Am J Hum Genet 39:11–24, 1986

DNA "Fingerprints" and Segregation Analysis of Multiple Markers in Human Pedigrees

ALEC J. JEFFREYS,[1] VICTORIA WILSON,[1] SWEE LAY THEIN,[2]
DAVID J. WEATHERALL,[2] AND BRUCE A. J. PONDER[3]

SUMMARY

Tandem-repetitive DNA hybridization probes based on a putative human recombination signal detect multiple polymorphic minisatellite fragments in human DNA. The genetic complexity of the resulting individual-specific DNA "fingerprints" was investigated by studying a large sibship affected by neurofibromatosis and a more extensive pedigree segregating for two different hemoglobinopathies. The segregation of up to 41 different heterozygous DNA fragments from each parent could be analyzed in a single sibship, using two different repeat probes. Most of these variable DNA fragments could not be paired as alleles, to an extent which suggests that the DNA fingerprints are together derived from ~ 60 heterozygous loci (~ 120 variable fragments), only a proportion of which can be scored in a given individual. Two or three of the DNA fragments detected by one probe showed tight linkage and may be derived from long minisatellite(s) that are cleaved to produce more than one polymorphic DNA fragment. Excluding allelic and linked DNA fragments, almost all remaining scorable fragments segregated independently, allowing up to 34 unlinked loci to be examined simultaneously. These loci are scattered over most or all of the human autosomes. Minisatellite probes are therefore suitable for rapid marker generation and can be applied to linkage analysis in human pedigrees.

Received December 2, 1985.

This work was supported by a grant to A. J. J. and a Training Fellowship to S. L. T. from the Medical Research Council (MRC). B. A. J. P. is supported by a joint grant to the Institute of Cancer Research from the MRC and Cancer Research Campaign. A. J. J. is a Lister Institute Research Fellow. The minisatellite probes are the subject of Patent Applications. Commercial inquiries should be addressed to the Lister Institute of Preventive Medicine, Royal National Orthopaedic Hospital, Brockley Hill, Stanmore, Middlesex HA7 4LP, U.K.

[1] Department of Genetics, University of Leicester, Leicester LE1 7RH, U.K.

[2] MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, U.K.

[3] Institute of Cancer Research, Royal Cancer Hospital, Sutton, Surrey SM2 5PX, U.K.

INTRODUCTION

Restriction fragment length polymorphisms (RFLPs) provide a rich source of genetic markers in man and have been extensively used in linkage analysis, studies of chromosome abnormalities in cancer, and antenatal diagnosis of genetic disease (see [1–4]). The overall level of genetic variability in human DNA is, however, low, and most single-copy DNA probes detect few if any RFLPs [5, 6]. Furthermore, most of these RFLPs result from base substitutional gain or loss of a restriction endonuclease cleavage site to produce dimorphisms that can be relatively uninformative in pedigree analysis.

Hypervariable minisatellite regions, consisting of tandem repeats of a short sequence and showing multiallelic variation in repeat copy number, exist in human DNA and provide highly informative genetic markers [7–10]. A subset of human minisatellites share a 10–16 base pair (bp) common "core" sequence embedded within each repeat unit [11]. This core sequence may be a recombination signal in human DNA implicated in the generation of hypervariable tandem-repetitive regions. Hybridization probes consisting of tandem repeats of the core sequence can detect multiple hypervariable DNA fragments in Southern blot analyses of human DNA [11]. Two repeated core probes have been developed, each of which detects a different pattern of variable DNA fragments to produce distinct DNA "fingerprints" that are individual-specific and show somatic and germ-line stability [12]. To estimate the number of hypervariable loci that can be studied using these probes, and to determine the feasibility of using DNA fingerprints for linkage analysis in man, we have investigated the DNA fingerprints of two large families, one segregating for neurofibromatosis and the other for hereditary persistence of fetal hemoglobin apparently determined by an autosomal dominant gene not linked to the β-globin gene cluster.

MATERIALS AND METHODS

DNA Isolation

Fresh blood was diluted with an equal volume of 1 × SSC (SSC, saline sodium citrate, 0.15 M NaCl, 15 mM trisodium citrate, pH 7.0), layered onto Histopaque-1077 (Sigma, St. Louis, Mo.), and nucleated cells collected by centrifugation. Alternatively, frozen blood was thawed in 2 vol 1 × SSC and nucleated cells plus nuclei pelleted by centrifugation at 10,000 *g* for 15 min. High molecular weight DNA was prepared as described [5].

