Approaches to localizing disease genes as applied to cystic fibrosis

Michael Dean*, Mitchell L.Drumm[1], Claudia Stewart, Bernard Gerrard, Anjanette Perry, Noriko Hidaka[1], Jeffery L.Cole[1], Francis S.Collins[1] and Michael C.Iannuzzi[1]
Biological Carcinogenesis and Development Program, Program Resources, Inc., NCI-Frederick Cancer Research Facility, Frederick, MD 21701 and [1]Howard Hughes Medical Institute, University of Michigan, Ann Arbor, MI 48109, USA

Received September 6, 1989; Revised and Accepted November 30, 1989

ABSTRACT

Using chromosome jumping and walking and restriction fragment length polymorphism (RFLP) analysis, we have defined the region which must contain the cystic fibrosis gene. DNA segments spanning approximately 250 kb in the direction of the gene were isolated and used to identify several new polymorphisms informative in cystic fibrosis families. These RFLPs include a highly polymorphic, CA/GT repeat, and a 10 bp insertion uncovered using the polymerase chain reaction. By analyzing a family with a recombination near the gene, we can exclude this region as containing the mutation. Data on the extent of linkage disequilibrium of these markers provides additional information on where the gene is located.

INTRODUCTION

The gene responsible for cystic fibrosis (CF) has been mapped to chromosome 7q21−7q31 by genetic linkage analysis (1−6). Additional sequences have been cloned in this region by chromosome mediated gene transfer (7), saturation cloning of fragments from a chromosome-specific library (8), chromosome walking, and chromosome jumping, which allows the directed cloning of sequences at a defined distance (9−12). RFLP analysis of families that show recombination between the markers and the disease, and long-range restriction mapping by pulsed-field gel electrophoresis, have been used to locate the CF gene to a region spanning approximately 700 kb, flanked by MP6d and W32 (12−16). Few markers obtained from the D7S8 side of the CF gene have been available for study and only one family that is recombinant between D7S8 and CF has been reported (CF 1380). We used chromosome jumping and walking in the region between D7S8 and CF to isolate additional RFLP markers and developed strategies for increasing the informativeness of these markers in the recombinant family using the polymerase chain reaction and CA/GT repeats.

METHODS

Isolation of chromosome jumping clones

The screening of the jumping library (9), analysis of jumping clones, isolation and mapping of genomic clones, preparation of probes and identification of RFLPs has been described previously (12).

Polymerase chain reaction

One μg of genomic DNA was mixed with 20 ng of each W46 primer (5′ TGGAGATGTAGAGTGGT, 5′ GATCAGAAAGCACTATTTCAG), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin, 10 mM Tris, pH 8.3, and 5 μ TaqI polymerase U.S. Biochemicals) in a final volume of 100 μl. The samples were denatured for 7 min at 94° and treated for 25 cycles of 94°, 30 sec, 55°, 1 min, 72°, 2 min in a Perkin-Elmer DNA Thermal Cycler. Ten μl of sample was digested with Rsa I and resolved on a 15 cm non-denaturing 5% acrylamide gel, and stained with ethidium bromide.

The W30 CA/GT repeat was amplified using the conditions described by Litt and Luty (18) using the primers 5′ AAGGCCCATCTTCAGTAG, 5′ TTCTCACTCCTTTACTAGT, and the products separated on 8% DNA sequencing gels.

RESULTS

We performed successive rounds of chromosome jumping and walking, starting from the W32 locus (D7S424) to clone sequences closer to the CF gene. All of the jumps traversed a distance greater than 35 kb. The average length of 20 jumps isolated on both sides of the gene was 75 kb. Approximately 140 kb of DNA has been cloned and the jumps span about 250 kb in this region as determined by pulsed-field gel electrophoresis. The map of the loci isolated is shown in Figure 1.

