Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 2701–2705, April 1988
Genetics

High recombination between two physically close human basement membrane collagen genes at the distal end of chromosome 13q

(DNA polymorphisms/chromosome 13/linkage/haplotypes/linkage disequilibrium)

ANNE M. BOWCOCK*, JOAN M. HEBERT*, ELLEN WIJSMAN*, INDER GADI†, LUCA L. CAVALLI-SFORZA*, AND CHARLES D. BOYD†

*Department of Genetics, Stanford University, Stanford, CA 94305; and †Department of Surgery, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, New Brunswick, NJ 08903

Contributed by Luca L. Cavalli-Sforza, December 29, 1987

ABSTRACT **Two basement membrane collagen genes coding for the proα1 chain and proα2 chain of type IV collagen map to 13q34 and are linked with a maximum likelihood estimate of recombination of 0.028 at a logarithm of odds (lod) score of 19.98. The single-copy sequence that identifies the locus *D13S3* is also closely linked to both collagen genes. Four enzymes reveal polymorphisms with *COL4A1*, and 10 haplotypes have been observed in Caucasoids. Within *COL4A1* a nonrandom association of alleles exists only between alleles defined by *Hae* III and those defined by the other three enzymes. A random association of alleles of *COL4A1* and *COL4A2* is observed. Between the two collagen genes were detected three meiotic recombination events that contributed to the estimate of 2.8% recombination. This is higher than expected for two genes that lie within 650 kilobases of each other. The lack of linkage disequilibrium between *COL4A1* and *COL4A2* is in agreement with the relatively high recombination that is observed.**

Collagens are the major structural protein in mammals. There are at least 10 different collagen families, which themselves can be composed of more than one type of collagen chain. So far at least 20 collagen chains are known to exist (1). The genes coding for these collagen chains are dispersed throughout the genome; so far, seven human collagen genes have been mapped to five chromosomes (2).

Type IV collagen is a nonfibrillar protein found exclusively in basement membranes (3). This basement membrane collagen molecule is a heterotrimer, composed of two genetically distinct polypeptide chains, proα1(IV) and proα2(IV) collagen (4). The genes coding for these human collagen chains have been mapped to the distal region of chromosome 13q. The proα1(IV) collagen gene (*COL4A1*) was mapped to chromosome 13 (4, 5) and localized to band q33-q34 by *in situ* hybridization (2, 6), and the proα2(IV) collagen gene (*COL4A2*) was recently also mapped to band q33-q34 by *in situ* hybridization (7–9). Close localization of these genes has prompted the suggestion that they may be physically linked and perhaps be as close as the genes comprising the multigene clusters of α- and β-globin (10). Close physical proximity between two collagen genes within the same family has not been reported previously. For example, in the case of the type 1 collagen family, the genes coding for the α1 and α2 chains lie on different chromosomes (11).

In this report we show that *COL4A1* and *COL4A2* are closely linked. Meiotic recombination between *COL4A1* and *COL4A2* is described. We have recently shown that *COL4A1* is closely linked to the locus *D13S3* (12, 13) and this marker is included in the analysis. Haplotypes for *COL4A1* that are used in the linkage analysis are reported, and linkage disequilibria within *COL4A1* and between the loci *D13S3*, *COL4A1*, and *COL4A2* are studied. The physical distance between *COL4A1* and *COL4A2* is investigated.

Although these data indicate a close proximity between the genes coding for the proα1 and proα2 chains of type IV collagen, significantly more recombination than expected is observed between them. This is supported by the failure to observe linkage disequilibrium between the two genes.

MATERIALS AND METHODS

Description and Source of Families and Probes. DNA from lymphoblastoid cell lines of 30 normal families was obtained from the Human Polymorphism Study Center [Centre d'Etude du Polymorphisme Humain (C.E.P.H.); Paris]. These families were used in the determination of *COL4A1* haplotypes and in pairwise and three-point linkage analysis between *COL4A1*, *COL4A2*, and *D13S3*. Probes identifying the basement membrane collagen genes were as follows: (*i*) A cDNA clone, HT39, which contains 1.7 kilobases (kb) of DNA sequences coding for the carboxyl-terminal noncollagenous domain of human proα2(IV) collagen (9) and the entire 3′ untranslated region of the proα2(IV) collagen mRNA. (*ii*) A cDNA clone, HT21, which contains 2.6 kb of DNA sequence coding for 185 amino acids of a collagenous portion of human proα1(IV) collagen, all of the globular carboxyl-terminal domain, and the 3′ untranslated region (4). (*iii*) The probe for the human single-copy sequence p9A7 (12), which identifies the locus *D13S3*.

