

Prenatal diagnosis and linkage disequilibrium with cystic fibrosis for markers surrounding D7S8

Michael Dean¹, Jean A. Amos², Jennifer Lynch³, Giovanni Romeo⁴, Marcella Devoto⁴, Ken Ward⁵, Dicky Halley⁶, Ben Oostra⁶, Maurizio Ferrari⁷, Silvia Russo⁷, Bruce S. Weir⁸, Paula B. Finn², Francis S. Collins⁹, and Michael C. Iannuzzi⁹

¹Program Resources Inc., Frederick Cancer Research and Development Center, Frederick, MD 21701, USA

²Boston University School of Medicine, Boston, MA 02148, USA

³John Radcliffe Hospital, Oxford OX3 9DU, UK

⁴Istituto Giannina Gaslini, I-16148 Genoa, Italy

⁵University of Utah, Salt Lake City, UT 84132, USA

⁶Erasmus University, NL-300 DR Rotterdam, The Netherlands

⁷Istituto Scientifico H. S. Raffaele, I-20132 Milan, Italy

⁸North Carolina State University, Raleigh, NC 27695, USA

⁹University of Michigan, Ann Arbor, MI 48109, USA

Received September 27, 1989 / Revised November 21, 1989

Summary. Three polymorphic DNA markers surrounding the D7S8 locus were tested for their usefulness in the diagnosis of cystic fibrosis (CF) by linkage analysis. The markers correspond to the loci D7S424 and D7S426. These polymorphisms were studied by centers in the U.S., the United Kingdom, the Netherlands, and Italy, using samples from populations throughout Europe and North America. The additional information provided by these probes increased the heterogeneity of the region from 50% to 58% and was essential for a completely informative diagnosis in one family. A very high degree of linkage disequilibrium was found between these markers, which span a distance of approximately 250 kb. In addition, linkage disequilibrium with CF was noted. Significant heterogeneity of linkage disequilibrium was found among the populations, both for the marker-marker pairs and between the markers and CF.

Introduction

The discovery of linkage between markers on chromosome 7 and cystic fibrosis (CF) made possible carrier detection and prenatal diagnosis for families with a history of the disease (Dean et al. 1985, 1987; Farrall et al. 1986; Knowlton et al. 1985; Tsui et al. 1985; Wainwright et al. 1985; White et al. 1985, 1986; Dean 1988). The CF gene is flanked on one side by the MET and INT1L1 genes, and on the other by the D7S8 locus (Lathrop et al. 1988; Beaudet et al. 1986; Estivill et al. 1989). Restriction fragment length polymorphisms (RFLPs) at these loci

have been used to analyze CF families from several countries and ethnic backgrounds. Markers surrounding the INT1L1 gene show strong associations, linkage disequilibrium, with the CF mutation (Beaudet et al. 1986; Estivill et al. 1989; Schmidtke et al. 1987). Because the CF mutation is often on a haplotype that is infrequent in the general population these RFLPs are almost always informative in CF families. Relatively few RFLPs have been reported near D7S8, and this region shows only weak associations with CF (Estivill et al. 1989).

Using chromosome jumping to isolate new sequences a large distance from D7S8, we recently reported the identification of several new RFLPs (Iannuzzi et al. 1989). Our initial data suggested that these new polymorphisms were in strong disequilibrium with D7S8. We now report data from several clinical centers on the usefulness of these probes to CF families.

Materials and methods

DNA probes were labeled and used for Southern hybridization using standard techniques (Southern 1975; Feinberg and Vogelstein 1984). Standardized linkage disequilibrium was calculated as described (Iannuzzi et al. 1989; Hill and Robertson 1968). Tests for homogeneity of linkage disequilibrium were performed as described (Laurie-Ahlberg and Weir 1979; Weir 1990).

Results

Three previously reported RFLPs at the W32 (D7S424), and J29 (D7S426) locus were used by seven centers to type families requesting genetic diagnosis. W32 detects a

Table 1. Frequency of D7S8 region restriction fragment length polymorphisms (RFLPs)

Center ^a	W32 <i>SacI</i> , allele			D7S8 <i>MspI</i> , allele			J29 <i>PvuII</i> , allele			J29 <i>EcoRV</i> , allele		
	1	2	n	1	2	n	1	2	n	1	2	n
ROT	0.39	0.61	113	0.46	0.54	142	0.48	0.52	109	0.48	0.52	105
OXF	0.34	0.66	99	0.41	0.59	99	0.37	0.63	98	0.43	0.57	98
BOS	0.39	0.61	142	0.40	0.60	144	0.39	0.61	142	0.42	0.58	145
SLR	0.40	0.60	103	0.53	0.47	118	0.45	0.55	95	0.46	0.54	50
SLC	0.41	0.59	98	0.45	0.55	104	0.42	0.58	102	-	-	-
MI	0.40	0.60	297	0.51	0.49	181	0.19	0.71	16	0.31	0.69	16
GE	0.53	0.47	36	0.38	0.62	239	0.44	0.56	100	-	-	-
Total	0.40	0.60	888	0.45	0.55	1027	0.41	0.59	662	0.44	0.56	414

