. J., Bell, G. M., Hawley, K. M. F., Lench, N. J., Bates, G. P., Kruyer, H. C., Frederick, P. A., Stanier, P., Watson, E. K., Williamson, R. & Wainwright, B. J. 1987 A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. *Nature, Lond.* **326**, 840–845.

Poustka, A. & Bates, G. 1987 *Hum. Gene Mapping* **9**.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

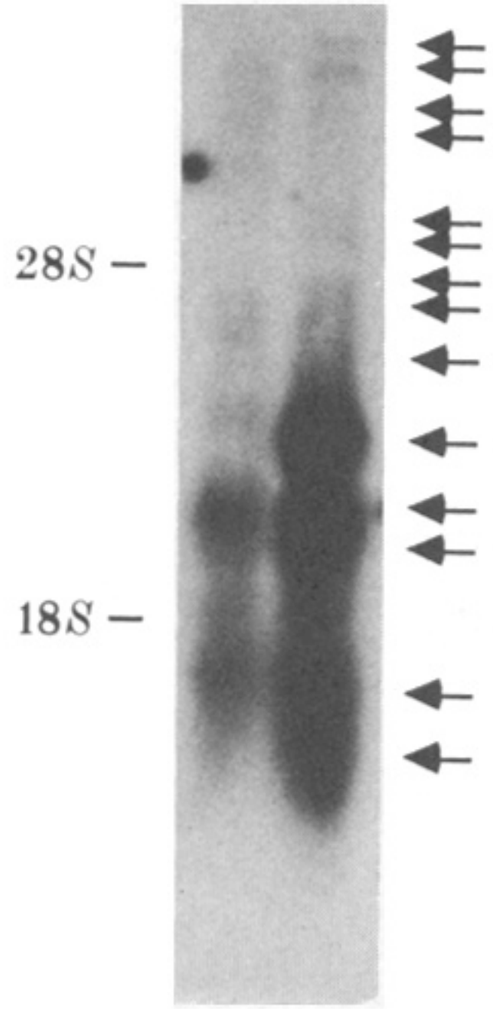

FIGURE 4. Subtraction hybridization analysis. Radioautograph shows the result of gel-blot hybridization of poly(A)$^+$ RNAs from cultured fibroblasts (left lane) and sweat-gland cells (right lane) with sweat-gland-specific ^{32}P-labelled cDNA probe pre-annealed with excess poly(A)$^+$ RNA from fibroblasts.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES