Am. J. Hum. Genet. 44:679–685, 1989

An Intronic Region within the Human Factor VIII Gene Is Duplicated within Xq28 and Is Homologous to the Polymorphic Locus DXS115 (767)

M. Patterson,*[,1] J. Gitschier,† J. Bloomfield,* M. Bell,* H. Dorkins,* U. Froster-Iskenius,‡ S. Sommer,† J. Sobell,† D. Schaid,† S. Thibodeau,§ and K. E. Davies*

*Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford: †Howard Hughes Medical Institute, University of California, San Francisco; ‡Institut für Humangenetik, Medizinische Hochschule Lübeck, Federal Republic of Germany; and §Mayo Clinic, Rochester, MN

Summary

The genomic sequences recognized by the anonymous probe 767 (DXS115) are localized to two sites within Xq28. One site lies within intron 22 of the factor VIII gene (FBC). Physical mapping suggests that the second site lies within 1.2 megabases of the F8C gene. The RFLPs detected by 767 are located within the second site. Genetic data suggest that F8C and DXS115 are tightly linked (θ_{max} = .04; Z_{max} = 8.30). Recombination events in meioses informative for DXS52 (St14), DXS115, and F8C suggest that DXS115 and F8C lie distal to DXS52.

Introduction

In the course of our analysis of the Xq27/28 region of the human X chromosome, we have been able to determine the order and physical separation of cloned genomic segments and genes by using pulsed-field gel electrophoresis (PFGE). We have shown, for example, that the probes St14 (DXS52), DX13 (DXS15), and MN12 (DXS33) lie within a 470-kb region (Patterson et al. 1987*a*). In addition, we have presented data suggesting that the gene encoding glucose-6-phosphate dehydrogenase (G6PD) lies within 500 kb of the factor VIII (F8C) gene (Patterson et al. 1987*b*). Thus it is becoming possible both to relate physical with genetic distance and to unravel the uncertainties, which have arisen from linkage analyses, regarding locus order in Xq28 (reviewed in Davies et al. 1988).

Received October 4, 1988; revision received December 14, 1988.

Address for correspondence and reprints: Dr. Kay E. Davies, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, England.

1. Present address: Division of Oncology, Stanford University Medical Center, Stanford, CA 94305–5306.

0002-9297/89/4405-0008$02.00

The anonymous probe, 767, was originally isolated from an X chromosome library and was localized by using somatic cell hybrids to Xq27-qter (Hofker et al. 1985). Recently, RFLPs have been reported for 767 by using the enzymes *Bst*XI and *Pst*I (Arveiler et al. 1988; Patterson et al. 1988*a*). To assess the utility of these RFLPs, it is important to establish the location of these polymorphisms relative to other markers in Xq28. Elsewhere we have presented data that suggest that 767 lies within 250 kb of F8C (Patterson et al. 1987*b*). However, in the present paper we present the results of a range of physical and genetic mapping experiments, which provide a more complete picture of the relationship between 767 and F8C. First, we show that 767 is homologous to a region within intron 22 of F8C. This region is duplicated in the human genome (Wion et al. 1986), and we show by PFGE and deletion-mapping experiments that both RFLPs associated with 767 derive from the duplicated site within 1.2 megabases (Mb) of F8C. Finally, we have used linkage analysis in normal pedigrees and in pedigrees affected by the fragile X syndrome to examine the location of the 767 RFLPs and F8C relative to the cluster of polymorphic loci around DXS52 (St14). The data suggest that DXS115 and F8C are distal to DXS52.

Material and Methods

DNA Probes

767 is a 1.3-kb *Hin*dIII fragment cloned in pAT153 and localized to Xq27-qter (Hofker et al. 1985).

Three probes from the factor VIII gene were used: F8.19 (450-bp *Eco*RI/*Sst*I fragment in pGemini 3), from the promoter (J. Gitschier, unpublished observations); 114.12 (647-bp *Stu*I/*Sca*I fragment in pUC12), containing part of exon 17 and all of exon 18 (Gitschier et al. 1985*b*); 625.8 (3.1-kb *Eco*RI fragment in pUC18) from the 3′ end of the gene (J. Gitschier, unpublished observations).