Southern Blot Analysis

Five-microgram samples of human DNA were digested with 20 U of *Hin*fI in the presence of 4 mM spermidine trichloride at 37° C for 2 hrs, and recovered after phenol extraction by ethanol precipitation. Restriction digests were dissolved in 16 μl H_2O plus 4 μl gel loading mix (12.5% ficoll 400, 0.2% bromophenol blue, 0.2 M Tris acetate, 0.1 M Na acetate, 1 mM EDTA, pH 8.3) and 2 μl 5mg/ml ethidium bromide, and loaded onto a horizontal agarose gel (0.7% Sigma Type I agarose in 40 mM Tris acetate, 20 mM Na acetate, 0.2 mM EDTA, 0.5 μg/ml ethidium bromide, pH 8.3; gels 0.7-cm thick by 20-cm or 35-cm long). After equilibration for 10 min, gels were electrophoresed at 2 V/cm for 24–48 hrs, until all DNA fragments less than 1.5 kilobases (kb) long had electrophoresed

off the gel. DNA was transferred by blotting onto a nitrocellulose filter (Sartorius, 0.45 μm pore size) [13]. [^{32}P]-labeled single-stranded probe DNA was prepared from the human minisatellite M13 recombinants 33.6 and 33.15, hybridized to Southern blots in 1 × SSC at 65° C, and autoradiographed as described [11, 12].

Data Analysis

Storage of segregation data and analysis of linkage were performed on a BBC Model B microcomputer, using programs written by A. J. J. Lod scores were calculated as described [14, 15].

RESULTS

DNA Fingerprint Probes

Two minisatellite hybridization probes used in this study are described by Jeffreys et al. [11, 12]. Probe 33.15 consists of a cloned human minisatellite comprised of 29 repeats of a 16-bp variant of the core sequence. The repeat unit of the minisatellite in probe 33.6 is a diverged trimer of the most conserved 11-bp 3′ end of the core sequence and is repeated 18 times. The sequences of the core and probe repeat units are:

core G G A G G T G G G C A G G A $\begin{matrix}A\\G\end{matrix}$ G

33.15 A G A G G T G G G C A G G T G G

33.6 $(A\ G\ G\ G\ C\ T\ G\ G\ A\ G\ G)_3$.

The difference both in sequence and in repeat length of probes 33.6 and 33.15 results in their detecting different patterns of long hypervariable minisatellite fragments in *Hin*fI digests of human DNA [12]. This 4-bp restriction endonuclease maximizes the resolution of allelic variation in minisatellites by releasing long tandem-repetitive minisatellites in DNA fragments with little flanking DNA.

Analysis of DNA Fingerprints in a Large Sibship

To investigate the segregation of individual minisatellite DNA fragments, we have studied a large sibship of 11 English individuals segregating for the peripheral form of neurofibromatosis (von Recklinghausen disease), an autosomal dominant disorder associated with tumors of the peripheral and central nervous system [16, 17]. No genetic markers have yet been linked to this disease [18]. Blood DNA fingerprints, detected by probes 33.6 and 33.15, of the 11 children (five affected, one probably affected, and five unaffected) are compared with their unaffected father in figure 1. Resolution of minisatellite DNA fragments was maximized by electrophoresis in 35-cm-long agarose gels.

We have previously shown that many of these hypervariable minisatellites, particularly the largest DNA fragments, have low allele frequencies and are seldom shared by unrelated individuals [12]. As predicted, almost all of the resolvable fragments in the neurofibromatosis family are present in the heterozygous state and are transmitted to only some of the progeny. Even though the

