



---

Cystic Fibrosis Locus Defined by a Genetically Linked Polymorphic DNA Marker  
Author(s): Lap-Chee Tsui, Manuel Buchwald, David Barker, Jeffrey C. Braman, Robert Knowlton, James W. Schumm, Hans Eiberg, Jan Mohr, Dara Kennedy, Natasa Plavsic, Martha Zsiga, Danuta Markiewicz, Gita Akots, Valerie Brown, Cynthia Helms, Thomas Gravius, Carol Parker, Kenneth Rediker, Helen Donis-Keller  
Source: *Science*, New Series, Vol. 230, No. 4729 (Nov. 29, 1985), pp. 1054-1057  
Published by: American Association for the Advancement of Science  
Stable URL: <http://www.jstor.org/stable/1695332>  
Accessed: 28/07/2009 14:26

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aaas>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

<http://www.jstor.org>

## Cystic Fibrosis Locus Defined by a Genetically Linked Polymorphic DNA Marker

**Abstract.** A polymorphic DNA marker has been found genetically linked, in a set of 39 human families, to an autosomal recessive gene that causes cystic fibrosis (CF), a disease affecting one in 2000 Caucasian children. The DNA marker (called D0CRI-917) is also linked to the PON locus, which by independent evidence is linked to the CF locus. The best estimates of the genetic distances are 5 centimorgans between the DNA marker and PON and 15 centimorgans between the DNA marker and the CF locus, meaning that the location of the disease gene has been narrowed to about 1 percent of the human genome (about 30 million base pairs). Although the data are consistent with the interpretation that a single locus causes cystic fibrosis, the possibility of genetic heterogeneity remains. The discovery of a linked DNA polymorphism is the first step in molecular analysis of the CF gene and its causative role in the disease.

LAP-CHEE TSUI

MANUEL BUCHWALD

Department of Genetics, Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8 and

Department of Medical Genetics and Medical Biophysics, University of Toronto, Ontario

DAVID BARKER

JEFFREY C. BRAMAN

ROBERT KNOWLTON

JAMES W. SCHUMM

Department of Human Genetics, Collaborative Research, Inc., Lexington, Massachusetts 02173

HANS EIBERG

JAN MOHR

University Institute of Medical Genetics, Copenhagen, Denmark

DARA KENNEDY

NATASA PLAVSIC

MARTHA ZSIGA

DANUTA MARKIEWICZ

Department of Genetics, Hospital for Sick Children

GITA AKOTS

VALERIE BROWN

CYNTHIA HELMS

THOMAS GRAVIUS

CAROL PARKER

KENNETH REDIKER

HELEN DONIS-KELLER

Department of Human Genetics, Collaborative Research, Inc.

Cystic fibrosis (CF) is the most common serious genetic disorder in Caucasian children; it is inherited as an autosomal recessive trait, and it occurs approximately in one out of 2000 live births in North America (1). The patients suffer from chronic pulmonary disease and pancreatic enzyme insufficiency. They also show increased sweat electrolyte level which, although of little clinical consequence itself, is the basis of the most reliable diagnostic test for the disease. All these clinical symptoms are

consistent with CF being an exocrine disorder. Despite extensive research effort, the basic defect in CF remains unidentified. Therefore, the metabolic abnormalities on which the biochemical studies are based are probably secondary or tertiary consequences of the primary defect, making it difficult to proceed from the phenotype observed in CF patients and their cells to the gene defect.

The simple autosomal recessive mode of inheritance of cystic fibrosis and extensive analysis of first cousin marriages (2) suggested that cystic fibrosis might be the result of defects in a single gene. Genetic analysis by linkage with restriction fragment length polymorphisms (RFLP's) (3) seemed feasible with random polymorphic DNA probes and a relatively small number of nuclear families containing two or more affected children. Supporting this idea, a close linkage has been detected between CF and PON (4-6), a genetic determinant for serum paraoxonase activity (7). The PON locus, however, unlike DNA polymorphisms, does not lead directly to simple chromosome localization (4). Since the availability of a DNA marker closely linked to CF would be extremely useful for the molecular genetic analysis of the disease, we and others (8) have, over the past few years, examined a large number of RFLP markers. While these studies did not result in the discovery of linkage to CF, it is estimated that about 40 percent of the human genome has been excluded as being the site of the CF gene.