Probes were radioactively labeled, after purification of the insert DNA, by the random priming method (Feinberg and Vogelstein 1983).

Southern Analysis

Standard methods were employed for conventional Southern analysis (Southern 1975; Davies et al. 1983).

PFGE

Sample preparation, restriction digestion, and associated methods were performed as described elsewhere (Kenwrick et al. 1987; Patterson et al. 1987*b*). A commercial apparatus was used for PFGE (Pharmacia-LKB, Ltd.). An orthogonal array of electrodes was used to produce the results shown in figure 1, and a hexagonal array was used to produce those shown in figure 4. The running conditions used for these gels are given in the figure legends.

Genetic Analysis

Fragile X pedigrees reported by Thibodeau et al. (1988) were analyzed with F8C and 767 (DXS115), by using the *Bst*XI RFLP (Arveiler et al. 1988). The informative pedigrees used in the linkage analysis are Fx7, Fx20, Fx32, Fx33, Fx62, Fx64, and Fx73. In addition, the pedigree presented by Patterson et al. (1988*a*) was analyzed with F8C. The data include one additional nuclear pedigree with three offspring, all of whom are affected. This pedigree was informative only at the F8C locus. The parameters used in the linkage calculations were as described elsewhere (Thibodeau et al. 1988).

Normal families from the Centre d'Ètude du Polymorphisme Humain collection were analyzed with 55E (DXS105; Heilig et al. 1988), St14 (DXS52; Oberle et al. 1985), F8C, and 767. Informative pedigrees are 2, 17, 28, 35, 37, and 45. The LINKAGE program was used for two-point and multipoint calculations (Lathrop and Lalouel 1984; Lathrop et al. 1985).

Results and Discussion

PFGE Mapping Using 767

PFGE has already provided evidence for the physical linkage of 767 and F8C. In particular, 767 identifies the same size *Nru*I fragments in partial digests as F8C, at 250 kb and 500 kb (Patterson et al. 1987*b*). However, with other enzymes, such as *Sfi*I and *Sst*II, 767 was found to hybridize to bands in addition to those recognized by F8C. An example of this is shown in figure 1, in which partial *Sst*II digests are used to construct a long-range map of the F8C region. Partial digestion was encouraged by using salt concentrations greater than those recommended by the manufacturers. The variability of the band intensities reflects the differing degrees of partial digestion under the different digestion conditions. While it is possible that altering the conditions of the restriction digest in this way may change the specificity of the restriction enzyme, this does not seem to be the case for *Sst*II or *Nru*I (Patterson et al. 1987*b*). The higher-molecular-weight bands observed in the partial digests at elevated salt concentration are also present, albeit at reduced intensity, in the digestion under the recommended conditions.

The F8C probes, 114.12 and F8-19, are known to lie approximately 140 kb apart, F8-19 being derived from the 5′ end of the F8C gene (J. Gitschier, unpublished observations) and 114.12 being derived from the exon 17/18 region (Gitschier et al. 1985*b*). Both probes produce identical patterns with *Sst*II digests, identifying the smallest band at 150 kb. Thus, the probes must lie at opposite ends of this fragment, as shown in figure 1*d*. The third F8C probe, 625.8, is a genomic probe which is 90 kb distal to 114.12 and hybridizes to smaller fragments in addition to fragments common to 114.12 and F8-19 (fig. 1*b*). From the sizes of the fragments a unique map of the *Sst*II sites can be constructed (fig. 1*d*), with 625.8 straddling a *Sst*II site. This prediction was upheld, since the 625.8 probe was found to be susceptible to cleavage by *Sst*II.