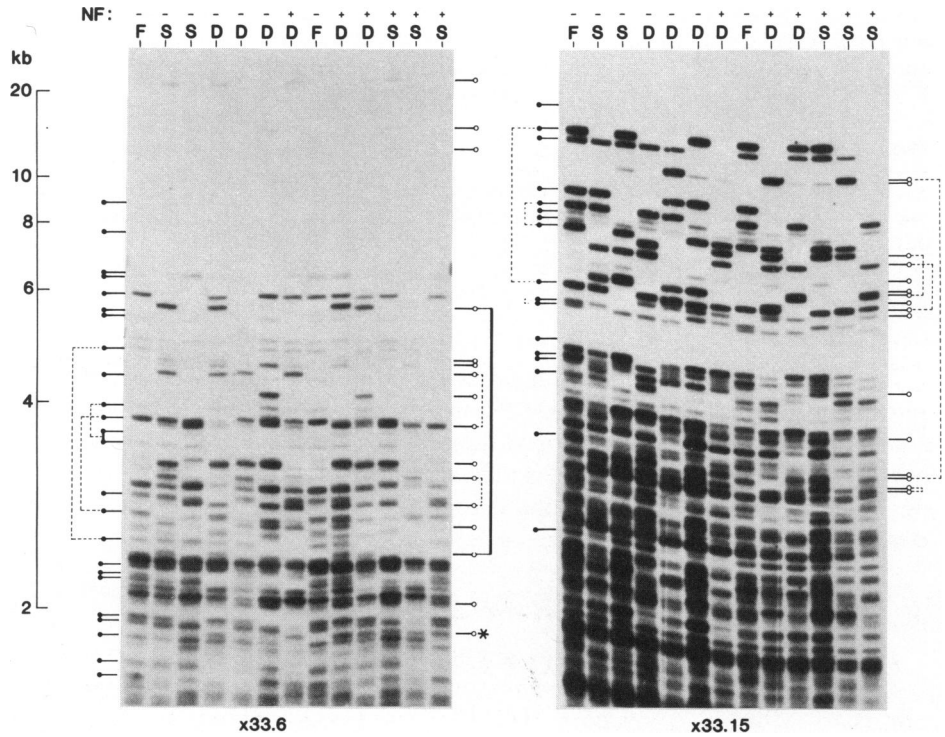

FIG. 1.—Segregation of hypervariable minisatellite fragments in the DNA fingerprints of a sibship affected by neurofibromatosis. Blood DNA samples digested with *Hin*fI were electrophoresed on a 35-cm-long 0.7% agarose gel and Southern blot hybridized to minisatellite probes 33.6 and 33.15. DNA fingerprints are shown for the unaffected father (*F*), five sons (*S*), and six daughters (*D*); the affected mother was not available for study. Offspring affected by multiple neurofibromata are indicated (+) (in the left-most daughter marked as affected, the diagnosis is suspected but awaits confirmation); the remaining offspring show no sign of neurofibromatosis. Resolved paternal (●) and maternal (○) heterozygous DNA fragments are indicated, and their segregation into offspring was scored directly from original autoradiographs taken at short, medium, and long exposures. Only those DNA fragments were scored whose positions and relative intensities in each offspring matched those in the parent. Linked pairs AB of parental DNA fragments which segregate AB or -- into offspring are joined by a continuous line; alleles that segregate A- and -B are joined by *a dotted line*. One maternal fragment that shows evidence of linkage in coupling to neurofibromatosis is marked with *an asterisk;* all five definitely affected offspring plus the one in whom neurofibromatosis is suspected have inherited this fragment, and four out of five unaffected children do not have this band. This gives a concordance of 9/10 between inheritance of this band and of neurofibromatosis, or 10/11 if the suspected case is scored as positive.

DNA of the affected mother is unavailable, maternally derived minisatellite fragments can be readily identified as fragments present in some offspring but absent from the father. Paternal fragments can similarly be identified. Using both probes 33.6 and 33.15, it is possible to score the segregation of 41 paternal and 32 maternal DNA fragments in this sibship (fig. 1, table 1). Numerous additional polymorphic fragments also exist, but have either been electrophoresed off the gel or are incompletely resolved and cannot therefore be reliably scored.

TABLE 1

SUMMARY OF MINISATELLITE MARKERS IN THE NEUROFIBROMATOSIS FAMILY

	FATHER		MOTHER	
PROBE	33.6	33.15	33.6	33.15
No. fragments scored (n)	24	17	16	16
No. allelic pairs (a)	3	3	2	4
No. linked pairs (b)	1	0	1	0
No. different loci scored (L)	20	14	13	12
Estimated total no. loci (N)	43	23	27	16

NOTE: The no. different loci (L) scored is given by $n - a - b$. The entire DNA fingerprint, including unresolved and therefore unscored fragments, is derived from N heterozygous loci ($2N$ fragments). Assuming that the $(n - b)$ distinct fragments scored are a random sample of the $2N$ bands in a DNA fingerprint, then the estimated total number of hypervariable loci N detected by a given probe is related to the number of allelic pairs a by

$$N = \frac{1}{2}\left[\frac{(n - b)(n - b - 1)}{2a} + 1\right]$$

(J. F. Y. Brookfield and B. English, personal communication).

Mendelian Inheritance of Hypervariable DNA Fragments

Heterozygous paternal DNA fragments are transmitted on average to 53% of the progeny. Similarly, maternal fragments showed 48% transmission, again consistent with 1:1 segregation (table 2). Furthermore, the number of children receiving each fragment followed the expected binomial distribution, in which the proportion of parental fragments that are transmitted to precisely r children in the sibship of 11 is $^{11}Cr/2^{11}$ (table 2). We conclude that these DNA fragments show Mendelian inheritance and that the scoring of parental bands, particularly the smaller and less well-resolved fragments, is not significantly influenced by possible cases of segregation of two or more superimposed bands that would give an apparent $\geq$ 75% transmission frequency. The correct maternity and paternity of all sibs is also established by these DNA fingerprints.

Identification of Allelic and Linked Pairs of DNA Fragments

By pairwise comparisons of the segregation patterns of all paternal or maternal DNA fragments in this large sibship, it is possible to identify allelic pairs of fragments plus pairs that show tight linkage in coupling (the odds of chance cosegregation of a given pair of bands in this sibship is 1/2,048). Several instances of allelic pairs of both paternal and maternal fragments could be identified with both probes (fig. 1). Probe 33.6 also detected a linked pair of fragments in both the mother and father. A similar linkage was found in a second pedigree (see below), which suggests that at least one of the hypervariable regions hybridizing to probe 33.6 is a long minisatellite/satellite that contains internal cleavage site(s) for *Hin*fI and is therefore cleaved to produce two or more fragments that cosegregate as a minisatellite "haplotype" in pedigrees. None of the polymorphic DNA fragments scored using probe 33.15 were present in the set of fragments detected by 33.6; any such fragment that hybridized