The chance of finding linkage depends on the number of probes tested and the degree of their polymorphism as well as the number and structure of families available. In order to facilitate such analyses of linkage, we isolated more than 200 new RFLP DNA markers with polymorphism information content (PIC) val-

ues (a measure of the degree of polymorphism) (3) of at least 0.3 (9). When applied to CF families as described here, we found that one of these is linked to the CF locus.

This DNA marker, provisionally designated as D0CRI-917, is defined by a human genomic fragment clone isolated from a Hae III-Alu I partial digest library in the lambda vector charon 4A (10). The recombinant phage, Lam4-917, contains about 18 kilobases (kb) of human DNA, which as expected from the method of isolation (11) contains no detectable repetitive sequences. It is one of 330 single-copy clones that were found to reveal RFLP's from a total of 1025 single-copy clones that have been isolated and tested for polymorphism in human DNA's digested with the enzymes Hind III and Hinc II (9, 12).

The restriction map of Lam4-917, as determined by the *cos*-oligonucleotide method (13) and by standard restriction analysis of subcloned Eco RI fragments, is shown in Fig. 1. The correspondence of the restriction sites in the clone with those in the human genome was confirmed by Southern gel-transfer hybridization analysis (14).

Human DNA's digested with Hind III and probed with Lam4-917 reveal allelic fragments of 6.3 kb (A1) or 5.3 plus 1.0 kb (A2) as well as constant fragments of 12.5 and 7.5 kb. This RFLP is due to the presence or absence of a Hind III site at the location indicated in Fig. 1. DNA samples from nuclear families reveal Mendelian segregation of the 6.3 kb band (A1) from the 5.3- and 1.0-kb bands (A2) (Fig. 2, left). The frequencies for A1 and A2 are 0.71 and 0.29, respectively, in a sample of 40 unrelated Caucasian individuals. These frequencies do not differ significantly from the frequencies of 0.73 (A1) and 0.27 (A2) found among 76 Caucasian parents of the CF families examined in this study. Family analysis of the Hinc II RFLP (Fig. 2, right) similarly reveals two clearly distinguishable alleles, B1 (4.3 kb) and B2 (2.3 and 1.8 kb) (constant fragments are 6.8, 5.1, and 3.9 kb) with respective frequencies of 0.48 and 0.52 in 40 randomly chosen individuals and 0.43 (B1) and 0.57 (B2) in 33 parents of CF families.

The PIC values (3) based on the combined gene frequency data from random individuals and CF parents are 0.32 for the Hind III polymorphism and 0.37 for the Hinc II polymorphism. To estimate the combined informativeness of the Hind III and Hinc II polymorphisms, we examined 58 individuals who were typed for both the Hind III and Hinc II RFLP.

Site haplotypes (A1B1, A1B2, A2B1, or A2B2) were inferred for 54 of these, allowing estimates of the site haplotype frequencies as 0.39 (A1B1), 0.44 (A1B2), 0.10 (A2B1), and 0.06 (A2B2). This result suggests that the Hinc II and Hind III RFLP sites are nearly randomly associated, that is, in linkage equilibrium, ( $\chi^2 = 1.25$ ,  $P > 0.2$ ). The combined PIC is approximately 0.57, and the observed heterozygosity was 62 percent.

The CF families in our study were two-generation families, each with both parents living and two or more affected children. The majority of the families were identified through the cooperation of the Canadian Cystic Fibrosis Foundation and of several CF clinics in Canada. These pedigrees, the diagnosis of CF patients, and the procedures for blood sample collection from participating individuals have been described (15). Additional CF families used in our study were obtained from the Human Genetic Mutant Cell Repository in Camden, New Jersey. DNA samples were prepared either directly from peripheral blood samples (16) or from lymphoblast cultures derived from individual members of each family (17).