^a Centers are: ROT, Rotterdam; OXF, Oxford; BOS, Boston; SLR and SLC, Salt Lake City reference and clinical families; MI, Milan; GE, Genoa

Table 2. Linkage disequilibrium values for marker-marker pairs^a

	<i>SacI-MspI</i> ^b			<i>SacI-PvuII</i> ^b			<i>SacI-EcoRV</i>			<i>MspI-PvuII</i> ^b			<i>MspI-EcoRV</i> ^b			<i>PvuII-EcoRV</i>		
	Δ	χ^2	n	Δ	χ^2	n	Δ	χ^2	n	Δ	χ^2	n	Δ	χ^2	n	Δ	χ^2	n
MI	0.67	64	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GE	1.0	28	28	0.62	12	32	-	-	-	0.76	55	94	-	-	-	-	-	-
ROT	0.81	80	105	0.92	69	81	0.85	64	89	0.92	86	101	0.96	93	101	0.90	62	77
OXF	0.83	57	87	0.89	69	86	0.79	54	87	0.70	42	86	0.78	53	87	0.88	67	86
UTR	0.72	44	84	0.90	62	77	-	-	-	0.69	39	81	-	-	-	-	-	-
UTC	0.88	76	98	0.94	86	98	-	-	-	0.90	79	98	-	-	-	-	-	-
BOS	0.73	76	141	0.83	96	139	0.80	89	141	0.65	65	142	0.64	59	144	0.87	108	142
Total			688			507	0.81	207	317			602			341	0.88	235	305

^a Probes used were as follows: *SacI*, W32 (D7S424); *MspI*, pJ3.11 (D7S8); *PvuII* and *EcoRV*, J29 (D7S426)

^b These populations showed significant heterogeneity for linkage disequilibrium and were not totaled

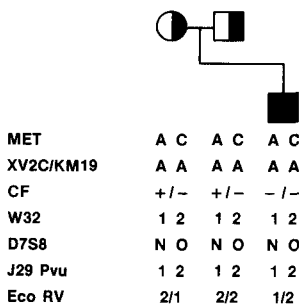


Fig. 1. Cystic fibrosis (CF) family informative for the J29 *EcoRV* polymorphism. Shown are the parents (*half-filled*) and CF child (*filled*) and the genotypes for several RFLPs. The mother is heterozygous for J29 *EcoRV*, whereas the father is homozygous. This allows the mother's CF and normal chromosomes to be followed

SacI RFLP and is 60 kb centromeric to D7S8; J29 detects *EcoRV* and *PvuII* polymorphisms and is 150 kb telomeric to D7S8 (Iannuzzi et al. 1989). The majority of families were nuclear, comprising both parents and one affected child. In total 306 families were studied, although each marker was not typed in every individual.

The frequency of the RFLPs were not significantly different (at the 5% level) between the study groups (Table 1). These new markers increased the heterozygosity of the D7S8 side of the gene from 50% to 58%, allowing a greater percentage of families to be informative. To examine the extent of linkage disequilibrium (Δ) between these markers and D7S8, haplotypes were generated. All marker-marker pairs showed strong association in all the populations studied, with standardized linkage disequilibrium values ranging from 0.62 to 1.0 (Table 2). However, a statistically significant degree of heterogeneity over populations was found for marker pairs (data not shown). No correlation was found between the extent of disequilibrium and the distance between the probes. By analyzing haplotypes we found that only 11 of 16 possible haplotypes were present in 149 CF chromosomes, with two haplotypes (1-1-1-1, 2-2-2-2) comprising 83% of the sample. On normal chromosomes, 12 haplotypes were found out of 145 chromosomes, with the same two haplotypes accounting for 85% of the total. Thus there is a significant distortion in the distribution of haplotypes in this region.