When 767 was hybridized to the *Sst*II filter, a complex pattern of bands was observed (fig. 1*c*). Some of the bands detected by 767 are shared with only 625.8 (blocked arrows), while others are common to all the factor VIII probes (clear arrows). This suggests very strongly, therefore, that 767 lies within the factor VIII gene at the location indicated in figure 1*d*. However, in addition to the F8C bands, 767 appears to hybridize to at least four other *Sst*II bands which remain after washing at high stringency. A similar observation is made when using the enzyme *Sfi*I, where 767 detects two bands; the first, at 280 kb, is coincident with the

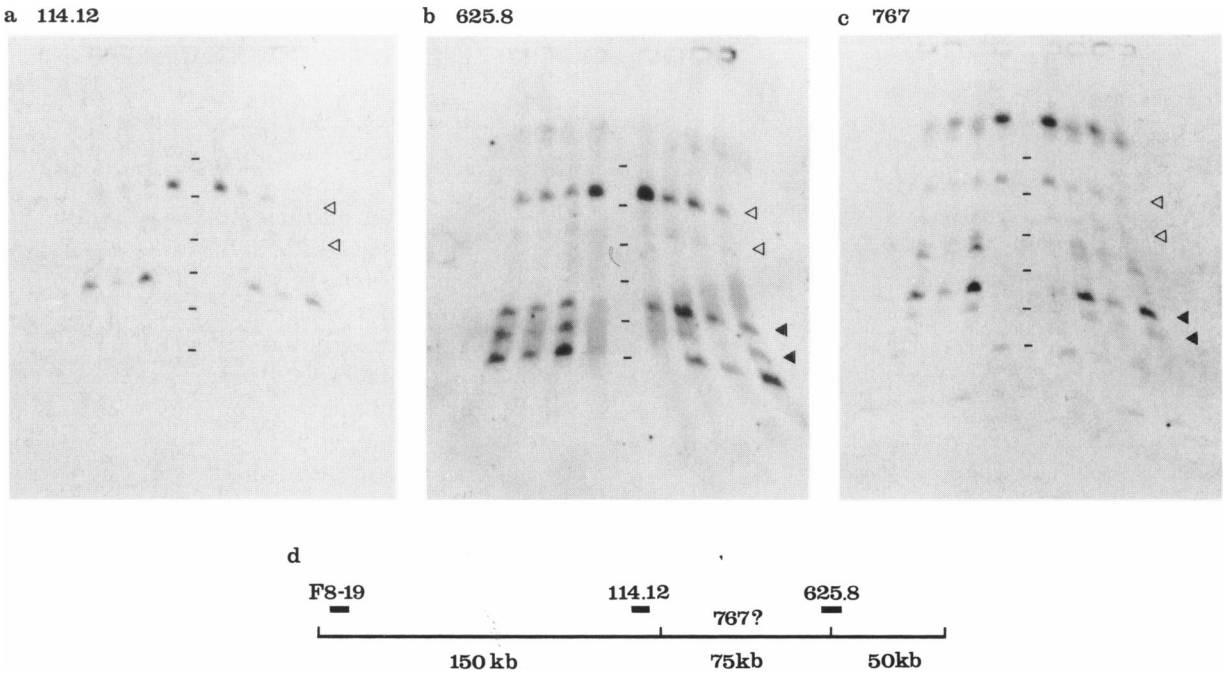

Figure 1 *Sst*II restriction-site map across the F8C gene. The DNA samples on the left half of the gel are from a normal male, and those on the right from a fragile X male. Samples were digested with *Sst*II, but with varying amounts of salt (NaCl) in the restriction buffer: the manufacturers' recommended buffer (containing 50 mM NaCl) was used for the samples run in the outermost two lanes on each side. The salt concentration was increased to 100 mM in the samples run second from the center and was increased to 150 mM for the most central lanes. The 1% gel was run at 200 V, with a pulse time of 35 s. The filter was hybridized successively to probes as indicated. Open and closed trinagles indicate bands common to probes. The lines in the center of the gel show the positions of the lambda concatamers (multiples of 50 kb) used as markers. A restriction-site map is shown in panel *d*, with the short vertical lines representing *Sst*II sites.

single band detected by all three F8C probes, and a second, at 310 kb, is shared with no other probes (results not shown). Since 767 does not contain sites for *Sst*II or *Sfi*I, these observations suggest that 767 hybridizes to two regions: one within F8C and the other elsewhere in the genome.

767 Hybridizes to a Region within F8C

To test directly whether 767 is located within F8C, we hybridized 767 to a series of genomic clones covering most of the factor VIII gene. 767 clearly identifies a 9.6-kb *Eco*RI fragment within cosmid 542, a result that localizes 767 within intron 22 of F8C (fig. 2). At this stage, it became possible to explain the additional bands detected by 767 in PFGE digests such as those in figure 1. This is because the intron 22 region to which 767 hybridized is known to be present at one other location in the human genome (Wion et al. 1986). 767 therefore identifies bands from both the F8C region and the second region. From the results presented previously, the location of the duplicated region, homologous to intron 22 could not, however, be determined.

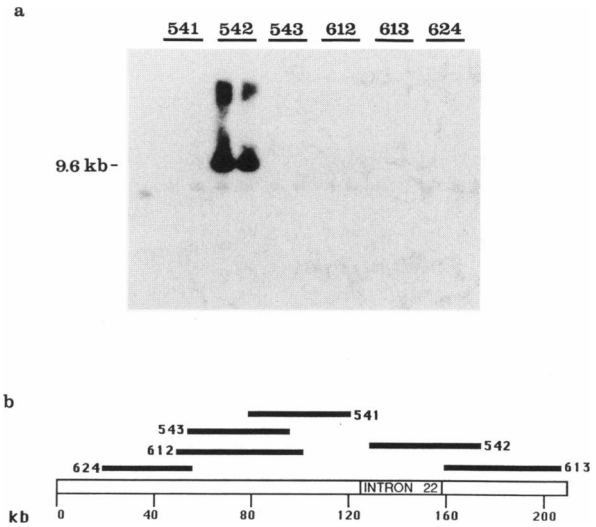

Figure 2 Hybridization of 767 to F8C genomic clones. *A*, *Eco*RI digests of cosmids covering most of the F8C gene were hybridized to 767. *B*, Schematic diagram showing the location of the genomic clones within F8C. The scale used is the same as that used by Gitschier et al. (1984), and the position of intron 22 which contain the *Eco*RI fragment to which 767 hybridizes is shown.

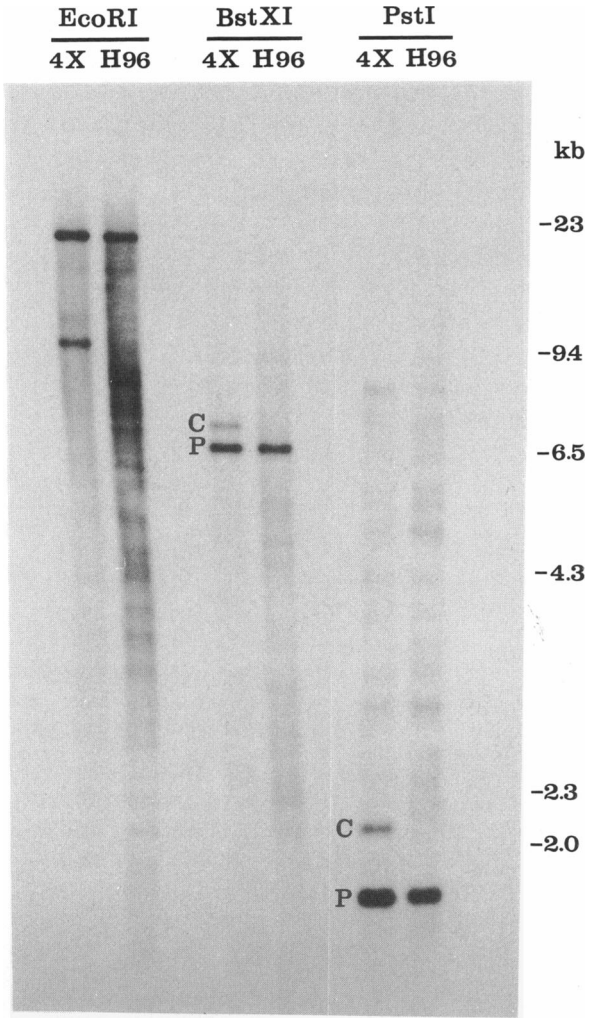

Figure 3 Mapping the RFLPs detected by 767. DNA samples from a control (4X) cell line and from a patient deleted from the intron 22 region of F8C were digested with enzymes as indicated. A 0.8% agarose gel was run, and the results of hybridization to 767 are shown. C and P indicate, respectively, the constant and polymorphic bands detected by 767.