To investigate the linkage relationship between *D0CRI-917* and *CF*, we examined samples from a total of 43 CF families for the inheritance pattern of the two RFLP's detectable with the Lam4-917 probe. Thirty-nine of them were informative for the analysis in that one or both of the parents in these families were heterozygous for either the Hind III or the Hinc II RFLP, or both. The calculation of the maximum likelihood and recombination fraction ( $\theta$ ) between the test marker and *CF* was performed by lod score analysis (18) with the LIPED computer program (19) (Table 1). If recombination frequencies are considered equal in male and female, a maximum likelihood estimate is obtained for  $\theta$  at 0.14 with a lod ( $z$ ) score of 3.96 (odds ratio of 9100:1). The confidence interval for  $\theta$  is between 0.07 and 0.25 (20). The recombination fraction is close to the frequency of recombination detected by counts of the number of apparent recombinant chromosomes divided by the number of informative chromosomes among CF children (13/101). However, a maximal lod score of 4.13 (odds ratio of 13,500:1) was obtained at male recombination fraction ( $\theta_M$ ) of 0.09 and female recombination fraction ( $\theta_F$ ) of 0.19 (21). Thus the score is much greater than 3 (odds ratio of 1000:1), the value accepted for proof of linkage (18).

A close linkage has been indicated

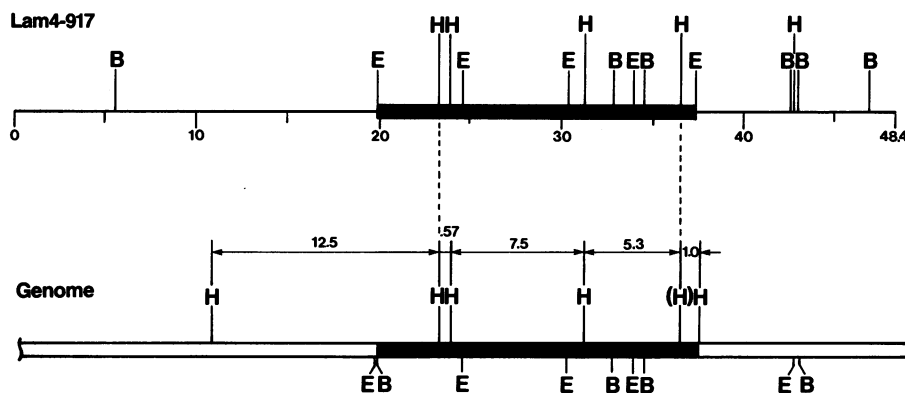


Fig. 1. Restriction maps of Lam4-917 and of the genomic DNA segment defined as the *D0CRI-917* locus. The 17.5-kb cloned human segment is represented by thick lines in both the Lam4-917 phage and the genomic representations. The letters B, E, and H indicate the sites for digestion with Bam HI, Eco RI, and Hind III, respectively. Numbers below the phage diagram indicate the distance, in kilobases, from the left end of Lam4-917; numbers in the genomic diagram indicate the size, in kilobases, of fragments generated by Hind III digestion and detected by Southern gel transfer hybridization analyses. The parentheses designate the polymorphic Hind III site revealed by probe Lam4-917.