The details of the walking and jumping are as follows: J35 was isolated by screening with a probe from the 5′ end of a 10 kb fragment in W32. Six independent phage clones were isolated using J35 as a probe. J35 contains 2.4 kb of a 2.5 kb genomic

* To whom correspondence should be addressed

Eco RI fragment and was therefore difficult to orient. The orientation was eventually determined by detailed mapping of the jumping clone and insert. The phage extending furthest in the 5′ direction contains a stretch of 6 kb which is present in about 20 copies in the human genome, but is not conserved in rodent DNA. Several of the fragments detected by this probe appeared to be polymorphic, but only one or two mapped to chromosome 7. Partial sequencing of this region failed to reveal any homology with sequences in the Genbank nucleotide sequence data base (data not shown). An attempt to walk through this region resulted in the isolation of several phage clones which did not map back to chromosome 7. A 1.5 kb fragment from the end of a 7 kb Eco RI fragment of W35 was used to screen the jumping library again, and J46 was isolated. J46 contains a portion of an L1 repeat and required competition with human DNA before use as a probe or to screen libraries (data not shown). Over 32 kb of overlapping phage clones were isolated at W46 and a 5.8 kb Eco RI fragment was chosen to continue jumping. Two independent jumps were isolated from W46, and designated J30 and J19. They were both used to screen lambda libraries and the clones overlapped, with J30 being 5′ to J19, about 10 kb apart (Figure 1). Over 60 kb of overlapping clones were isolated at W30/19, and a 4.5 kb Eco RI fragment was used to isolate J91. Five independent phage clones were isolated at W91 encoding 15 kb of DNA.

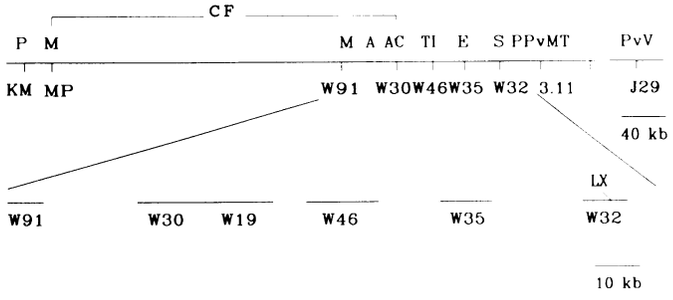

Figure 1. Map of the Cystic Fibrosis Locus. A, Diagram of the CF region showing the location of the gene as determined by the analysis of families displaying recombination. The approximate position of the jump clones was determined by PFGE (data not shown). In our conversion, J refers to the insert of the jumping clone and W to the genomic sequences isolated in that region. The symbols above the solid line indicte the presence of RFLPs. Data for KM 19 (D7S23) and MP6d (D7S399) have been reported previously (14). A, Acc I; CA, CA/GT repeat; E, Eco RI; I, insertion/deletion; M, Msp I; P, Pst I; Pv, Pvu II; S, Sac I; T, Taq I; V, Eco RV. B, Map of the overlapping genomic regions isolated at each locus. L and X indicate the location of Sal I and Xho I restriction sites, respectively.

At each locus we searched for the presence of rare cutting restriction sites and RFLPs, and used probes to hybridize to pulsed-field blots. The W32 (D7S424) locus contains Sal I and Xho I sites, but no other clones contained rare enzymes sites. Most probes tested revealed genomic Msp I and Taq I fragments of 5−20 kb, indicating that this region has a low abundance of CpG base pairs. This is consistent with the long-range restriction maps of this region (17).

Single-copy clones from each locus were used to screen for RFLPs by hybridizing to DNA from 9 unrelated, mostly Caucasian individuals. DNA was digested with 10−40 enzymes and putative RFLPs were tested on larger population samples; segregation was analyzed in CF or normal Caucasian pedigrees. The p35ES3.0 probe (D7S431) detects a low frequency Eco RI polymorphism, which is in a high degree of linkage disequilibrium with the Taq I RFLP at the D7S8 locus (Tables 1 and 2). Additional very rare Nco I and Taq I alleles were detected, principally in Black individuals, but these probes were uninformative in all parents tested (data not shown). The p46E3.6 (D7S432) clone revealed a Taq I RFLP whose alleles have frequencies of 0.46 and 0.54. This probe contains several repetitive sequences and even with competition using excess human DNA does not reliably yield readable blots. Data obtained from 48 unrelated individuals showed that this RFLP splits up haplotypes of previously isolated markers on this side of the CF gene; 43% of individuals uninformative for markers in the D7S8 locus are informative for this marker (data not shown). The probe p30E6.5 (D7S433) revealed a polymorphism with Acc I; the frequency of the alleles being 0.78 and 0.22 (Table 1). This RFLP was also informative in several individuals uninformative for D7S8 (data not shown). In addition, a relatively rare Msp I RFLP (0.92/0.08) was detected by J91. These new markers increase the heterozygosity on the telomeric side of the gene by over 20%.