Despite the modest increase in heterozygosity observed with these new probes, use has been made of the

Table 3. Linkage disequilibrium with the cystic fibrosis locus^a

	<i>SacI</i> ^b			<i>MspI</i> ^b			<i>PvuII</i>			<i>EcoRV</i> ^b		
	Δ	χ^2	<i>n</i>	Δ	χ^2	<i>n</i>	Δ	χ^2	<i>n</i>	Δ	χ^2	<i>n</i>
MI	0.07	1.2	251	-0.09	1.2	145	-	-	-	-	-	-
GE	-0.17	1.0	36	0.15	2.1	94	0	0	100	-	-	-
ROT	-0.23	5.5	105	-0.31	5.8	129	-0.34	11	101	-0.27	7.3	101
UTR	-0.10	0.9	95	-0.01	0	105	-0.05	0.2	92	-	-	-
OXF	-0.12	1.2	87	-0.11	1.0	87	-0.13	1.5	86	-0.13	1.5	87
BOS	-0.29	12	141	-0.30	13	144	-0.23	7.6	142	0.20	6	144
UTC	-0.18	3.2	98	-0.13	1.7	98	-0.15	2.5	98	-	-	-
Total			811			802	-0.16	16	637			334

^a Probes used: *SacI*, W32 (D7S424); *MspI*, pJ3.11 (D7S8); *PvuII* and *EcoRV*, J29 (D7S426)

^b Populations heterogeneous for linkage disequilibrium

J29 *EcoRV* RFLP to achieve a fully informative diagnosis in one case. Here both parents and the affected child were heterozygous for several markers commonly used for prenatal diagnosis (Fig. 1), thus marking it impossible to distinguish an affected from an unaffected fetus. The *EcoRV* RFLP was informative only in the mother, which allowed the maternal CF chromosome to be traced in the fetal DNA, and by combining this marker with the others, a prediction can be made from all possible genotypes.

Low, but significant, values of linkage disequilibrium with CF have been reported for D7S8 (Estivill et al. 1989; Schmidtke et al. 1987), and most of the centers in this study showed low Δ values for markers in this region and CF (Table 3). Significant heterogeneity over populations was found for associations of CF with all markers except for the J29 *PvuII* RFLP. Two of the centers (BOS, ROT) show strikingly higher values than all the others. Both of these centers received patients from a dispersed area with no obvious concentration of any one geographical region. The BOS families were taken from all over North America, and the ROT sample includes families from several European countries (J.A. Amos, D. Halley, and B. Oostra, data not shown). Because all of the patients are Caucasians of European origin this discrepancy is hard to explain.

Due to the strong linkage disequilibrium exhibited between these new markers and D7S8, the RFLPs studied here will not be useful as primary markers for the analysis of CF families. These markers could certainly be used in the place of D7S8, and in the hands of some groups gave clearer results. However, as shown in Fig. 1, these probes do provide additional information in certain cases. We estimate that 16% of individuals uninformative for D7S8 will be informative with one of these polymorphisms.

Discussion

Following the discovery of linkage between an RFLP marker and a disease gene, the next step is to identify markers flanking the gene, and further closely linked,

highly polymorphic RFLPs. The rationale behind this strategy is to increase the applicability of genetic testing and pedigree analysis. With flanking markers, each 10 cM from the disease gene, a fully informative family can be diagnosed with an accuracy approaching 96%. Locus expansion, (isolating additional sequences flanking a marker using chromosome walking) usually results in the identification of additional polymorphisms, which increase the heterozygosity of the locus. For example, this strategy was used to increase the informativeness of markers linked to Huntington's chorea (Gilliam et al. 1987). However, as reported here, large regions of DNA in linkage disequilibrium may frustrate such efforts. While the sequences surrounding D7S8 may not be typical of most loci, only a small portion of the human genome has been examined in this fashion, making it hard to predict the extent of disequilibrium in other regions. Rommens et al. (1989) have found a similar large stretch of DNA in disequilibrium between the INT1L1 gene and CF. Large regions of DNA in disequilibrium are more likely to occur for diseases like CF, which are restricted to specific racial groups, and have a low mutation rate. On the other hand, the presence of large regions in disequilibrium would aid the analysis of complex diseases that involve both genetic and environmental factors, and may require population genetics approaches.

The range of associations between D7S8 and CF in the different centers was surprising, given that almost all the families are Caucasian. This difference can be most likely explained by either the presence of actual differences in the populations or an artefact of the sample size. Since both the BOS and ROT populations are quite heterogeneous, the difference is most likely due to sample size. These two centers studied 42 and 35 families, respectively.

While this manuscript was in preparation the cloning of the CF gene, and the identification of a mutation present on 70% of CF chromosomes was described (Rommens et al. 1989; Riordan et al. 1989). Presumably the remaining mutations will be identified, and the diagnosis of CF will involve the direct identification of mutations. In this case the markers described here will no longer be of use in prenatal diagnosis. However, the approaches

and problems described will apply to the diagnosis of other genetic diseases as they progress from linkage to the identification of mutations.