TABLE 2

SEGREGATION OF HYPERVARIABLE FRAGMENTS IN THE NEUROFIBROMATOSIS FAMILY

	FATHER							MOTHER						
	TRANSMISSION OF													
	SINGLE FRAGMENT		PAIR (AB or --)					SINGLE FRAGMENT		PAIR (AB or --)				
				Exp.							Exp.			
TRANSMISSION TO NO. CHILDREN (*r*)	Obs.	Exp.	Obs.	U	30	20	10	Obs.	Exp.	Obs.	U	30	20	10
0	0	0.02	(0)	0	1	2	9	(0)	0.02	(0)	0	0	1	6
1	0	0.2	4	3	5	8	19	1	0.2	0	2	3	5	13
2	1	1.1	16	15	18	22	31	2	0.9	5	8	10	13	19
3	1	3.3	48	45	47	49	51	3	2.6	26	24	26	27	28
4	5	6.7	90	90	89	85	76	5	5.2	48	48	47	45	38
5	8	9.2	113	127	121	114	95	6	7.2	61	68	64	59	45
6	12	9.2	120	127	121	114	95	5	7.2	77	68	64	59	45
7	10	6.7	89	90	89	85	76	6	5.2	52	48	47	45	38
8	4	3.3	58	45	47	49	51	4	2.6	23	24	26	27	28
9	0	1.1	19	15	18	22	31	0	0.9	7	8	10	13	19
10	0	0.2	4	3	5	8	19	0	0.2	1	2	3	5	13
11	(0)	0.02	(0)	0	1	2	9	(0)	0.02	(0)	0	0	1	6
Transmission frequency	53.0 ± 2.4%							47.7 ± 2.7%						

NOTE: To study the transmission frequency of hypervariable fragments (fig. 1), the no. fragments detected by probes 33.6 and 33.15, out of n scored, which were transmitted to precisely r children in the sibship of 11 was compared with the expected no. given by the binomial distribution $({}^{11}Cr/2^{11}) \cdot n$ assuming 50% transmission. Fragments present in all children may be from homozygous loci and were ignored. Maternal fragments not transmitted to any children could not be scored since maternal DNA was not available. The mean transmission frequencies (± SEM) are also given.

Linkage between pairs of fragments AB in each parent was investigated by scoring the no. offspring who were concordant for AB (either AB or --), using all possible pairwise comparisons of paternal or maternal fragments having first excluded alleles and linked bands (i.e., L loci were analyzed in each parent, see table 1, giving $[L(L-1)/2]$ pairwise comparisons). Pairs of fragments that fall into the zero- or 11-children classes represent alleles or tightly linked pairs, respectively; by definition, no pairs fall into either class. The observed distribution is compared with that expected if all L loci are unlinked (U), in which case the no. pairwise comparisons that give precisely r (AB or --) offspring is given by the binomial distribution $({}^{11}Cr/2^{11}) \cdot [L(L-1)/2]$. The distribution is also compared with that expected if the L loci are clustered and spaced uniformly, with adjacent loci being separated by a recombination frequency θ (10, 20 or 30 cM apart). The cluster will therefore be spread over $(L-1)\psi$ map units, where $\psi = -\frac{1}{2}\ln(1-2\theta)$. For L loci (sampled at one or other allele at random), the no. pairwise comparisons that give precisely r (AB and --) offspring in the sibship of 11 is given by:

$$\sum_{i=1}^{i=L-1} (L-i) \cdot {}^{11}C_r \left[\frac{x_i^{11-r}(1-x_i)^r + x_i^r(1-x_i)^{11-r}}{2} \right] ,$$

where x_i is the recombination fraction between two loci $i\psi$ map units apart; x_i is given by the mapping function [19]

$$x_i = \frac{1}{2}(1 - e^{-2i\psi}) .$$

to both probes would have been detected as bands of equal size that were transmitted from the same parent to the same children (i.e., "linked"). These two probes therefore hybridize to essentially completely different subsets of human minisatellites. In addition, no bands detected by probe 33.6 were allelic to, or linked with, any fragments hybridizing to 33.15.

Number and Distribution of Hypervariable Loci Detected by Probes 33.6 and 33.15

By eliminating alleles and linked fragments, we conclude that 34 and 25 distinct loci separable by at least one recombination event have been scored in the father and mother, respectively (table 1). For ~ 80% of loci, only one of the two alleles is resolved (fig. 1), and the second allele is probably located in the poorly resolved complex of shorter minisatellite fragments. This implies that large differences in minisatellite allele lengths must exist, arising presumably by unequal exchange in these tandem repetitive regions; several allelic pairs identified in figure 1 do indeed show substantial length differences. From the proportion of bands that can be paired into alleles, it is possible to estimate that the total number of heterozygous loci present in the entire DNA fingerprints detected by probes 33.6 and 33.15 is approximately 43–66, of which roughly half can be scored in each parent (table 1). It is not possible to determine allelism between paternal and maternal fragments in this sibship.