Southern Analysis. DNA digestion, electrophoresis, Southern transfer, hybridization, and autoradiography were carried out as described (14). Cloned DNA fragments were isolated by electrophoresis in low-melting (FMC) agarose gels and were labeled with ^{32}P by the method of Feinberg and Vogelstein (15).

Linkage Analysis. Two- and three-point linkage analyses were performed by using the program MLINK from the LINKAGE package (16).

Pulsed-Field Gel Analysis. Preparation of high molecular weight DNA, restriction endonuclease digestion, electrophoresis, Southern transfer, hybridization, and autoradiography were carried out as described (17), except that 5% dextran sulfate was included in the hybridization solution. Probes for the *COL4A1* and *COL4A2* genes were used and are described above.

RESULTS

Restriction Fragment Length Polymorphism (RFLP) Screen. To locate RFLPs in the *COL4A2* gene, the probe

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RFLP, restriction fragment length polymorphism; lod, logarithm of odds; C.E.P.H., Centre d'Etude du Polymorphisme Humain.

Table 1. Frequency of RFLP haplotypes at human *COL4A1* in the Caucasoid population

Haplotype				
Hae III	*Hin*dIII	*Xmn* I	*Hin*fI	Frequency
1.65	2.1	5.0	1.95	0.297
0.95	3.2	5.0	1.95	0.297
0.95	2.1	5.0	1.95	0.183
0.95	3.2	5.0	1.1	0.063
1.20	2.1	4.7	1.95	0.029
1.65	3.2	5.0	1.95	0.063
0.95	2.1	5.0	1.1	0.046
1.20	3.2	4.7	1.95	0.010
1.65	3.2	4.7	1.95	0.006
0.95	3.2	4.7	1.1	0.006

One-hundred seventy-five chromosomes were studied. The alleles detected are indicated in kb.

HT39 was used on a panel of six unrelated individuals digested separately with 29 restriction enzymes. HT39 detected a simple two-allele polymorphism with the restriction endonuclease *Taq* I. Allele lengths are 2.4 kb and 2.0 kb. A sample of 97 chromosomes gave allele frequencies of 0.26 and 0.74, respectively. Constant bands revealed by this enzyme are 3.7, 3.2, 1.5, 0.77, and 0.64 kb. No polymorphism was detected with the following enzymes: *Alu* I, *Apa* I, *Bam*HI, *Ban* II, *Bcl* I, *Bgl* I, *Bgl* II, *Bsp*1286, *Dde* I, *Dra* I, *Eco*RI, *Hae* III, *Hin*cII, *Hin*fI, *Hin*dIII, *Hph* I, *Kpn* I, *Mbo* II, *Msp* I, *Pst* I, *Pvu* II, *Sac* I, *Sca* I, *Xba* I, *Xma* I, and *Xmn* I. RFLPs were also revealed with *Eco*RV and *Nci* I, but these enzymes were not used in the analysis, since the polymorphisms were less clear. *Eco*RV-defined alleles were 20 kb and 15 kb with frequencies of 0.76 and 0.24 and *Nci* I allele lengths were 5.2 kb and 1.35 kb. Both alleles were present at a frequency of 0.5.

Haplotype Analysis at the *COL4A1* Locus. Within *COL4A1* four RFLPs have been described (13): *Hae* III reveals a three-allele polymorphism with alleles of 1.65 kb, 1.2 kb, and 0.95 kb and frequencies of 0.39, 0.05, and 0.56, respectively; *Hin*dIII, *Hin*fI, and *Xmn* I each reveal two-allele polymorphisms; *Hin*dIII-defined alleles are 3.6 kb and 2.4 kb with frequencies of 0.55 and 0.45, respectively; *Hin*fI-defined alleles are 1.95 kb and 1.1 kb with frequencies of 0.81 and 0.19, respectively; and *Xmn* I-defined alleles are 5.0 kb and 4.7 kb with frequencies of 0.93 and 0.07, respectively.