Two polymorphisms detectable only by the polymerase chain reaction (PCR) were also found in this region. An insertion of about 10 bp was detected in the W46 locus upon amplification of a 900 bp segment (Figure 2). This RFLP was shown to segregate in 5 CF families and was in significant linkage disequilibrium with the Taq I RFLP of D7S8 (Table 2). Upon sequencing a fragment in the W30 locus we identified a region which contains a CA/GT repeat (Figure 3,4). Since these repeats have recently been shown to be polymorphic (18,19), we used flanking primers to amplify this region in genomic DNA. At least

TABLE 1 RFLPs Linked to Cystic Fibrosis

Locus	Probe	Enzyme	Alleles	Size (kb)	Freq.	# Chrom.
D7S434	J91	Msp I	1	2.5	0.08	180
			2	1.4+1.1	0.92	
D7S433	p30E6.5	Acc I	1	3.0	0.78	78
			2	2.5	0.22	
		CA/GT Repeat		Greater than 5 alleles		
D7S432	p46E3.6	Taq I	1	4.0	0.46	80
			2	2.5	0.54	
		Ins.	1	0.11	0.08	88
			2	0.10	0.92	
D7S431	p35ES3.0	Eco RI	1	13	0.11	109
			2	6	0.89	

The frequency (Freq.) of each allele was determined by analyzing both normal and CF chromosomes (Chrom) in unrelated Caucasians. Ins. refers to an insertion/deletion RFLP detected by digesting amplified DNA with Rsa I.

