Proc. Natl. Acad. Sci. USA
Vol. 77, No. 11, pp. 6754–6758, November 1980
Genetics

A highly polymorphic locus in human DNA

(human genetics/recombinant DNA/genetic marker loci)

ARLENE R. WYMAN* AND RAY WHITE†

Department of Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01605

Communicated by Franklin W. Stahl, August 19, 1980

ABSTRACT **A locus in the human genome, not associated with any specific gene, has been found to be a site of restriction fragment length polymorphism. The polymorphism was found by hybridizing a 16-kilobase-pair segment of single-copy human DNA, selected from the human genome library cloned in phage λCH4A, to a Southern transfer of total human DNA digested with *Eco*RI. DNAs from a number of individuals from within Mormon pedigrees as well as random individuals have been examined. The locus is highly variable, with at least eight alleles present, homozygotes accounting for less than 25% of the individuals examined. The polymorphism appears to be the result of DNA rearrangements rather than base-pair substitutions or modifications. Examination of the DNA from seven members of a family revealed fragment lengths that are consistent with their inheritance as Mendelian alleles through three generations.**

Because of the scarcity of useful genetic marker loci, the human genetic linkage map has relatively few entries. For a marker locus to be useful in linkage studies, it must be the site of several variants (alleles). With several alleles occurring at reasonable frequencies in the population, parental individuals heterozygous at different marker loci will be common and cosegregation of markers can be scored among their progeny.

Almost all polymorphic marker loci thus far identified in humans represent biochemical variations, such as isozymes or cell surface markers. Most of these loci do not have a sufficient number and frequency of alleles to be of general use in linkage studies in families, and many are not appropriate for prenatal diagnosis. A few biochemical and immunological loci having reasonably frequent variants within the human population have been described, however, and shown to be useful for linkage studies and for the resolution of genetic models (1).

We have argued (2) that a large number of DNA sequence variations exist in the human population. Some of these should be detectable as variants in the lengths of the DNA fragments produced by a restriction enzyme (restriction fragment length polymorphism or RFLP). These loci will define arbitrary genetic sites, not necessarily associated with any specific genes, and should prove highly suitable as genetic markers. A linkage map of the human can be constructed by determining the pattern of cosegregation of these genetic markers in human pedigrees. A map based on several hundred such loci will have a sufficient marker density to allow the placement of any other locus with reasonable certainty (2, 3). RFLP markers specifically associated with the human β- and γ-globin loci have already been reported (4, 5). We will describe here an *arbitrary* locus showing a significant frequency of DNA sequence variation.

In these experiments, the restriction fragments from the DNA of a locus are revealed by hybridization with a radiolabeled recombinant DNA probe. The search for polymorphisms has been facilitated by choosing as probes recombinant DNA segments that have two characteristics. First, it is important that the segments be entirely single copy because repetitious sequences obscure the fragments representing the locus and reveal a large and bewildering number of unlinked sequences of only partial homology. Second, the use of long probes permits long segments of DNA and many restriction sites to be examined in a single experiment.

Unfortunately, most long segments of human DNA contain repetitive sequences because the DNA organization of the human genome follows the general pattern of interspersed unique and repeated elements found in most eukaryotes (6). However, these studies do suggest that some sequences of length more than 13 kilobase pairs (kb) are completely unique. Therefore, our approach has been to screen with radiolabeled human repetitious DNA a library consisting of human genome segments of length 15–20 kb carried in the λ phage vector Charon 4A (7) in order to detect clones that are free of repetitive human DNA. We have identified a number of such clones and have tested five of them as probes against restriction endonuclease-digested DNA of individual humans. The polymorphism revealed by one of these, λCH4A-rHs18, has been examined in detail and is the subject of this report.

MATERIALS AND METHODS

Preparation of Human DNA. Human DNA was prepared from peripheral white blood cells by lysis in 0.1% $NaDodSO_4$ and proteinase K (8). Typically, 20 ml of heparinized blood was gently mixed with 20 ml of 1.8% NaCl and 20 ml of 6% (wt/vol) Dextran 70 (Macrodex, Pharmacia) in 1.8% NaCl and the erythrocytes were allowed to settle. The supernatant was centrifuged at low speed at 4°C and the pellet was resuspended in 7 ml of H_2O for 45 sec to lyse residual erythrocytes. Seven milliliters of 1.8% NaCl was then added, and the remaining cells were pelleted and resuspended in 10 ml of lysis buffer (0.1 M NaCl/10 mM Tris·HCl, pH 7.8/1 mM EDTA). The cells were lysed by dropwise addition into 10 ml of lysis buffer containing 1% $NaDodSO_4$ and 0.4 mg of proteinase K per ml (Boehringer Mannheim) at room temperature. Shaking of this mixture for 1–2 hr at room temperature was followed by two extractions with equal volumes of phenol and two extractions with $CHCl_3$/isoamyl alcohol, 24:1 (vol/vol). DNA, in the aqueous phase, was precipitated by addition of 2.5 vol of isopropanol and mixed by pouring back and forth in beakers. The fibrous DNA precipitate, typically 200 μg, was mechanically removed

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: kb, kilobase pairs; RFLP, restriction fragment length polymorphism.

* Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.
† Present address: Department of Cellular, Viral and Molecular Biology, University of Utah Medical School, Salt Lake City, UT 84132.

and redissolved in 1–2 ml of 10 mM Tris·HCl, pH 7.8/1 mM EDTA. The DNA was further purified by equilibrium sedimentation in CsCl/ethidium bromide gradients as follows: 1 ml of the DNA solution was added to 12 ml of 50 mM Tris·HCl (pH 8.0), 0.2 M NaCl, 1 mM EDTA, 6% (wt/vol) sucrose, 0.3 mg of ethidium bromide per ml, and 0.87 g of CsCl per ml, and the mixture was centrifuged for 48–65 hr at 40,000 rpm in a 50Ti rotor. The fluorescent DNA bands were removed from the gradients by puncture with a 16-gauge needle under long-wavelength ultraviolet illumination. Ethidium bromide was removed by successive washes with isopropanol, and CsCl was removed by dialysis against 10 mM Tris·HCl, pH 7.8/0.1 mM EDTA.

Digestion, Electrophoresis, Transfer, and Hybridization of DNA. Restriction enzymes were obtained from Bethesda Research Laboratories (Rockville, MD), Boehringer Mannheim, and New England BioLabs and used as directed. Restriction digestions of human DNA were carried out at a DNA concentration of 100 μg/ml. Human DNA was monitored for complete digestion by removal of a volume containing 1 μg of human DNA (immediately after enzyme addition) into 1 μg of λ phage DNA (New England BioLabs) for parallel digestion. Digestion of the human DNA was judged to be complete if the λ DNA was fully digested. Electrophoresis in 0.4% or 0.5% agarose (FMC, Rockland, ME) gels was carried out according to Hayward and Smith (9). Nick-translations (10) were carried out for 90 min at 15°C in a final volume of 10 μl, containing 15 μM each [^{32}P]dATP, [^{32}P]dCTP, and [^{32}P]dTTP (approximately 400 Ci/mmol, 1 Ci = 3.7 × 10^{10} becquerels; Amersham and New England Nuclear), 30 mM dGTP (P-L Biochemicals), 50 mM Tris·HCl (pH 7.4), 5 mM $MgCl_2$, 20 mM 2-mercaptoethanol, 2 μg of DNase I (Worthington; stored and activated according to ref. 10), and 1 μl of DNase-free *Escherichia coli* DNA polymerase I (a gift from Ronald Davis). Human DNA was transferred from gel to diazobenzyloxymethyl-paper and hybridized with nick-translated probes by the methods of Alwine *et al.* (11) and Wahl *et al.* (12), with a final wash of the paper in 15 mM NaCl/1.5 mM trisodium citrate/0.1% Na$DodSO_4$, pH 7.4, for 30 min at 51–53°C. Autoradiography was for 16 hr to 10 days at −80°C with Kodak XR-5 film, backed by a Lightning-Plus intensifying screen (DuPont).