Mapping of the 767 RFLPs outside F8C

Recently, 767 has been shown to detect two RFLPs with the enzymes *Bst*XI and *Pst*I (Arveiler et al. 1988; Patterson et al. 1988*a*). After it was established that 767 hybridizes to two separate regions of the human genome it became important to determine which of these regions contained the polymorphic sites. This question was addressed using a hemophilia A patient who carries a 39-kb deletion encompassing all of intron 22 of F8C (patient H96; Gitschier et al. 1985*a*). The results are shown in figure 3. First, in an *Eco*RI digest, 767 identifies 9.6-kb and 21-kb fragments in control DNA (4X), whereas the 9.6-kb fragment is absent in the patient DNA. In *Bst*XI and *Pst*I digests 767 detects an upper constant band and a lower polymorphic pair of fragments. For both enzymes the patient DNA reveals one of the polymorphic bands but is deleted for the constant bands. Thus, both the *Pst*I and the *Bst*XI polymorphisms which constitute the genetic locus, DXS115, derive from the non-F8C region to which 767 hybridizes. This is also consistent with the almost complete linkage disequilibrium that we have observed between the *Bst*XI and *Pst*I polymorphisms.

Both the constant and the polymorphic bands detected by 767 cosegregate in a number of somatic cell hybrids, including hybrids whose only human X component is that from Xq26-qter (results not shown). In addition, all the bands detected by 767 are present in a single dose in three unrelated females who carry a terminal deletion of the distal tip of the long arm of one of their X chromosomes (Patterson et al. 1987*b*). Thus, the polymorphisms detected by 767 are likely to be located distal to the fragile site at Xq27.3.

Physical mapping suggests that the duplicated regions may lie within a single large *Not*I fragment. F8C probes such as 114.12 detect a 1.2-Mb *Not*I fragment (fig. 4). As expected, 767 also detects this fragment, but it detects no other fragment below 7 Mb. By using gels that resolve in the kilobase range, we have excluded the possibility that 767 hybridizes to a small *Not*I fragment. It is possible that 767 lies within a very large *Not*I fragment of greater than 7 Mb, although we consider this unlikely given the relatively high density of sites for rare-cutter enzymes on the distal side of the fragile site (Patterson et al. 1988*b*). These results therefore suggest that the polymorphisms detected by 767 are within 1.2 Mb of F8C. This mapping also increases, to approximately 1.5 Mb, the region that has been mapped around G6PD and F8C (Patterson et al. 1987*b*). However, the physical distance between this region and the cluster of polymorphic loci around DXS52 (St14) has not been established.

Genetic Analysis

Linkage analysis both in normal pedigrees and in families affected by the fragile X syndrome provides evidence for tight linkage between the loci DXS115, DXS52, and F8C. The two-point lod scores are shown in table 1, and the results are in good agreement with previous data (Arveiler et al. 1988).