Haplotypes within *COL4A1* were constructed by typing individuals in the C.E.P.H. pedigrees for the four *COL4A1* RFLPs described above. Of 24 possible haplotypes, 10 were observed. Table 1 describes these haplotypes and provides their frequencies in the Caucasoid population. The 3 most common haplotypes are also frequent in African and Chinese populations, though with differing frequencies (unpublished results).

Linkage Analysis. The linkage relationship between *COL4A1* and *COL4A2* was determined by using *COL4A1* haplotypes and the *Taq* I alleles of *COL4A2*. The *Hin*dIII alleles of *D13S3* were also included in the linkage analysis. The data on *D13S3* was partly determined by us and partly derived from data provided by the C.E.P.H. collaboration that had been used in the construction of a primary map of chromosome 13 (18). Two-point linkage analysis was performed in 30 families for all three pairs of markers and the results are given in Table 2. All markers are very closely linked, and in fact no recombinants were observed between *COL4A2* and *D13S3*, which demonstrated linkage with a logarithm of odds (lod) score of 6.86 at a maximum likelihood estimate of recombination ($\hat{\theta}$) at 0.00. The confidence interval is between $\theta = 0.000$ and 0.080 [the two recombination fractions that have lod scores $\approx$1 unit (lod scale) below the maximum lod score (19)]. The *COL4A1* and *COL4A2* genes are tightly linked, giving a maximum likelihood estimate of recombination ($\hat{\theta}$) of 0.028 with a lod score of 19.98. The confidence interval for recombination between these two genes is between 0.006 and 0.078. Three recombination events were detected between *COL4A1* and *COL4A2*. One event was of paternal origin; the other two events were observed together in a different family and were identical. The parental origin of the two recombination events that had occurred in this family could not be ascertained. Linkage between *COL4A1* and *D13S3* has been reported (13); however, inclusion of more data has increased the lod score to 28.52 at a maximum likelihood estimate of recombination of 0.025 and a confidence interval of $\theta = 0.007$–0.062. Four recombination events were detected between *COL4A1* and *D13S3*: two were of parental origin, one was of maternal origin, and one could have been from either parent. Sex differences in recombination rate were not significant.

Multipoint Analysis. Odds for the three possible orders were very similar and no order could be claimed to be most favored (Table 3).

Pulsed-Field Gel Electrophoresis. This technique was used to investigate the physical proximity of the two collagen genes, since it allows resolution of DNA fragments of at least between 50 and 3000 kb (20, 21). Hybridizations with the two probes were carried out using the same filters, since fragment sizes are dependent on several factors and can vary from filter to filter. Fig. 1 shows the same filter containing human genomic DNA digested with a variety of restriction endonucleases and hybridized to probes for *COL4A1* (Fig. 1*A*) and *COL4A2* (Fig. 1*B*). Fragment sizes observed with the two probes on this and other filters are listed in Table 4. Both probes hybridized to *Mlu* I fragments of 850 and 650 kb (lane 1) and to *Sac* II fragments of 900 and 825 kb (lane 7). *COL4A1* also hybridized to *Sac* II fragments of 725 and 600 kb, and *COL4A2* also hybridized to a *Sac* II fragment of 200 kb (not seen in Fig. 1). Both probes also hybridized to very large *Nru* I and *Not* I fragments (>1000 kb). Fragment lengths observed with the enzymes *Sfi* I, *Nae* I, and *Nar* I were smaller and of different lengths, except for a 90-kb *Nar* I fragment.

These results suggest that the two collagen genes may lie on the same *Mlu* I fragment of 650 kb and possibly the same *Sac* II fragment of 825 kb. Both may also be located on large *Nru* I and *Not* I fragments (>1000 kb). The larger 850-kb *Mlu* I and 900-kb *Sac* II fragments may be the results of incomplete digestion or may be due to differences in the methylation of *Mlu* I and *Sac* II sites in this region. A large number of smaller fragments is observed with the enzymes *Nae* I and *Nar* I and these may be due to partial digestion or methylation at these sites. The close proximity of the two collagen genes has been independently observed by Cutting *et al.* (22), who have used

Table 2. Results of two-point linkage analysis between *COL4A1*, *COL4A2*, and *D13S3*

Marker pair	Recombination fraction (θ)									
	0.00	0.001	0.01	0.05	0.1	0.2	0.3	0.4	$\hat{\theta}$	$\hat{Z}$
COL4A2–D13S3	6.86	6.85	6.74	6.24	5.60	4.23	2.76	1.25	0.00	6.86
COL4A1–COL4A2	−00	16.87	15.46	19.71	18.22	14.06	9.17	4.01	0.028	19.98
D13S3–COL4A1	−00	24.57	27.96	27.98	25.62	19.40	12.35	5.23	0.025	28.52