All of the paternal loci scored are autosomal and do not show specific transmission either into daughters (X linkage) or sons (Y linkage). Furthermore, all pairs AB of paternal DNA fragments (excluding allelic and tightly linked bands) apparently segregate independently into offspring, to give on average equal numbers of (AB, --) and (A- , -B) progeny; precise numbers followed the expected binomial distribution for unlinked loci (table 2). Maternal DNA fragments behaved similarly. More detailed analysis suggests that if these loci are clustered, then the minimal locus-to-locus spacing must be at least 30 cM (46 map units); any closer spacing would generate significant numbers of pairs of fragments that tend to cosegregate (linked in coupling) or segregate as pseudoalleles (linked in repulsion) (table 2). The resolvable minisatellite loci must therefore be spread over at least half of the 3,300-cM-long human genome [20] and, additionally, must therefore be scattered over many or all of the human autosomes.

One minisatellite fragment from the affected mother shows weak evidence of linkage in coupling with neurofibromatosis (fig. 1). If the one individual in whom neurofibromatosis is suspected but not confirmed were scored, for the sake of argument, as positive, then 10/11 children would be concordant for this fragment and the disease ($\hat{z} = 1.55$ at $\hat{\theta} = 0.1$).

DNA Fingerprints of an Extended Pedigree: Possible Linkage to HPFH

We have extended this analysis of DNA fingerprints to a more extensive 4-generation pedigree of Gujarati Asians that is segregating both for β-thalassemia and for the heterocellular form of hereditary persistence of fetal hemoglobin (HPFH). As shown in figure 2, elevation of HbF is transmitted

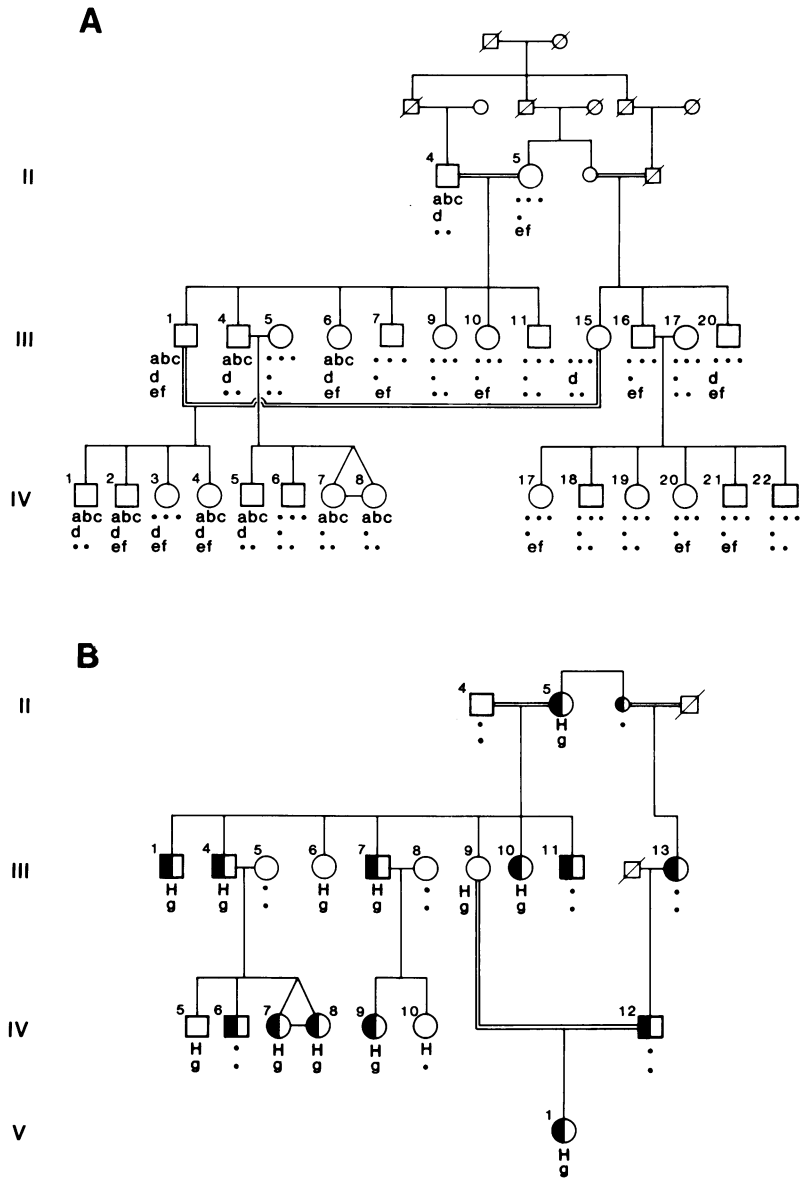

FIG. 2.—Cosegregation of minisatellite fragments, and of HPFH and a minisatellite, in an extended Gujarati pedigree. *A*, The segregation of 30 hypervariable fragments from II 4 and 27 fragments from II 5 into offspring III 1–11 was screened for possible linkage of pairs AB of parental fragments; possible examples of linkage showing at least 6/7 (AB, --) offspring were further examined in additional relatives. The two clearest examples of linkage are shown (*a–f*, presence of fragments *a–f* in an individual; •, fragment absent). Fragments *a–c* and *e,f* each show perfect cosegregation; fragment *d* tends to cosegregate with *a–c*, but sibship IV 1–4 is uninformative and identical twins IV 7,8 are recombinant, having inherited *a–c* but not *d*. *B*, Inheritance of β-thalassemia trait (◐, ◧), HPFH (*H*), and minisatellite fragment *g*. Individuals were scored as having HPFH if they showed > 1% HbF (normal) or > 3% HbF (β-thalassemia trait). HPFH and β-thalassemia trait segregate independently in III 1–11 and IV 5–8 and are therefore determined by unlinked loci. Fragment *g* cosegregates with HPFH in all individuals examined, except for individual IV 10 who is recombinant.

independently of β-thalassemia trait and is apparently determined by an autosomal dominant gene unlinked to the β-globin gene cluster (S. L. T. and D. J. W., manuscript in preparation). A similar Sardinian pedigree has been reported by Gianni et al. [21].