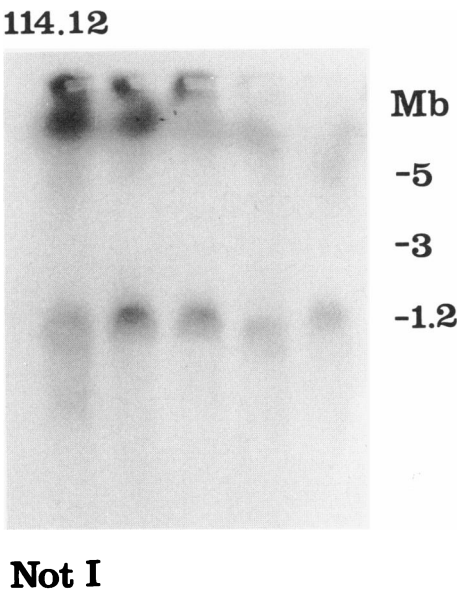

Figure 4 1.2-Mb *Not*I band detected by F8C probes. Male DNA samples from five independent EBV-transformed cell lines were digested from *Not*I and were subjected to PFGE through a 0.6% agarose gel at 40 V with a pulse time of 60 min. After Southern transfer the filter was hybridized to 114.12. The size markers used were *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* chromosomes with sizes as estimated by Vollrath and Davis (1987).

In order to construct a complete physical and genetic map of Xq27/28, we are interested in the relative positions of DXS52, DXS115, and F8C, and three multiply informative recombinants shed some light on this question. In the normal families a recombinant is observed between DXS115 and F8C (fig. 5*a*, II.4). Since this family is also informative for DXS105 (55E) (Heilig et al. 1988), the results suggest that DXS115 is distal to F8C. Unfortunately, this family is not informative at DXS52, and, overall, multipoint analysis of the normal data set does not distinguish significantly between different orders for the three distal loci. Two recombinants are also observed in the fragile X families. The recombination between DXS115 and DXS52 shown in figure 5*b* (III.2) has been reported elsewhere (Patterson et al. 1988*a*). The information at the F8C locus presented here suggests that both DXS115 and F8C are distal to DXS52. Finally, figure 5*c* shows a recombination event between DXS52 and F8C, providing further evidence for a distal localization of F8C with respect to DXS52, although this assumes that individual III.1 is not a nonpenetrant male carrier. Multipoint analysis of the fragile X family data suggests that the order FRAXA–DXS52–(F8C,DXS115) is approximately 12 times more likely than any other.

Conclusions

A variety of physical and genetic mapping strategies have been used to determine the relationship in Xq28 between DXS115 (detected by 767) and F8C. First, the locus DXS115 does not lie within 250 kb of the F8C locus as has been suggested elsewhere (Patterson et al. 1987*b*), since the probe 767 hybridizes to two regions within Xq28. One of these regions is within intron 22 of the factor VIII gene. The second region, which contains both RFLPs detected by 767 and can therefore be considered to correspond to DXS115, probably lies within 1.2 Mb of F8C. Linkage analysis reveals tight linkage between the loci DXS115, DXS52, and F8C, and multiply informative recombination events suggest that the most likely locus order is FRAXA–DXS52–F8C–DXS115–Xqter.

Table 1

Linkage Analysis between Loci in Xq27-qter

Locus 1	Locus 2	Z_{max}	θ_{max}	Recombination Fraction (θ)							
				.001	.01	.05	.1	.15	.2	.3	.4
FRAXA	DXS52 (St14)	1.91	.19	−13.12	−5.84	−.46	1.26	1.83	1.90	1.37	.53
	F8C (F8)	1.18	.13	−1.90	−.27	.86	1.15	1.17	1.06	.68	.24
	DXS115 (767)	.49	.20	−4.73	−2.17	−.36	.23	.44	.49	.36	.12
DXS52 (St14)	F8C	6.27	.07	−1.93	4.77	6.18	6.18	5.76	5.15	3.55	1.62
	DXS115 (767)	5.80	.06	3.11	4.97	5.78	5.56	5.14	4.54	3.02	1.28
F8C	DXS115 (767)	8.25	.02	7.28	8.12	8.09	7.42	6.61	5.70	3.65	1.49