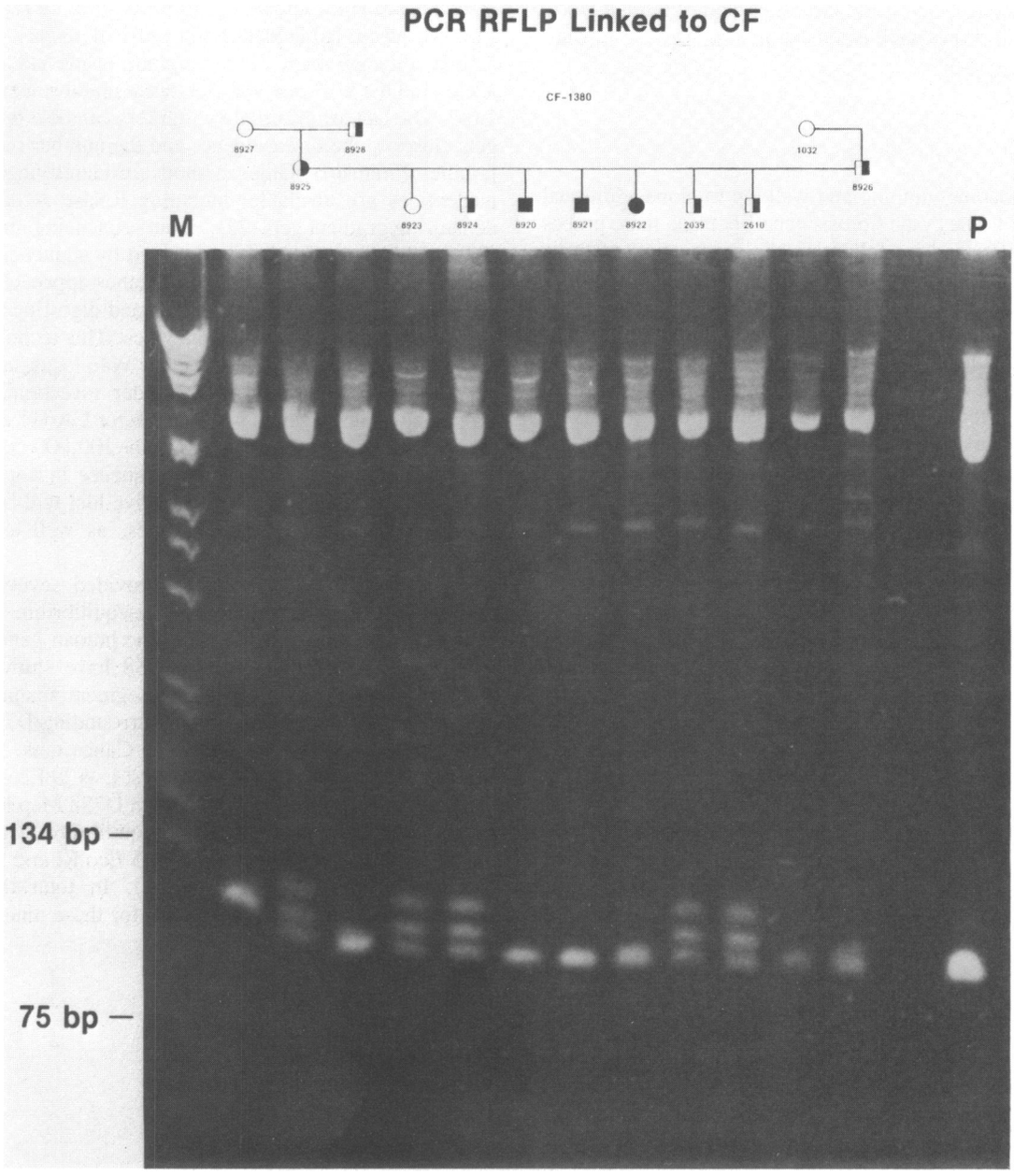

Figure 2. RFLP in the W46 Locus Detected by Amplification and Digestion with Rsa I. Primers that amplify a 0.9 kb DNA fragment were used on DNA from individuals of pedigree 1380. The amplified DNA was digested with Rsa I, separated on a 5% acrylamide gel, and stained with ethidium bromide. The upper band observed in the heterozygotes is a heteroduplex of the two alleles. Genotypes are displayed in Figure 5. M, marker; P, plasmid.

5 alleles have been detected in 8 unrelated parents of CF pedigrees, with virtually every parent informative (data not shown). Due to the presence of background bands we have been unable to identify all the alleles at this locus, however, the genotype of CF and normal children within a family could usually be distinguished. We are attempting to improve the resolution of these alleles so that this polymorphism can be used for diagnostic and population studies. We conclude that the polymerase chain reaction can be used to identify RFLPs undetectable by standard techniques.

The CF pedigree 1380, collected at the University of Utah, contains three affected children (8920−2), one of whom (8920) is an obligate recombinant, in the paternal chromosomes, between D7S8 and CF (Figs. 2,3,5; ref. 20). This individual was also shown to be recombinant with a marker at the W32 locus (Fig. 5, 12), and we have continued to use it to exclude the CF mutation from segments of DNA in this region. The father of this pedigree (8926) was uninformative for the W91 Msp I, W32 Eco RI, and W46 insertion RFLPs, but is informative for the W46 Taq I and W30 Acc I RFLPs. As shown in Figure 5, W46 Taq I and W30 Acc I are recombinant in this family. Using the W30 CA/GT repeat, we could also see that the recombinant child has inherited a different paternal chromosome then his other affected siblings (Figure 3,5). These data demonstrate that we have not crossed this recombination, and exclude the region between W30 and W32 as containing the mutation responsible for cystic fibrosis. None of these markers displays levels of linkage disequilibrium with CF that are dramatically different from those at D7S8 (Table

2). To our knowledge, this is the largest stretch of human DNA characterized that shows such extensive marker-marker linkage disequilibrium.