DNA fingerprints obtained with probes 33.6 and 33.15 have been studied in this pedigree (examples shown in fig. 3). Thirty variable fragments were scored in the grandfather (II 4) and 27 fragments in the grandmother (II 5). Study of their seven offspring (III 1–11) indicated that these fragments were derived from at least 22 distinct unlinked paternal and 18 maternal autosomal loci, using the criteria described for the neurofibromatosis family. The remaining DNA fragments showed evidence of allelism or linkage to other fragments, although proof with this small sibship is not possible (a given pair of parental DNA fragments has a 1/64 chance of fortuitously being transmitted either linked or as alleles in a sibship of seven).

Further evidence of linkage was sought in additional members of the pedigree, and the two strongest cases of linkage are shown in figures 2 and 3. Fragments a, b, and c detected by probe 33.6 are transmitted in perfect linkage from II 4 into his children (III 1–11) and, thence, grandchildren (IV 1–8) ($\hat{z} = 7.82$ at $\hat{\theta} = 0$ for three markers). As discussed above, this suggests that fragments a–c represent a minisatellite "haplotype" derived from a single hypervariable locus. Band d detected by probe 33.15 also shows evidence of linkage to bands a–c; however, one sibship (IV 1–4) is uninformative since both parents carry fragment d, and another (IV 5–8) contains a recombinant (identical twins IV 7 and 8). The evidence of linkage between band d and the a–c cluster is therefore weak ($\hat{z} = 1.30$ at $\hat{\theta} = .1$). Maternal bands e (detected by 33.6) and f (detected by 33.15) also show tight linkage both in the descendants of II 4 and II 5, and in an additional related sibship IV 17–22 ($\hat{z} = 4.52$ at $\hat{\theta} = 0$). Since probes 33.6 and 33.15 detect different sets of minisatellites and do not cross-hybridize to fragments e and f, these fragments may represent an example of authentic linkage between two different autosomal minisatellite loci. Finally, the two linkage groups (fragments a–c and e–f) are not alleles of the same locus. Individual III 1 is a compound heterozygote carrying both the paternal a–c cluster and the maternal e–f pair; both clusters are transmitted to two of his four children, establishing that they are not segregating as alleles but instead must be derived from two unlinked hypervariable regions.

None of the maternal (II 5) minisatellite fragments showed significant linkage to β-thalassemia trait and are therefore not closely linked to the β-globin gene cluster on chromosome 11. In contrast, one maternal fragment (g) 8.6 kb in length tended to cosegregate with HPFH in the seven offspring and in three informative sibships of grandchildren (fig. 2). Only one recombinant was seen in 13 progeny, suggesting close linkage ($\hat{z} = 2.06$ at $\hat{\theta} = .1$). Even allowing for the fact that 18 loci in II 5 have been investigated, this linkage still appears significant (the probability that at least one of 24 unlinked bands derived from at least 18 different loci would show this degree of linkage in coupling with HPFH *by chance* is .04). We have further checked that fragment g in II 5 is a single minisatellite allele, and not two superimposed segregating DNA fragments, by