Table 3. Results of multipoint analysis for the three loci *COL4A1*, *COL4A2*, and *D13S3*

Locus order	Recombination estimate		Odds
D13S3–COL4A1–COL4A2	0.023	0.025	1
COL4A1–D13S3–COL4A2	0.024	0.011	1.1
COL4A1–COL4A2–D13S3	0.029	0.000	1.3

The odds reported are for a given order versus the odds for the order *D13S3–COL4A1–COL4A2*.

pulsed-field analysis to construct a map of this region, placing the two genes within 400 kb of each other.

Nonrandom Association of Alleles. Pairwise association of alleles within the *COL4A1* locus is shown in Table 5. There is an association between alleles defined by *Hae* III and those defined by the other three enzymes, *Hin*fI, *Xmn* I, and *Hin*dIII. This can be seen not only in the above three pairwise comparisons but also in comparisons of *Hae* III alleles versus the haplotypes constructed with the remaining three enzymes (χ^2 = 55.2, 3 df, data not shown).

The only significant between-locus association occurs for the *D13S3* locus and the *Hin*dIII alleles of *COL4A1* (χ^2 = 4.33, 1 df, $P < 0.05$). Linkage disequilibrium is not observed between any alleles belonging to *COL4A1* and *COL4A2* or between *COL4A2* and *D13S3* (Table 5).

DISCUSSION

Linkage analysis has shown that *COL4A1* and *COL4A2* are tightly linked with a maximum likelihood estimate of recombination between them of 2.8%. On average, the physical distance represented by 2.8% recombination is 2800 kb. This is derived from the estimate that 1 centimorgan (which represents 1% recombination) is about 1000 kb (24). The confidence limits for recombination between *COL4A1* and *COL4A2* of 0.6% and 7.8% can be interpreted in a similar fashion, to represent, on average, a distance of between 600 and 7800 kb. However, our physical mapping data place *COL4A1* and *COL4A2* to within 650 kb of each other, and Cutting *et al.* (22) have mapped them to within 400 kb of each other. The distance between *COL4A1* and *COL4A2* is therefore less than that set by the lower confidence limit, suggesting that recombination in this region is higher than average. Further evidence for this increased recombination is provided by the absence of linkage disequilibrium between the two genes, which is unlikely to be due entirely to the low level of information provided by the *Taq* I polymorphism of *COL4A2*.

Table 4. Sizes of fragments (kb) generated by pulsed-field gel electrophoresis with the *COL4A1* and *COL4A2* probes

Enzyme	Locus *COL4A1*	Locus *COL4A2*
Mlu I	850	850
	650	650
Nae I	600	450
	180	320
		250
Nar I	600	370
	290	280
	230	170
	90	90
Nru I	>1000	>1000
Sfi I	150	<50
Sac II	900	900
	825	825
	720	200
	600	
Not I	>1000	>1000

A possible explanation for the local increase in recombination is the presence of a recombination hot spot between the two collagen genes. It is difficult, however, to distinguish this hypothesis from that of a general expectation of increased recombination in terminal regions, which is observed at least in some chromosomes (see ref. 25 for examples). It is not known, in any case, if the increased recombination in terminal regions results from a local increase in the number or effectiveness of recombination hot spots, thereby generating greater discontinuity in the probability of recombination per megabase, or if the increased recombination is relatively continuously distributed.

Recombination hot spots may be a feature of DNA sequences between gene family members. For example, in the β-globin gene cluster on chromosome 11p, two independent meiotic recombination events have been observed (26, 27), and at the population level a region 5′ to the β-globin structural gene has 3–30 times more recombination than expected (28).