DISCUSSION

We used chromosome jumping and walking to clone additional sequences closer to the cystic fibrosis gene. By using these probes to identify RFLPs, and by following the segregation of these polymorphisms in a family recombinant with CF, we have been able to narrow the location of CF by about 200 kb. Genetic data places the CF gene between MP6d and W30, which span 500 kb. Chromosome jumping has advantages over the use of walking alone to travel large distances along the length of a chromosome. We have found several regions where screening both phage and cosmid libraries failed to result in a clone that extended in the desired direction. The presence of large repetitive sequences could also hamper walking efforts. While we do not know the length or nature of the low copy repetitive element encountered in the W35 locus, this region probably could have been traversed with a cosmid clone or a high stringency approach. However, jumping over this sequence allowed us to proceed more rapidly past this site. We succeeded in identifying an RFLP in each of the regions cloned, although many of the polymorphisms had low frequency alleles and the W91 and W35 loci were uninformative in pedigree 1380. The lack of informative markers can be a problem as one gets closer to the disease gene, and the number of recombinant families diminishes. Other methods for identifying polymorphic nucleotides are available, including RNase A digestion (21), denaturing gradient gels (22), direct sequencing and oligotyping (23), and blotting fragments resolved by sequencing gels (24). We have developed and applied another approach, amplifying segments up to 1 kb with the PCR, and digesting the amplified DNA with frequently cutting enzymes. This technique revealed the presence of a small insertion in the W46 locus, and the general applicability of this strategy is under investigation. A more promising strategy would be to search for CA/GT repeats. Since these elements are present in 50,000−100,000 copies, there is a good chance of finding such a sequence in a given genomic clone (18,19). Such highly informative loci will be very useful in characterizing individual families, as well as for genetic diagnosis.

TABLE 2 Linkage Disequilibrium of Markers between D7S8 and CF

Markers	# Chrom.	Δ	X^2
W35 Eco vs. CF	20	0.15	0.4
W46 Ins. vs. CF	88	0.04	0.1
W30 AccI vs. CF	112	0.13	1.9
W91 Msp vs. CF	118	−0.04	0.2
W46 Ins. vs. D7S8 Taq I	82	−0.47	18
W35 Eco vs. D7S8 Taq I	25	−1.0	25

Standardized linkage disequilibrium values were calculated as described (12). In this population the value for D7S8 Msp I vs. CF is 0.02 (13).

The cystic fibrosis locus has provided several interesting examples of long range linkage disequilibrium (non-random association of genetic markers) in the human genome. Both of the flanking markers *met* and D7S8 have shown significant associations with CF (1,4,6,12), a segment spanning 2−4 cM and 2000 kb. The polymorphisms surrounding D7S8 also show dramatic associations, in unrelated Caucasians. Four RFLPs spanning 300 kb (W32 Sac I, D7S8 Pst I, W29 Eco RV and W29 Pvu II) are all in tight association with D7S8 Msp I (12,25). The D7S8 Taq I RFLP is in linkage disequilibrium with three other polymorphisms (W46 insertion, W35 Eco RI and D7S8 Pvu II) spanning about 250 kb (Table 2). In total the region is approximately 450 kb in length and for these nine RFLPs (512