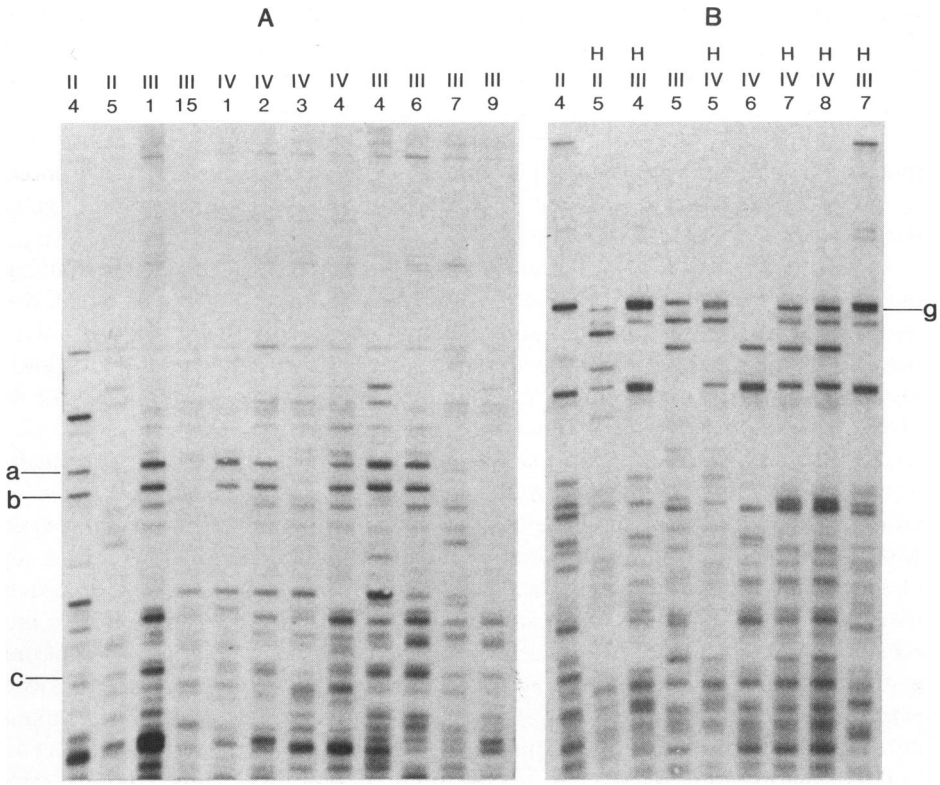

FIG. 3.—Examples of linkage in the DNA fingerprints of a Gujarati pedigree segregating for HPFH. Ten-microgram samples of blood DNA were digested with *Hin*fI, and DNA fingerprints were produced as described in figure 1, using probe 33.6 (*A*) or 33.15 (*B*). Electrophoresis was performed in a relative short (20 cm) agarose gel. The relationship between individuals is given in figure 2. *A*, Hypervariable fragments *a*, *b*, and *c* are closely linked and are either all present or all absent in each individual in the pedigree. *B*, Cosegregation of band *g* and HPFH (*H*). Individuals IV 7 and IV 8 are identical twins and have indistinguishable DNA fingerprints. Detailed marker segregation patterns are given in figure 2.

investigating the DNA fingerprints of all individuals shown in figure 2, digested with *Sau*3A instead of *Hin*fI; every positive fingerprint contained a corresponding *Sau*3A fragment of size similar to that of fragment g (8.2 kb vs. 8.6 kb), as expected for a single minisatellite fragment [12] (data not shown).

DISCUSSION

Human pedigree analysis shows that the DNA fingerprints detected by minisatellite probes can be reliably used for studying the segregation of multiple heterozygous DNA fragments, even in families where one or the other parent is unavailable for study. Using two such probes, it is possible to analyze up to 34 hypervariable loci simultaneously in a single individual, a rate of genetic marker generation that is far higher than that obtained by conventional methods, including RFLPs, in human genetics. The stable inheritance of vari-

able minisatellite fragments [11], together with the low population frequency of individual fragments [12], makes them well suited to linkage analysis, as shown by the examples of linkage discovered in the two pedigrees analyzed. We should stress that, while these hypervariable minisatellites may be recombination hotspots [11], the estimated rates of unequal exchange occurring at long minisatellite loci ($\sim$ 0.001 per locus per gamete) is not sufficient to perturb significantly the linkage between a minisatellite locus and a neighboring gene such as a disease locus.

We estimate that the total number of hypervariable loci detected together by minisatellite probes 33.6 and 33.15 is very approximately 60. At least one of the two alleles of about half of these loci can be resolved in a given DNA fingerprint, and it therefore follows that in DNA fingerprints of different individuals the spectrum of scorable loci examined will not necessarily be identical. Most or all of these loci are genetically unlinked and must therefore be scattered over a substantial proportion of the human genome. Their precise location is not known and must await the cloning and regional localization of individual hypervariable minisatellite loci. Curiously, no minisatellites have yet been found on the X or Y chromosome in either pedigree studied. We estimate that approximately 43 different loci have been scored for possible sex linkage in the father of the neurofibromatosis family together with individual II 4 in the HPFH family. Since the X and Y chromosomes together constitute $\sim$ 5% of the genome of a male, then the probability that none of 43 randomly dispersed loci resides on these chromosomes is $(.95)^{43} = .1$; the apparent lack of sex-linked minisatellites is therefore not significant.

These dispersed hypervariable minisatellite loci are well suited to the search for markers linked to disease loci, as shown by the provisional example of linkage to HPFH. Unlike conventional single-locus genetic analysis, linkage data cannot be pooled between unrelated small pedigrees, since a different minisatellite allele is likely to be associated with the disease locus in each pedigree. Instead, DNA fingerprints are only suitable for studying linkage, particularly of dominant disorders, in an extensive pedigree and most ideally in a single large sibship.

So far, probes 33.6 and 33.15 permit up to 34 autosomal hypervariable loci to be scored in an individual. The chance that at least one of these loci is closely linked to a given disease locus (within 10 cM) is 20%, assuming random dispersal of minisatellites throughout the 3,300-cM-long human linkage map. For extended pedigrees such as the HPFH family, this probability falls to $\sim$ 10% since only one allele of most loci is scorable, and to detect linkage, this allele must be linked *in coupling* to the disease locus. To raise these probabilities above 50% would require the scoring of $>$ 104 hypervariable loci in a single large sibship and $>$ 208 loci in an extended pedigree (see also [22]). These numbers exceed the estimated total number of loci detected by the two minisatellite probes used so far. However, probes 33.6 and 33.15 detect essentially totally different sets of hypervariable loci, which suggests that the total number of human minisatellites that contain various versions of the core sequence may be large.