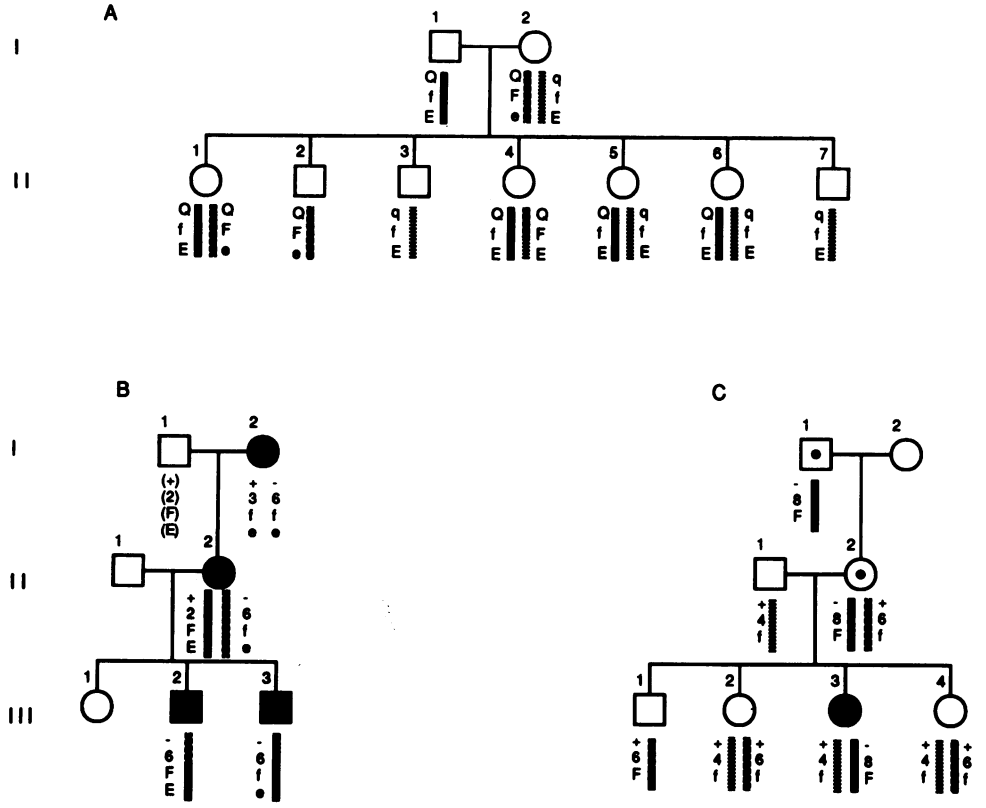

Figure 5 Recombination events that help to order loci within Xq28. *A*, Normal pedigree from the CEPH panel (CEPH 28). *B*, Part of the fragile X pedigree analyzed by Patterson et al. (1988*a*). *C*, Part of the fragile X pedigree (Fx7) analyzed by Thibodeau et al. (1988). Loci are abbreviated as follows: Q, q = DXS105 (55E); 2, 3, 4, 6, and 8 = DXS52 (St14); F, f = F8C; E, e = DXS115 (767). Shading is used to indicate the most likely haplotypes and the positions at which recombination is thought to have occurred. Affected individuals are indicated by a black symbol, and phenotypically normal carriers are indicated by a dot within the symbol.

Acknowledgements

We would like to thank Dr. P. Pearson for the contribution of the 767 probe, Dr. J.-L. Mandel for 55E, Tracy Flint for expert technical assistance, Helen Blaber and Janet Watt for help in the preparation of the manuscript, and Drs. W. Cookson and S. Kenwrick for many helpful discussions. This research was supported by The Medical Research Council.

References

Arveiler, B., I. Oberlé, A. Vincent, M. H. Hofker, P. L. Pearson, and J. L. Mandel. 1988. Genetic mapping of the Xq27-q28 region: new RFLP markers useful for diagnostic applications in fragile X and hemophilia-B families. Am. J. Hum. Genet. **42**:380–389.

Davies, K. C., J.-L. Mandel, J. Weissenbach, and M. Fellous. 1988. Cytogenet. Cell Genet. **46**:277–315.

Davies, K. E., P. L. Pearson, P. S. Harper, J. M. Murray, T. O'Brian, M. Sarfarazi, and R. Williamson. 1983. Linkage analysis of two cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. Nucleic Acids Res. **11**:2303–2312.

Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.