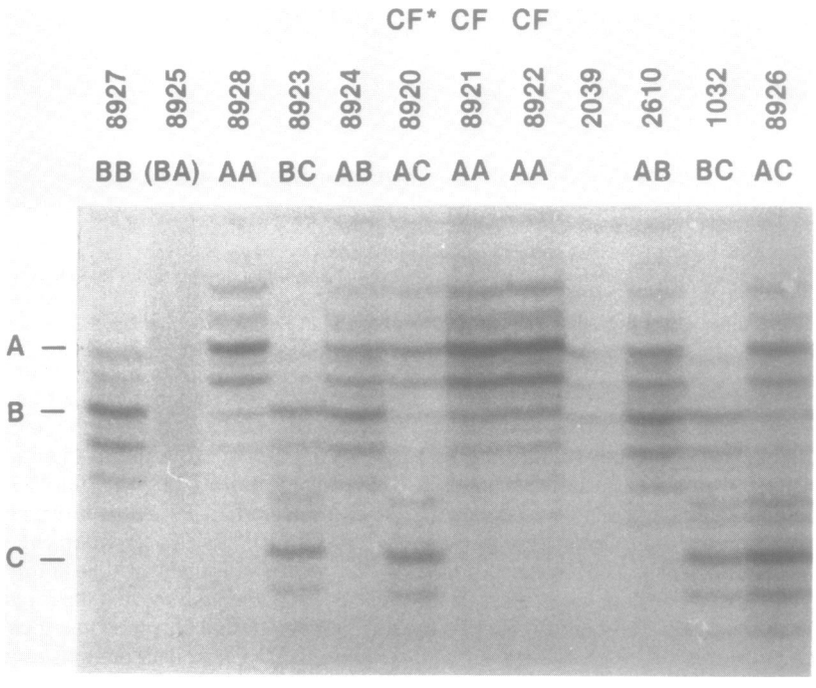

Figure 3. Segregation of the W30 CA/GT repeat. 1 μg of DNA from each individual was amplified with primers flanking the repeat and the DNA separated on an 8% sequencing gel. See Figs. 2 and 4 for relationships between individuals. Genotype of individual 8925 is derived. CF* represents the patient (8920) previously shown to be recombinant between D758 and CF and here shown to be recombinant for W30 (see Figure 5).

```
        10                  30                  50
TCAAAGCTATCACTTGTATCCCTCTCCCTGCTCAAGGCCCATCTTCAGTAGACTAAACAA
        70                  90                  110
ATACCAGAATCACACACACACACACACACACACACACACACACACACACACACACAAT
        130                 150                 170
CCTGTATACTAGTAAAGGAGTGAGAAGGAAAATACTCTAAGATAAAAGGATAATAAGTTT
        190                 210                 230
TTACACAGTAATTCTGTTTAACAGGTAAATAAATGCAGGATAACAGAAGGCAGAAACCTG
        250                 270                 290
ATTTATTGCAGAAAACGGATGGAAAAAATTCTAAACACAACAGATGAACTCCAGAGCCCT
        310                 330                 350
GCTGTGAAAACAAGGAATTTCAATTCTGGTCAAATTCTCAACAGTGGCATAACTCTTAGG
        370                 390                 410
AAATCTCAGACTTCTCAGTCTTCAGATACTTTCTGTCTAGGTCTACTTTTCTGCATATTC
        430                 450                 470
CTAGATTCTTATCTATTCAGACATGCATCTGGGTGACAGTCATTTTAAGTCAAGATAGGA
        490                 510                 530
AGGATTGTCTTTCCTTATGTCTCCAGGCCCTTAGGGATTAAATAGACTTTATTGGTTATA
        550                 570                 590
TAACTGTCCTTTAACCTGCAACTTCGGGGCCACAGCCTTTTTTTGAGGCTTACTCTGGGA
```

Figure 4. Partial Nucleotide Sequence of a Hind III Fragment in the W30 Region. The position of the CA repeat 1s nucleotides 71–118. EMBL Accession number X16414.

met		2|1		2|2	2|1	1|1	1|1	1|1	2|1			2|1
E6	2|2	2|2	1|2	2|1	2|2	2|2	2|2	2|2	2|2	2|2		1|1	1|2
W3Hin	2|1	1|1	1|1	1|1	1|1	1|1	1|1	1|1	1|1	1|1		1|1	1|1
KM19 Pst													
H2.8 Nco	1|2	2|1	2|1	2|2	2|1	1|1	1|1	1|1		2|1		2|1	2|1
E5.1 Msp	2|1	1|2	1|2	1|1	1|2	2|2	2|2	2|2		1|2		1|2	1|2
CF	|+	+|-	+|-	+|+	+|-	-|-	-|-	-|-	+|-	+|-		+|	+|-
W30 MS	B|B	(B|A)	A|A	B|C	B|A	A|C	A|A	A|A		B|A		C|B	C|A
Acc	1|1	1|1	1|1	1|2	1|1	1|2	1|1		1|1	1|1		2|1	2|1
W46 Taq	2|2	2|1	2|1	2|2	2|1	1|2	1|1	1|1	2|1	2|1		2|1	2|1
Ins	1|1	1|2	2|2	1|2	1|2	2|2	2|2	2|2	1|2	1|2		2|2	2|2
W35 Eco	1|1	1|2	2|2	1|2	1|2	2|2	2|2	2|2		1|2		2|2	2|2
W32 Sac	2|2	2|2	1|2	2|1	2|2	2|1	2|2	2|2		2|2		1|2	1|2
311 Msp	1|1	1|2	1|2	1|1	1|2	1|2	2|2	2|2	1|2	1|2		1|1	1|2
Taq	2|2	2|1	1|1	2|1	2|1	1|1	1|1	1|1	2|1	2|1		1|2	1|1
W29 Pvu	2|2	2|2	1|2	2|1	2|2	2|1	2|2	2|2	2|2	1|2		1|2	
RV	1|1	1|2	1|2	1|1	1|2	2|1	2|2	2|2	1|2	1|2		1|2	1|2