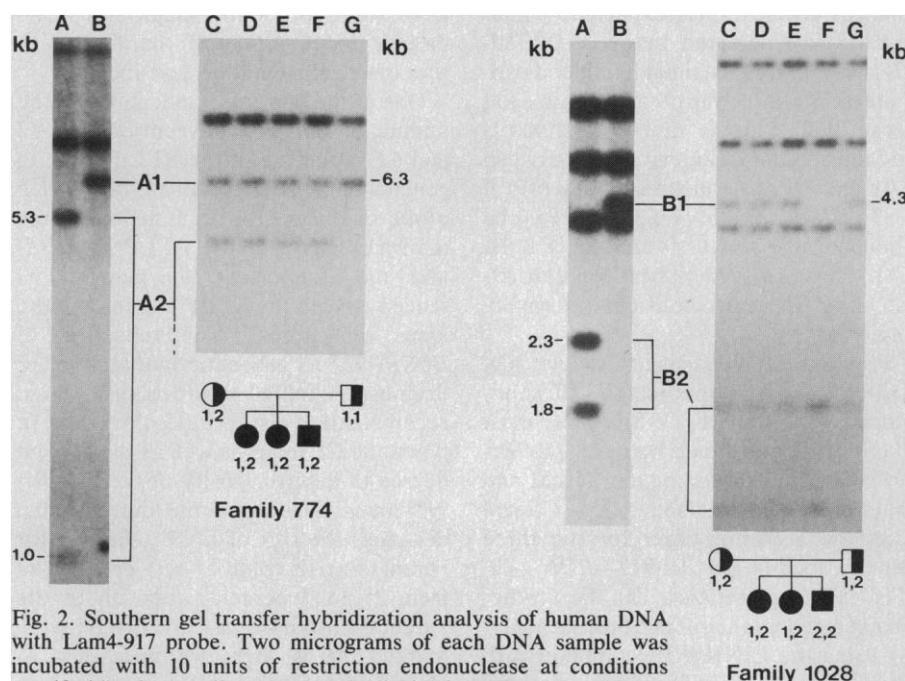


Fig. 2. Southern transfer hybridization analysis of human DNA with Lam4-917 probe. Two micrograms of each DNA sample was incubated with 10 units of restriction endonuclease at conditions specified by the supplier (New England Biolabs). The digested DNA samples were size-fractionated by electrophoresis in 0.8 percent agarose gels, transferred to Zetabind membranes (AMF Cuno, manufacturer) and hybridized with radioactive DNA probes as described (30).  $^{32}$ P-Labeled Lam4-917 probe was prepared by nick translation (31) of Lam4-917 DNA to a specific activity of approximately  $2 \times 10^8$  cpm/ $\mu$ g. After overnight hybridization with probe (approximately  $1 \times 10^5$  to  $2 \times 10^5$  cpm/cm $^2$ ) (32), the membranes were washed twice at room temperature in  $2 \times$  SSC (saline sodium citrate), and twice at 65°C in  $0.1 \times$  SSC containing 0.2 percent sodium dodecyl sulfate. Autoradiography was carried out at  $-70^\circ$  (Kodak XAR5 film with Dupont Lightning-Plus intensifying screens). (Left) Lanes A and B contain Hind III-digested DNA from randomly selected individuals. The DNA in lane A illustrates the A2 allele (variable fragment sizes of 5.3 and 1.0 kb) and the DNA in lane B reveals the A1 allele (variable fragment size of 6.3 kb). Lanes C to G contain Hind III-digested DNA samples from individuals in family 774 (only the 5.3-kb fragment of the A2 allele is shown). As indicated in the diagram of the pedigree below the autoradiogram, lanes C and G contain DNA sample from the mother and father, respectively, both carriers for cystic fibrosis. Lanes D, E, and F contain samples from two affected daughters and an affected son in this family, respectively. (Right) Lanes A and B contain Hinc II-digested DNA from randomly selected individuals. The DNA in lane A is characteristic of the B2 allele (variable fragment sizes of 2.3 and 1.8 kb), and the DNA in lane B reveals the B1 allele (variable fragment size of 4.3 kb). Lanes C to G contain DNA samples of individuals in family 1028. As indicated in the diagram of the pedigree below the autoradiogram, lanes C and G contain DNA samples from the mother and father, respectively, both carriers for cystic fibrosis. Lanes D, E, and F contain DNA samples from two affected daughters and an affected son in this family, respectively. Family 1028 includes a crossover between *CF* and *D0CRI-917*.