In conventional human pedigree analysis using defined single-locus markers, evidence of linkage between a marker and a disease locus usually directly gives the approximate genomic location of the disease gene and can be further established by analyzing additional pedigrees. The converse is true for DNA fingerprints, and further analysis of possible linkage between a hypervariable DNA fragment and a disease is possible only via isolation of the fragment by preparative gel electrophoresis and cloning. Locus-specific hybridization probes could then be designed from the isolated minisatellite, either by using unique sequence DNA segments immediately flanking the minisatellite or by using the entire minisatellite in high stringency hybridizations. Such locus-specific probes could be used both to extend the linkage data in additional families and to localize the minisatellite within the human genome. This approach is currently being tested by cloning the 8.2-kb *Sau*3A minisatellite fragment apparently linked to the HPFH locus in the Gujarati pedigree and the 2.0-kb *Hin*fI fragment that tends to cosegregate with neurofibromatosis.

ACKNOWLEDGMENTS

We are grateful to John F. Y. Brookfield, Robert Semeonoff, and Brian English for helpful discussions on statistics and to Jim Wainscoat for discussions that led to initiating the HPFH linkage analysis project. We also thank Margaret Ponder for tracing the NF family and the family themselves and "LINK" (the NF patients' association) for their help.

REFERENCES

1. WEATHERALL DJ: *The New Genetics and Clinical Practice*. London, Nuffield Provincial Hospitals Trust, 1982
2. CAVANEE WK, DRYJA TP, PHILLIPS RA, ET AL.: Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* 305:779–784, 1983
3. GUSELLA JF, WEXLER NS, CONNEALLY PM, ET AL.: A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306:234–238, 1983
4. WHITE R, LEPPERT M, BISHOP DT, ET AL.: Construction of linkage maps with DNA markers for human chromosomes. *Nature* 313:101–105, 1985
5. JEFFREYS AJ: DNA sequence variants in the $^{G}\gamma$-, $^{A}\gamma$-, δ- and β-globin genes of man. *Cell* 18:1–10, 1979
6. COOPER DN, SCHMIDTKE J: DNA restriction fragment length polymorphisms and heterozygosity in the human genome. *Hum Genet* 66:1–16, 1984
7. BELL GI, SELBY MJ, RUTTER WJ: The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 295:31–35, 1982
8. PROUDFOOT NJ, GIL A, MANIATIS T: The structure of the human zeta-globin gene and a closely linked, nearly identical, pseudogene. *Cell* 31:553–563, 1982
9. CAPON DJ, CHEN EY, LEVINSON AD, SEEBURG PH, GOEDDEL DV: Complete nucleotide sequence of the T24 human bladder carcinoma oncogene and its normal homologue. *Nature* 302:33–37, 1983
10. GOODBOURN SEY, HIGGS DR, CLEGG JB, WEATHERALL DJ: Molecular basis of length polymorphism in the human ζ-globin gene complex. *Proc Natl Acad Sci USA* 80:5022–5026, 1983
11. JEFFREYS AJ, WILSON V, THEIN SL: Hypervariable "minisatellite" regions in human DNA. *Nature* 314:67–73, 1985
12. JEFFREYS AJ, WILSON V, THEIN SL: Individual-specific 'fingerprints' of human DNA. *Nature* 316:76–79, 1985

13. SOUTHERN EM: Gel electrophoresis of restriction fragments. *Methods Enzymol* 68:152–176, 1980
14. SMITH CAB: Linkage scores and corrections in simple two- and three-generation families. *Ann Hum Genet* 32:127–150, 1968
15. EMERY AEH: *Methodology in Medical Genetics*. Edinburgh, Churchill Livingstone, 1976
16. HOPE DG, MULVIHILL JJ: Malignancy in neurofibromatosis. *Adv Neurol* 29:33–56, 1981
17. RICCARDI VM: Von Recklinghausen neurofibromatosis. *N Engl J Med* 305:1617–1627, 1981
18. SPENCE MA, BADER JL, PARRY DM, ET AL.: Linkage analysis of neurofibromatosis (von Recklinghausen disease). *J Med Genet* 20:334–337, 1983
19. HALDANE JBS: The probable errors of calculated linkage values and the most accurate method of determining gametic from certain zygotic series. *J Genet* 8:291–297, 1919
20. RENWICK JH: Progress in mapping human autosomes. *Br Med Bull* 25:65–73, 1969
21. GIANNI AM, BREGNI M, CAPPELLINI MD, ET AL.: A gene controlling foetal haemoglobin expression in adults is not linked to the non-α-globin cluster. *EMBO J* 2:921–925, 1983
22. BOTSTEIN D, WHITE RL, SKOLNICK M, DAVIS RW: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331, 1980