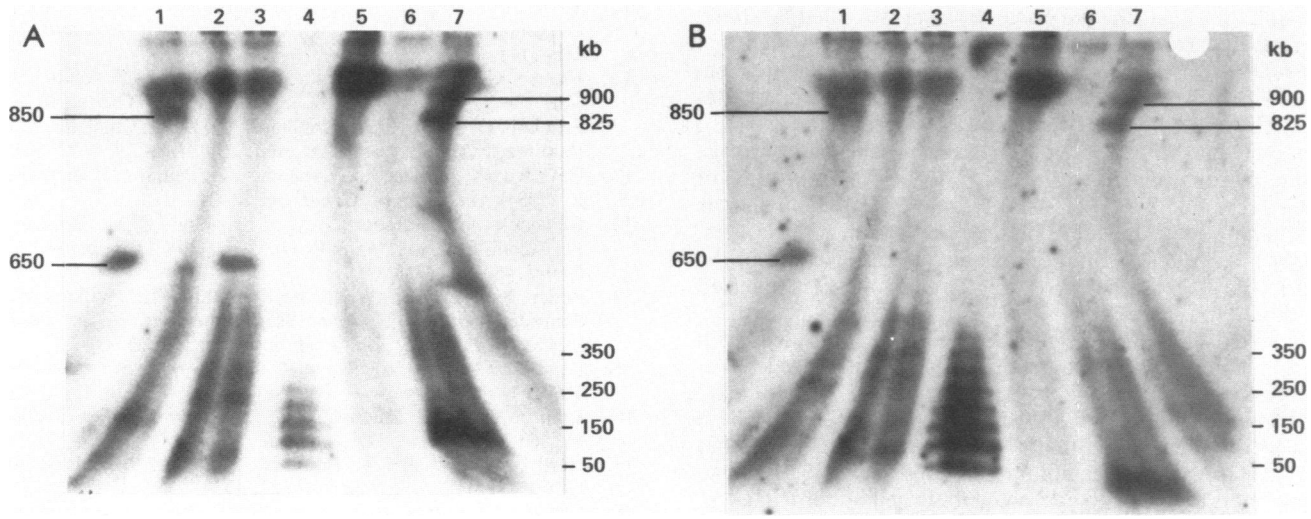

FIG. 1. Pulsed-field gel analysis of the human basement membrane collagen gene region. Genomic DNA was from the human lymphoblastoid line WT/51. The hybridization patterns with a *COL4A1* probe (*A*) and a *COL4A2* probe (*B*) on the same blot are shown. Genomic DNA was digested with the following restriction enzymes: *Mlu* I (lanes 1), *Nae* I (lanes 2), *Nar* I (lanes 3), *Nru* I (lanes 5), *Sfi* I (lanes 6), and *Sac* II (lanes 7). Lanes 4, ladder of λ concatamers used as size markers.

Table 5. Pairwise association of alleles in the *COL4A1/COL4A2/D13S3* region in Caucasoids

	COL4A1 *Hae* III	*COL4A1* *Hin*dIII	*COL4A1* *Xmn* I	*COL4A1* *Hin*fI	*COL4A2* *Taq* I	*D13S3* *Hin*dIII
COL4A1 *Hae* III	—	0.09 (40.8)*	0.15 (73.3)*	0.04 (13.9)†	0.06 (0.88)	0.12 (2.44)
COL4A1 *Hin*dIII	—	—	0.10 (2.32)	0.05 (0.44)	0.06 (0.85)	0.15 (4.33)‡
COL4A1 *Xmn* I	—	—	—	0.04 (0.25)	0.05 (0.67)	0.04 (0.25)
COL4A1 *Hin*fI	—	—	—	—	0.05 (0.37)	0.12 (2.18)
COL4A2 *Taq* I	—	—	—	—	—	0.05 (0.55)

The correlation coefficient r is shown for each enzyme pair. Its corresponding χ^2 or a measure of similar meaning is shown underneath in parentheses. N, the number of chromosomes included in the analysis, was from 140 to 248. In all rows except the first, $\chi^2 = Nr^2$ with 1 df. In the first row, with *Hae* III there are three alleles and χ^2 values with 2 df; the measure of correlation is that used by Migone *et al.* (23).
*$P << 0.001$.
†$P < 0.001$.
‡$P < 0.05$.

All together three recombinants were observed between *COL4A1* and *COL4A2*, and, interestingly, two of the events occurred in the same family and were identical. The genotypes of these individuals differed at other loci, eliminating the possibility that the same DNA sample was mistakenly used. Further experiments will help clarify whether or not the same recombination event occurred twice in the same family due to chance.