Figure 5. Segregation of Markers in the Pedigree 1380 Which Contains a Recombination on the Telomeric Side of CF. The genotypes for the RFLPs described in this report as well as several recently described by Tsui et al. (17) are displayed. The order of the markers read from centromere to telomere, top to bottom. The left-handed haplotypes are maternal, right-handed paternal. MS corresponds to the CA repeat microsatellite.

possible haplotypes) three haplotypes predominate. This phenomenon has been observed on both normal and CF chromosomes, suggesting that the Caucasian population has a significant restriction in the number of haplotypes present in this region. Such large expanses of DNA in linkage disequilibrium pose problems for using locus expansion to increase the heterozygosity of specific regions. However, it does suggest that linkage disequilibrium may be useful in locating genetic components for diseases in which pedigrees are unavailable. Linkage disequilibrium also provides independent evidence on the location of mutations, although the use of this data is not widely agreed on. Because the markers on the other side of the gene show much greater association with the CF mutation (14), our data suggests that the gene will lie closer to MP6d than to J91.

The techniques and strategies developed for cloning the CF gene will greatly aid the progress in cloning other disease genes. We have shown that chromosome jumping and walking can be used to clone large stretches of DNA. By using the polymerase chain reaction to identify RFLPs in regions not found to be polymorphic by Southern blotting, virtually any region of the genome should be amenable to genetic analysis.

While this manuscript was being prepared, the cystic fibrosis gene was cloned and identified (17,29,30). The location of the CF gene is entirely consistent with both the recombination data and linkage disequilibrium data presented in this manuscript. Both parents of pedigree 1380 carry the common CF mutation, and all three of their affected children are homozygous for this mutation (data not shown).