Acknowledgements. We thank Maria Whittaker for assistance in compiling the data and Julie Cashour for typing the manuscript. M. C. I. and F. S. C. acknowledge support from the Cystic Fibrosis Foundation and NIH grant DK39690-01. B. S. W. acknowledges support from NIH grant GM32518. 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

- Beaudet A, Bowcock A, Buchwald M, Cavalli-Sforza L, Farrall M, King MC, Klinger K, Lalouel JM, Lathrop G, Naylor S, Ott J, Tsui LC, 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
- Dean M (1988) Molecular and genetic analysis of cystic fibrosis. *Genomics* 3:93-99
- Dean M, Park M, Le Beau MM, Robins TS, Diaz MO, Rowley JD, Blair DG, Vande Woude GF (1985) The human *met* oncogene is related to the tyrosine kinase oncogenes. *Nature* 318:385-388
- Dean M, O'Connell P, Leppert M, Park M, Amos JA, Phillips DG, White R, Vande Woude GF (1987) Three additional DNA polymorphisms in the *met* gene and D7S8 locus: use in prenatal diagnosis of cystic fibrosis. *J Pediatr* 111:490-495
- Estivill X, McClean C, Nunes V, Casais T, Gallano P, Scambler P, Williamson R (1989) Isolation of a new DNA marker in linkage disequilibrium with cystic fibrosis, situated between J3.11 (D7S8) and IRP. *Am J Hum Genet* 44:701-710
- Farrall M, Law HY, Rodeck CH, Warren R, Stanier P, Super M, Lissens W, Scambler P, Watson E, Wainwright B, Williamson R (1986) First-trimester prenatal diagnosis of cystic fibrosis with linked DNA probes. *Lancet* I:1402-1405
- Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction fragments to high specific activity. *Addendum. Anal Biochem* 137:266-267
- Gilliam TC, Bucan M, MacDonald ME, Zimmer M, Haines JL, Cheng SV, Pohl TM, Meyers RH, Whaley WL, Allitto BA, Faryniarz A, Wasmuth JJ, Frischaur AM, Conneally PM, Lehrach H, Gusella JF (1987) A DNA segment encoding two genes very tightly linked to Huntington's disease. *Science* 238:950-952
- Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. *Theor Appl Genet* 38:226-231
- Iannuzzi MC, Dean M, Drumm ML, Hidaka N, Cole JL, Perry A, Stewart C, Gerrard B, Collins FS (1989) Isolation of additional polymorphic clones from the cystic fibrosis region, using chromosome jumping from D7S8. *Am J Hum Genet* 44:695-703
- Knowlton RG, Cohen-Haguenauer O, Nguyen VC, Frézal J, Brown VA, Barker D, Braman JC, Schumm JW, Tsui LC, Buchwald M, Donis-Keller H (1985) A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature* 318:380-382
- Lathrop GM, Farrall M, O'Connell P, Wainwright B, Leppert M, Nakamura Y, Lench N, Kruyer H, Dean M, Park M, Vande Woude G, Lalouel JM, Williamson R, White R (1988) Refined linkage map of chromosome 7 in the region of the cystic fibrosis gene. *Am J Hum Genet* 42:38-44
- Laurie-Ahlberg CC, Weir BS (1979) Allozymic variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster*. *Genetics* 92:1295-1314
- Riordan JR, Rommens JM, Kerem BS, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, Drumm ML, Iannuzzi C, Collins FS, Tsui LC (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245:1066-1073
- Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989) Identification of the cystic fibrosis locus: chromosome walking and jumping. *Science* 245:1059-1065
- Schmidtko J, Krawczak M, Schwartz M, Alkan M, Bonduelle M, Bühler E, Chemke M, Darnedde T, Domagk J, Engel W, Frey D, Fryberg K, Halley D, Hundrieser J, Ladanyi L, Libaers I, Lissens W, Mächler M, Malik NJ, Morreau J, Neubauer V, Oostra B, Pape B, Poncin JE, Schinzel A, Simon P, Trefz FK, Tümmler B, Vassart G, Voss R (1987) Linkage relationships and allelic associations of the cystic fibrosis locus and four marker loci. *Hum Genet* 76:337-343
- Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503-517
- Tsui LC, Buchwald M, Barker D, Braman JC, Knowlton R, Schumm JW, 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) Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science* 230:1054-1057
- Wainwright BJ, Scambler PJ, Schmidtko J, Watson EA, Law HY, Farrall M, Cooke HJ, Eiberg H, Williamson R (1985) Localization of cystic fibrosis locus to human chromosome 7cen-q22. *Nature* 318:384-385
- Weir BS (1990) Genetic data analysis. Sinauer, Sunderland, Mass
- White R, Woodward S, Leppert M, O'Connell P, Hoff M, Herbst J, Lalouel JM, Dean M, Vande Woude G (1985) A closely linked genetic marker for cystic fibrosis. *Nature* 318:382-384
- White R, Leppert M, O'Connell P, Nakamura Y, Woodward S, Hoff M, Herbst J, Dean M, Vande Woude G, Lathrop GM, Lalouel JM (1986) Further linkage data on cystic fibrosis: the Utah study. *Am J Hum Genet* 39:694-698