Significant linkage disequilibrium is detected within the *COL4A1* gene between alleles defined by *Hae* III and those defined by the other three enzymes, *Hin*fI, *Xmn* I, and *Hin*dIII. This nonrandom association is confined to comparisons involving *Hae* III, and a random association of alleles revealed by the other three enzymes is observed. *Hae* III is also in significant disequilibrium with the haplotypes constructed with the three enzymes. These results suggest that the mutations causing the polymorphisms revealed by *Hin*fI, *Xmn* I, and *Hin*dIII occurred earlier in evolution than that revealed by *Hae* III. Alleles revealed by these three enzymes then had time to reach equilibrium before the mutation causing the *Hae* III polymorphism occurred. It is also possible that natural selection has maintained the linkage disequilibrium related to *Hae* III. The lack of association of alleles revealed by *Hin*fI, *Xmn* I, and *Hin*dIII could also be due to mutational hot spots or gene conversion at these sites. Weak linkage disequilibrium has also been observed in the proα2(I) collagen gene (*COL1A2*) for pairs of alleles revealed by three enzymes (29).

In spite of the existence of 2.5% recombination, there is a suggestion of linkage disequilibrium at a significance level of 5% between *D13S3* and the *Hin*dIII alleles of *COL4A1*. Perhaps the estimate of 2.5% recombination between these two loci is too high so that recombination is closer to that provided by the lower confidence limit of 0.7%. Linkage disequilibrium was not detected between other *COL4A1* alleles and *D13S3*, between *COL4A1* and *COL4A2*, or between *COL4A2* and *D13S3*.

Multipoint analysis did not favor any particular order for the three loci. Pulsed-field mapping involving *D13S3* may determine which order is correct.

Close linkage between two members of a collagen gene family has not been described previously. The close proximity of these two basement membrane collagen genes at human chromosome 13q33-34 and the high level of recombination between them may be of considerable significance to the evolution of these complex multi-exon genes. Furthermore, this detailed linkage analysis will serve as the basis for future considerations of the role of the 13q33-34 region in a variety of human diseases, involving not only basement membrane collagen but possibly also other genes known to occur at this region, such as those coding for factors VII and X (30).

We are most grateful to Dr. Jean Dausset and the Centre d'Etude du Polymorphisme Humain consortium (Paris) for the family DNA used in the study. Dr. W. Cavenee (Ontario, Canada) kindly provided probe p9A7. We also thank Drs. P. Byers (Seattle), R. Spielman (Philadelphia), and J. Rosenbloom (Philadelphia) for comments on the manuscript. Mrs. Berta Young and Miss Laleh Daneshvar provided excellent technical help. This work was supported in part by National Institutes of Health Grants GM 28428 and AM 22051.