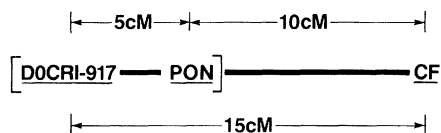


Fig. 3. Proposed linear order of genetic loci *D0CRI-917-PON-CF*. The brackets indicate that the given order as opposed to *PON-D0CRI-917-CF* is only favored by odds of 9:5.

between *CF* and *PON* (4). If the linkage between *CF* and *D0CRI-917* and that between *CF* and *PON* are both true, linkage should also be detectable between *D0CRI-917* and *PON*. Since paroxonase typing data were available for 31 of the Canadian families used in our study (5, 6, 22), it was possible to examine the linkage relationship between *D0CRI-917* and *PON*. Eleven families were found to be informative for this analysis (Table 1). As expected, a tight linkage was detected between *D0CRI-917* and *PON*. A maximal likelihood estimate for  $\theta$  at 0.05 was obtained with a lod score of 5.01 (odds ratio of 102,000:1) and the confidence interval was between 0.01 and 0.17. A maximal lod score of 5.13 (odds ratio of 132,000:1) can be obtained at  $\theta_M$  of 0.10 and  $\theta_F$  of 0.04. (21). Thus the linkage between *D0CRI-917* and *PON* is also considered formally demonstrated.

Previous studies suggested that *CF* and *PON* are approximately 10 centimorgans (cM) apart (4). Since the most likely genetic distance between *D0CRI-917* and *CF* is approximately 15 cM and that for *D0CRI-917* and *PON* is 5 cM (23), the apparent order for the three genetic loci is *D0CRI-917-PON-CF* (Fig. 3). This relation can be further investigated by multilocus linkage analysis using the LINKAGE computer program (24). The result of this analysis

showed that the relative odds of orders *D0CRI-917-PON-CF*, *CF-D0CRI-917-PON*, and *D0CRI-917-CF-PON* are 4500:2500:1 (Fig. 4). While the relative odds calculation does not greatly favor the order *D0CRI-917-PON-CF* over the order *CF-D0CRI-917-PON*, it clearly indicates that *CF* is not located between *D0CRI-917* and *PON*.

The discovery of a DNA marker closely linked to the cystic fibrosis locus effectively eliminates 99 percent of the human genome from consideration as the site for *CF*, since the marker is approximately 15 cM from the *CF* locus and the size of the human genome is approximately 3000 cM (25). This marker (*D0CRI-917*) is also linked to a serum isoenzyme (*PON*) recently shown independently to be linked to the cystic fibrosis locus. Our results with *CF* represent the third case of linkage of a random DNA RFLP probe to a major autosomal disease locus (26) and the first to a recessive autosomal disease locus.

One of the possible applications of the finding of linkage between *D0CRI-917* and *CF* is the use of the RFLP probe in genetic diagnosis. Usefulness of a RFLP probe in diagnosis is a function of the informativeness of the RFLP locus (3) and, more important, the genetic distance between the RFLP marker and the locus of interest. The reliability of *D0CRI-917* as a genetic marker for *CF* diagnosis is limited by the relatively high recombination distance, 15 cM, between it and the *CF* locus as well as the current degree of heterozygosity of the *D0CRI-917* locus. Based on the current map distance, the risk of a *CF* offspring for parents at risk could at best be reduced from 25 to 7 percent when all of the parental chromosomes are distinctively marked. It is therefore necessary to identify polymorphic DNA markers that

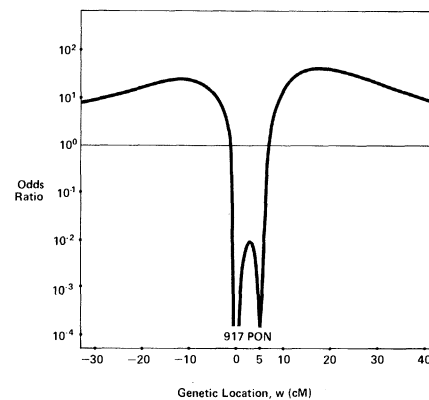


Fig. 4. Likelihood of the map location of *CF* with respect to *D0CRI-917* and *PON*. Data are selected from 11 *CF* families which are informative for both *D0CRI-917* and *PON* polymorphisms. Calculations were carried out using the LINKAGE computer program assuming *D0CRI-917* and *PON* were 5 cM apart by varying the distance between *CF* and the two markers (23). Horizontal axis, genetic distance ( $w$ ) from *D0CRI-917*; vertical axis, odds ratio for location of *CF* at  $W_{CF}$  versus *CF* at infinite distance (that is, no linkage).

are closer or flank the cystic fibrosis locus.