Gitschier, J., D. Drayna, E. G. D. Tuddenham, R. L. White, and R. M. Lawn. 1985*a*. Genetic mapping and diagnosis of haemophilia A achieved through a *BclI* polymorphism in the factor VIII gene. Nature **314**:738–740.

Gitschier, J., W. I. Wood, T. M. Goralka, K. L. Wion, E. Y. Chen, D. H. Eaton, G. A. Vehar, D. J. Capon, and R. M. Lawn. 1984. Characterisation of human factor VIII gene. Nature **312**:326–330.

Gitschier, J., W. I. Wood, E. G. D. Tuddenham, M. A. Shuman, T. M. Goralka, E. Y. Chen, and R. M. Lawn. 1985*b*. Detection and sequence of mutations in the factor VIII gene of haemophiliacs. Nature **315**:427–430.

Heilig, R., I. Oberle, B. Arveiler, A. Hanauer, M. Vivaud, and J.-L. Mandel. 1988. Improved DNA markers for efficient analysis of fragile X families. Am. J. Med. Genet. **30**:543–550.

Hofker, M. H., M. C. Wapenaar, N. Goor, E. Bakker, G.-J. B. van Ommen, and P. L. Pearson. 1985. Isolation of probes detecting restriction fragment length polymorphisms from X chromosome–specific libraries: potential use for diagnosis of Duchenne muscular dystrophy. Hum. Genet. **70:** 148–156.

Kenwrick, S., M. N. Patterson, A. Speer, K. Fischbeck, and K. Davies. 1987. Molecular analysis of the Duchenne muscular dystrophy region using pulsed-field gel electrophoresis. Cell **48:**351–357.

Lathrop, G. M., and J. M. Lalouel. 1984. Easy calculations of lod scores and genetic risks on small computers. Am. J. Hum. Genet. **36:**460–465.

Lathrop, G. M., J. M. Lalouel, C. Julier, and J. Ott. 1985. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am. J. Hum. Genet. **37:**482–498.

Oberle, I., D. Drayna, G. Camerino, R. White, and J.-L. Mandel. 1985. The telomeric region of the human X chromosome long arm: presence of a highly polymorphic DNA marker and analysis of recombination frequency. Proc. Natl. Acad. Sci. USA **82:**2824–2828.

Patterson, M., M. Bell, W. Kress, K. E. Davies, and U. Froster-Iskenius. 1988*a*. Linkage studies in a large fragile X family. Am. J. Hum. Genet. **43:**684–688.

Patterson, M., M. Bell, C. Schwartz, and K. Davies. 1988*b*. Pulsed-field gel mapping studies in the vicinity of the fragile site at Xq27.3. Am. J. Med. Genet. **30:**581–591.

Patterson, M., S. Kenwrick, S. Thibodeau, K. Faulk, M. G. Mattei, J. F. Mattei, and K. E. Davies. 1987*a*. Mapping of DNA markers close to the fragile site on the human X chromosome at Xq27.3. Nucleic Acids Res. **15:**2639–2651.

Patterson, M., C. Schwartz, M. Bell, S. Sauer, M. Hofker, B. Trask, G. van den Engh, and K. E. Davies. 1987*b*. Physical mapping studies on the human X chromosome in the region Xq27-Xqter. Genomics **1:**297–306.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.

Thibodeau, S. N., H. R. Dorkins, K. R. Faulk, R. Berry, A. C. M. Smith, R. Hagerman. A. King, and K. E. Davies. 1988. Linkage analysis using multiple DNA polymorphic markers in normal families and in families with fragile X syndrome. Hum. Genet. **79:**219–227.

Vollrath, D., and R. Davis. 1987. Resolution of DNA molecules greater than 5 megabases by contour-clamped homogeneous electric fields. Nucleic Acids Res. **15:**7865–7876.

Wion, K. L., E. G. D. Tuddenham, and R. M. Lawn. 1986. A new polymorphism in the factor VIII gene for prenatal diagnosis of hemophilia A. Nucleic Acids Res. **14:**4535–4542.