ACKNOWLEDGEMENTS

We thank Mike Litt for helpful advice on resolving the alleles of CA repeats. M.C.I. and F.S.C. acknowledge support from the Cystic Fibrosis Foundation and NIH grant DK39690-01. This project has been funded at least in part with federal funds from the Department of Health and Human Services under contract N01-CO-74102 with Program Resources, Inc. F.S.C. is an Associate Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. White, R., Woodward, S., Leppert, M., O'Connell, P., Hoff, M., Herbst, J., Lalouel, J-M., Dean, M., and Vande Woude, G. *(1985) Nature (London)* **318**, 382–384 .
2. Dean, M., Park, M., Le Beau, M.M., Robins, T.S., Diaz, M. O., Rowley, J.D., Blair, D.G., and Vande Woude, G.F. (1985) *Nature (London)* **318**, 385–388.
3. Tsui, L-C., Buchwald, M., Barker, D., Braman, J.C., Knowlton, R., Schumm, J.W., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Brown, V., Helms, C., Gravius, T., Parker, C., Rediker, K., and Donis-Keller, H. (1985) *Science* **230**, 1054–1057.
4. Wainwright, B.J., Scrambler, P.J., Schmidtke, J., Watson, E. A., Law, H-Y., Farrall, M., Cook, H.J., Eiberg, H., and Williamson, R. (1985) *Nature (London)* **318**, 384–385.
5. Lathrop, G.M., Farrall, M., O'Connell, P., Wainwright, B., Leppert, M., Nakamura, Y., Lench, N., Kruyer, H., Dean, M., Park, M., Vande Woude, G., Lalouel, J.-M., Williamson, R., and White, R. (1988) *Amer. J. Hum. Genet.* **42**, 38–44.
6. Dean, M. (1988) *Genomics*, **3**, 93–99.
7. Scambler, P. J., Law, H-Y., Williamson, R., and Cooper, C. S. (1986) *Nucleic Acids Res.*, **14**, 7159–7174.
8. Zengerling, S., Olek, K., Tsui, L.-C., Grzeschik, K.-H., Riordan, J.R., and Buchwald, M. (1987) *Amer. J. Hum. Genet.* **40**, 228–236.
9. Collins, F.S., Drumm, M.L., Cole, J.L., Lockwood, W.K., Vande Woude, G.F., and Iannuzzi, M.C. (1987) *Science* **235**, 1046–1049.
10. Poutska, A. and Lehrach, H. (1986) *Trends Genet.* **2**, 174–179.
11. Collins. F.S. and Weissman, S.M. *(1984) Proc. Natl. Acad. Aci. USA* **81**, 6812–6816.
12. Iannuzzi, M.C., Dean, M., Drumm, M.L., Hidaka, N., Cole, J. L., Perry, A., Stewart, C., Gerrard, B., and Collins, F.S. *(1989) Am. J. Hum. Genet.* **44**, 695–703.
13. 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., and Williamson, R. (1986) *Amer. J. Hum. Genet.* **39**, 681–693.
14. Estivill, X., McLean, C., Nunes, V., Casals, T., Gallano, P., Scambler, P., Williamson, R. (1989) *Am. J. Hum. Genet.* **44**, 704–710.
15. Drumm, M.L., Smith, C.L., Dean, M., Cole, J.L., Iannuzzi, M. C., and Collins, F.S. (1988) *Genomics* **2**, 346–354.
16. Poustka, A.-M., Lehrach, H., Williamson, R., and Bates, G. (1988) *Genomics* **2**, 337–345.
17. Kerem et.al., (1989) *Science* **245**, 1073–1079.
18. Weber, J.L. and May, P.E. (1989) *Am. J. Hum. Genet.*, 44, 388–396.
19. Litt, M. and Luty, J.A. (1989) *Am. J. Hum. Genet.*, 44, 397–401.
20. White, R., Leppert, M., O'Connell, P., Nakamura, Y., Woodward, S., Hoff, M., Herbst, J., Dean, M., Vande Woude, G., Lathrop, G.M., and Lalouel, J-M. (1986) *Amer. J. Hum. Genet.* **39**, 694–698.
21. Myers, R.M., Larin, Z., and Maniatis, T. (1985) *Science* **230**, 1242–1246.
22. Myers, R.M., Lumelsky, N., Lerman, L.S., Maniatis, T. (1985) *Nature (London)* **313**, 495.

23. Martell, M., Le Gall, I., Millasseau, P., Dausset, J., and Cohen, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2682–2865.
24. Kreitman, M., and Aguade, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3562–3566.
25. Dean, M., et. al., *Hum. Genetics* (1990) (in press).
26. Northrup, H., Rosenbloom, C., O'Brien, W. E., Beaudet, A. L. (1989) *Nucleic Acids Res.* **17**, 1784.
27. Farrall, M., Wainwright, B.J., Feldman, G.L., Beauset, A., Sretenovic, Z., Halley, D., Simon, M., Dickerman, L., Devoto, M., Romeo, G., Kaplan, J.-C., Kitzis, A., and Williamson, R. (1988) *Am. J. Hum. Genet.* **43**, 471–475.
28. Rommens, J.M., Zengerling, S., Burns, J., Melmer, G., Kerem, B-S., Plavsic, N., Zsiga, M., Kennedy, D., Markiewicz, D., Roamahel, R., Riordan, J.R., Buchwald, M., and Tsui, L.-C. *(1989) Am. J. Hum. Genet.* **43**, 1–13.
29. Rommens, J.M., Iannuzzi, M.C., Bat-sheva, K., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.-C., Collins, F.S. (1989) *Science* **245**, 1059–1065.
30. Riordan et. al. (1989) *Science* **245**, 1066–1072.