Genetic heterogeneity in *CF* has been considered as one of the possible explanations for the high frequency of the disease (27). Previous evidence was consistent with the assumption that *CF* is due to defects in a single gene (2). The availability of a polymorphic DNA marker genetically linked to *CF* allowed us to examine the possibility of the existence of other genetic loci for cystic fibrosis. If there were more than one *CF* locus, families which carry mutations in the *CF* locus linked to *D0CRI-917* would appear as one group (the linked group) while families with mutations in other *CF* loci would form another group (the unlinked group). The test for heterogeneity was performed by the admixture method (28) with the present *D0CRI-917-CF* linkage data from 39 families with the HOMOG computer program (29). The result of the analysis was consistent with the existence of a single *CF* locus responsible for the disease in all affected families. However, the size of the present data set is too small to analyze heterogeneity with confidence. The investigation of a larger family set which is in progress and the use of additional closer and flanking markers should allow resolution of this issue.

The estimated distance of 15 cM means that about 15 million base pairs lie between the *D0CRI-917* and *CF* loci. This distance precludes the "chromosome walking" strategy to identify the *CF* gene. Nevertheless, the resolution of

Table 1. Linkage relationships of *D0CRI-917-CF* and *D0CRI-917-PON*.

Loci	Number of informative families	LOD (z) scores at recombinant fractions ( $\theta$ ) of:									
		0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	
<i>D0CRI-917-CF</i>	36 (Can)	-6.02	0.98	2.84	3.20	2.96	2.44	1.79	1.13	0.55	
	3 (HGCMR)	0.14	0.69	0.79	0.75	0.66	0.53	0.39	0.25	0.12	
	39 (Total)	-5.88	1.67	3.63	3.95	3.62	2.97	2.18	1.38	0.67	
		$\hat{\theta} = 0.14$ ( $z = 3.96$ ); confidence interval: 0.07-0.25; $\hat{z} = 4.13$ ( $\theta_M = 0.09$ ; $\theta_F = 0.19$ )									
<i>D0CRI-917-PON</i>	11 (Can)	4.27	5.01	4.78	4.28	3.66	2.97	2.25	1.51	0.81	
			$\hat{\theta} = 0.05$ ( $z = 5.01$ ); confidence interval: 0.01-0.17; $\hat{z} = 5.12$ ( $\theta_M = 0.10$ ; $\theta_F = 0.04$ )								

the CF gene location to less than 1 percent of the human genome followed by chromosomal localization will facilitate the subsequent isolation and characterization of the gene that causes cystic fibrosis.