1. Bornstein, P. & Sage, H. (1980) *Annu. Rev. Biochem.* **49,** 957–1003.
2. Boyd, C. D., Weliky, K., Toth-Fejel, S. E., Deak, S. B., Christiano, A. M., Mackenzie, J. W., Sandell, L. J., Tryggvason, K. & Magenis, E. (1986) *Hum. Genet.* **74,** 121–125.
3. Timpl, R., Wiedemann, H., Van Delden, V., Furthmayr, H. & Kuhn, K. (1981) *Eur. J. Biochem.* **120,** 203–211.
4. Pihlajaniemi, T., Tryggvason, K., Myers, J. C., Kurkinen, M., Lebo, R., Cheung, M.-C., Prockop, D. J. & Boyd, C. D. (1985) *J. Biol. Chem.* **260,** 7681–7687.
5. Solomon, E., Hiorns, L. R., Spurr, N., Kurkinen, M., Barlow, D., Hogan, B. L. M. & Dagleish, R. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3330–3334.
6. Emanuel, B. S., Sellinger, B. T., Gudas, L. J. & Myers, J. C. (1986) *Am. J. Hum. Genet.* **38,** 38–44.
7. Solomon, E., Hall, V. & Kurkinen, M. (1987) *Ann. Hum. Genet.* **51,** 125–127.
8. Griffin, C., Emanuel, B. S., Hansen, J. R., Cavenee, W. K. & Myers, J. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 512–516.
9. Boyd, C. D., Toth-Fejel, S. E., Gadi, I. K., Kurkinen, M., Kolbe, M., Hagen, I. K., Mackenzie, J. W. & Magenis, E. (1987) in *ICSU Short Reports 7*, eds. Voellmy, R. W., Ahmad, F., Black, S., Burgess, D. R., Rotundo, R., Scott, W. A. & Whelan, W. J. (Cambridge Univ. Press, Cambridge, U.K.), p. 142.
10. Flavell, R. A. & Grosveld, F. G. (1983) in *Eukaryotic Genes: Their Structure, Activity and Regulation*, eds. Maclaen, N., Gregory, S. P. & Flavell, R. A. (Butterworth, London), pp. 299–314.
11. Huerre, C., Junien, C., Weil, D., Chu, M.-L., Morobito, M., Van Cong, N., Myers, J. C., Foubert, C., Gross, M.-S., Prockop, D. J., Bove, A., Kaplan, J.-C., De La Chapelle, A. & Ramirez, F. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6627–6630.
12. Cavenee, W., Leach, R., Mohandas, T., Pearson, P. & White, R. (1984) *Am. J. Hum. Genet.* **36,** 10–24.
13. Bowcock, A. M., Hebert, J. M., Christiano, A. M., Wijsman, E., Cavalli-Sforza, L. L. & Boyd, C. D. (1987) *Cytogenet. Cell Genet.* **45,** 234–236.
14. Feder, J., Yen, L., Wijsman, E., Wang, L., Wilkins, L., Schroder, J., Spurr, N., Cann, H., Blumenberg, M. & Cavalli-Sforza, L. L. (1985) *Am. J. Hum. Genet.* **37,** 635–649.
15. Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137,** 266–267.
16. Lathrop, G. M. & Lalouel, J.-M. (1984) *Am. J. Hum. Genet.* **36,** 460–465.
17. Hardy, D. A., Bell, J. I., Long, E., Lindsten, T. & McDevitt, H. O. (1986) *Nature (London)* **323,** 453–455.
18. Leppert, M., Cavenee, W., Callahan, P., Holm, T., O'Connell, P., Thompson, K., Lathrop, G. M., Lalouel, J.-M. & White, R. (1986) *Am. J. Hum. Genet.* **39,** 425–437.
19. Conneally, P. M., Edwards, J. H., Kidd, K. K., Lalouel, J. M., Morton, N. E., Ott, J. & White, R. (1985) *Cytogenet. Cell Genet.* **40,** 356–359.

20. Schwartz, D. C. & Cantor, C. R. (1984) *Cell* **37,** 67–75.
21. Carle, G. F. & Olson, M. V. (1984) *Nucleic Acids Res.* **12,** 5647–5664.
22. Cutting, G. R., Kazazian, H. H., Jr., Antonarakis, S. E., Killen, P. D. & Yamada, Y. (1987) *Am. J. Hum. Genet.* **41,** A613 (abstr.).
23. Migone, N., De Lange, G., Piazza, A. & Cavalli-Sforza, L. L. (1985) *Am. J. Hum. Genet.* **37,** 1146–1163.
24. Renwick, J. H. (1969) *Br. Med. Bull.* **25,** 65–73.
25. Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T. P., Bowden, D. W., Smith, D. R., Lander, E. S., Botstein, D., Akots, G., Rediker, K. S., Gravius, T., Brown, V. A., Rising, M. B., Parker, C., Powers, J. A., Watt, D. E., Kauffman, E. R., Bricker, A., Phipps, P., Muller-Kahle, H., Fulton, T. R., Ng, S., Schumm, J. W., Braman, J. C., Knowlton, R. G., Barker, D. F., Crooks, S. M., Lincoln, S. E., Daly, M. J. & Abrahamson, J. (1987) *Cell* **51,** 319–337.
26. Gerhard, D. S., Kidd, K. K., Kidd, J. R., Egeland, J. A. & Housman, D. I. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7875–7879.
27. Old, J. M., Heath, C., Fitches, A., Thein, S. L., Jeffreys, A. J., Petrov, M., Modell, B. & Weatherall, D. J. (1986) *J. Med. Genet.* **23,** 14–18.
28. Chakravarti, A., Buetow, K. H., Antonarakis, S. E., Waber, P. G., Boehm, C. D. & Kazazian, H. H. (1984) *Am. J. Hum. Genet.* **36,** 1239–1258.
29. Borresen, A.-L., Berg, K., Tsipouras, P., Dickson, L. A., Prockop, D. J. & Ramirez, F. (1985) in *Medical Genetics: Past, Present, Future* (Liss, New York), pp. 37–51.
30. Pfeiffer, R. A., Ott, R., Gilenkrantz, S. & Alexandre, P. (1982) *Hum. Genet.* **62,** 358–360.