#### References and Notes

- R. C. Talamo, B. J. Rosenstein, R. W. Benninger, in *The Metabolic Basis of Inherited Disease*, J. B. Stanbury *et al.* Eds. (McGraw-Hill, New York, ed. 5, 1983), pp. 1889-1917.
- D. M. Danks, P. D. Phelan, C. Chapman, *Am. J. Hum. Genet.* **36**, 1389 (1984); G. Romeo *et al.*, *ibid.* **37**, 338 (1985).
- Polymorphism information content (PIC) is a mathematical measure of the usefulness of a polymorphism in linkage studies. It is the probability of identifying in any given offspring the parental chromosome contribution. For example, a two-allele locus with each allele present at 50 percent frequency in the population has a PIC value of 0.37. D. Botstein, R. L. White, M. Skolnick, R. W. Davis, *Am. J. Hum. Genet.* **32**, 314 (1980).
- H. Eiberg, J. Mohr, K. Schmiegelow, L. S. Nielsen, *Clin. Genet.* **28**, 265 (1985).
- H. Eiberg *et al.*, abstract in the 8th Human Gene Mapping Workshop, 4 to 10 August 1985, Helsinki, Finland, *Cytogen. Cell Genet.*, in press.
- K. Schmiegelow *et al.*, *Clin. Genet.*, in press.
- M. Geldmacher-von Mallinckrodt, *Hum. Genet. Suppl.* **1**, 65 (1978); H. Eiberg and J. Mohr, *Ann. Hum. Genet.* **45**, 323 (1981); R. F. Mueller *et al.*, *Am. J. Hum. Genet.* **35**, 393 (1983); H. W. Eckerson, C. M. Wytte, B. N. La Du, *ibid.*, p. 1126.
- L.-C. Tsui *et al.*, in *Cystic Fibrosis: Horizons*, D. Lawson, Ed. (Wiley, New York, 1984), p. 166; L.-C. Tsui *et al.*, *Am. J. Hum. Genet.* **36**, 156S (1984); M. Buchwald *et al.*, abstract in the 8th Human Gene Mapping Workshop, 4 to 10 August 1985, Helsinki, Finland; P. Scambler *et al.*, *Hum. Genet.* **69**, 250 (1985); L.-C. Tsui *et al.*, *Cytogenet. Cell Genet.*, in press; P. Scambler *et al.*, *ibid.*, in press; P. Watkins *et al.*, *Am. J. Hum. Genet.* **37**, A181 (1985).
- J. Schumm *et al.*, abstract in the 8th Human Gene Mapping Workshop, 4 to 10 August 1985, Helsinki, Finland; J. Schumm *et al.*, *Cytogenet. Cell Genet.*, in press.
- R. M. Lawn *et al.*, *Cell* **15**, 1157 (1978).
- A. Wyman and R. White, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6754 (1980).
- Nineteen other enzymes, including Alu I, Apa I, Bam HI, Bcl I, Bgl I, Bst XI, Eco RI, Hinf I, Hae III, Mbo I, Msp I, Nco I, Pst I, Pvu II, Rsa I, Taq I, Xba I, and Xmn I do not reveal frequent polymorphism at this locus.
- H. Rackwitz, G. Zehetner, A. Frischauf, H. Lehrach, *Gene* **30**, 195 (1984).
- E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
- M. Buchwald, L.-C. Tsui, M. Corey, in preparation.
- N. Blin and D. W. Stafford, *Nucleic Acids Res.* **3**, 2303 (1976).
- M. A. Anderson and J. F. Gusella, *In Vitro* **20**, 856 (1984).
- The method of lod score analysis was described by N. E. Morton [*Am. J. Hum. Genet.* **7**, 277 (1955)]; based on the recessive mode of inheritance of cystic fibrosis, the affected children in each family were assumed to have inherited two copies of the defective CF gene, one from each parent (obligatory heterozygotes). The unaffected sibs in the families were also included in the analysis; although their CF status was undetermined, the algorithm considered them to be either homozygous normal or heterozygous carrier.
- J. Ott, *ibid.* **26**, 588 (1974).
- The confidence interval was estimated by taking the  $\theta$  values which corresponded to a lod ( $z$ ) score value of ( $z - 1$ ), as recommended at the 8th Human Gene Mapping Workshop, 4 to 10 August 1985, Helsinki, Finland.
- In accordance with the recommendation at the 8th Human Gene Mapping Workshop, 4 to 10 August 1985, Helsinki, Finland, a maximal lod score was calculated for males and females separately. However, because the difference in lod score is very small, reflecting the fact that the total number of crossovers is only 13, no statistically significant inference can be drawn about differences in recombination in male and female meioses.
- H. Eiberg, J. Mohr, L.-C. Tsui, M. Buchwald, unpublished observations.
- The genetic distance was estimated from Haldane's [J. B. S. Haldane, *J. Genet.* **8**, 299 (1919)] function:  $d = -0.5 \ln(1 - 2r)$ .
- G. M. Lathrop, J. M. Lalouel, C. Julier, J. Ott, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3443 (1984).
- J. H. Renwick, *Br. Med. Bull.* **25**, 65 (1969).
- J. F. Gusella *et al.*, *Nature (London)* **306**, 234 (1983); S. T. Reeders *et al.*, *ibid.* **317**, 542 (1985).
- M. W. Thompson, in *Perspectives in Cystic Fibrosis*, J. M. Sturgess, Ed. (Canadian Cystic Fibrosis Foundation, Toronto, 1980), pp. 281-291.
- C. A. B. Smith, *Ann. Hum. Genet.* **27**, 175 (1963); S. E. Hodge, C. E. Anderson, K. Neiswanger, R. S. Sparkes, D. L. Rimoim, *Am. J. Hum. Genet.* **35**, 1139 (1983).
- J. Ott, *Ann. Hum. Genet.* **47**, 311 (1983).
- D. Barker, T. Holm, R. White, *Am. J. Hum. Genet.* **36**, 1159 (1984).
- P. W. J. Rigby *et al.*, *J. Mol. Biol.* **113**, 237 (1977).
- T. Maniatis, E. F. Fritsch, J. Sambrook, in *Molecular Cloning—A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), pp. 383-389.
- We thank David Botstein and Ronald Worton for discussions and comments on the manuscript; David Botstein for his support in establishing the RFLP's mapping effort and its applications at CRI; the CF families who contributed blood samples and the Canadian CF clinics and their staff who assisted in the identification of the families in the collection of samples. Supported by the Canadian Cystic Fibrosis Foundation (CCFF), the Sellers Fund of the Hospital for Sick Children, Toronto (L.-C.T. and M.B.), and the National Institutes of Health.

30 September 1985; accepted 9 October 1985

## Retroviral Vector-Mediated Gene Transfer into Human Hematopoietic Progenitor Cells

**Abstract.** *The transfer of the human gene for hypoxanthine phosphoribosyltransferase (HPRT) into human bone marrow cells was accomplished by use of a retroviral vector. The cells were infected in vitro with a replication-incompetent murine retroviral vector that carried and expressed a mutant HPRT complementary DNA. The infected cells were superinfected with a helper virus and maintained in long-term culture. The production of progeny HPRT virus by the bone marrow cells was demonstrated with a colony formation assay on cultured HPRT-deficient, ouabain-resistant murine fibroblasts. Hematopoietic progenitor cells able to form colonies of granulocytes or macrophages (or both) in semisolid medium in the presence of colony stimulating factor were present in the nonadherent cell population. Colony forming units cloned in agar and subsequently cultured in liquid medium produced progeny HPRT virus, indicating infection of this class of hematopoietic progenitor cell.*

**HARRY E. GRUBER**  
**KIM D. FINLEY**  
**ROBERT M. HERSHBERG**  
**SCOTT S. KATZMAN**  
**PAUL K. LAIKIND**  
**J. EDWIN SEEGMILLER**  
*Department of Medicine,*  
*University of California, San Diego,*  
*La Jolla, California 92093*  
**THEODORE FRIEDMANN**  
**JIING-KUAN YEE**  
**DOUGLAS J. JOLLY**  
*Department of Pediatrics,*  
*University of California, San Diego*

infect and genetically transform virtually 100 percent of recipient cells in a population of susceptible cells. Completely helper-free transmissible vector can be produced by transfection of packaging-defective viral DNA constructs that supply helper functions in *trans* (3).

We have previously used retroviral vectors carrying the complementary DNA (cDNA) (2) for human hypoxanthine phosphoribosyltransferase (HPRT; E.C. 2.4.2.8) (4), to demonstrate the *in vitro* infection of murine bone marrow cells and the partial repopulation of the bone marrows of whole mice with these genetically transformed cells (5). Other groups have obtained similar results with other retroviral vectors (6) and have also shown that murine pluripotential hematopoietic progenitor cells can be infected by such vectors.

The ability to insert genes functionally and efficiently into human cells (7) allows consideration of somatic cell manipulation as therapy for certain kinds of human disease (8). One potential pathway toward such an application in humans would be the genetic, and therefore phenotypic, alteration of a self-renewing stem cell population in a readily accessible major organ such as the human bone