Subcloning. Phage λCH4A-rHs18 and plasmid vector pBR322 were digested to completion with *Eco*RI. The linearized pBR322 was treated with bacterial alkaline phosphatase (Worthington) to prevent recyclization (13); 2.5 μg of *Eco*RI fragments of the phage and 0.25 μg of vector DNA were ligated by using 5 units of T4 ligase (Bethesda Research Laboratories, as directed except that 2 mM ATP was used) in 50 μl. Incubation was at 15°C for 2 hr; then the mixture was diluted 1:10 and incubation was continued for 16 hr. The ligation products were used directly to transfect *E. coli* strain HB101 (14). The recombinant plasmid that contained a 5-kb insert in pBR322, corresponding to fragment D (Fig. 1), was designated pAW101. In accordance with the Guidelines for Recombinant DNA Research of the National Institutes of Health, λCH4A-rHs18 was grown in DP50supF, an EK2 host–vector system, and pAW101 was grown in HB101, an EK1 host–vector system, both under P2 containment.

RESULTS

Detection of RFLP in Human DNA. Variability in the lengths of *Eco*RI fragments at an arbitrary locus in the DNA from individual humans was examined by digesting the DNA with restriction endonucleases and transferring it to diazobenzyloxymethyl-paper (12, 15). Recombinant phage λCH4A-rHs18, containing a 16-kb single-copy human DNA segment bounded by synthetic *Eco*RI sites, was used as probe. Fig. 1 shows the restriction site map of the human insert of the phage for the restriction enzymes *Eco*RI and *Hin*dIII. The result with human DNA was consistent with *Eco*RI restriction fragment D from the phage having homology with a polymorphic (and much longer) *Eco*RI fragment in the DNA from individual humans and fragments A, B, and C having homology with nonpolymorphic fragments. Restriction fragment D was subcloned in the plasmid vector pBR322 and, when used as a probe, revealed the polymorphic restriction fragment pattern shown in Fig. 2. Our interpretation of these data is that individuals with two bands are heterozygous for two fragment lengths (alleles) and individuals with only one band are homozygous for a single allele or contain two alleles that are not yet resolved.

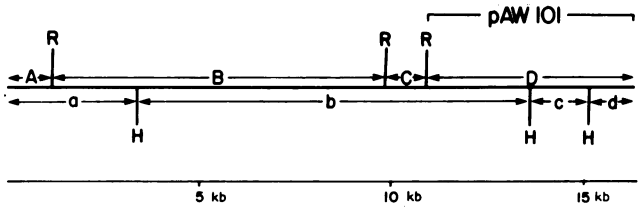

FIG. 1. Restriction site map of the human insert of phage λCH4A-rHs18. A–D are *Eco*RI fragments; a–d are *Hin*dIII fragments. The subclone of fragment D in the plasmid vector pBR322, used in these studies, is designated pAW101. Sites for *Eco*RI and *Hin*dIII were determined by combined cleavage with the two enzymes, and fragment lengths were determined by comparison with fragments of λ DNA produced by cleavage with *Hin*dIII or *Sal* I. The map was confirmed by two-dimensional electrophoresis blotting hybridization (16) of phage λCH4A-rHs18, with unlabeled *Eco*RI fragments and ^{32}P-labeled *Hin*dIII fragments (data not shown).

Number and Frequency of Alleles. By comparing bands in the same or adjacent lanes (Fig. 2), at least eight different fragment lengths could be clearly discriminated. The size range of fragment lengths was from 14 to 29 kb. Three alleles in the molecular size range of 14–17 kb are seen in Fig. 2 among individuals 4621, 4625, 4609, and 4629. Four alleles in the molecular size range of 19–22 kb are seen among individuals 4454, 4202, 4451, 4571, and 4610. An additional 29-kb *Eco*RI fragment is seen in individuals 4616 and 4202. Table 1 gives the lengths of the *Eco*RI fragments from 43 individuals. Because

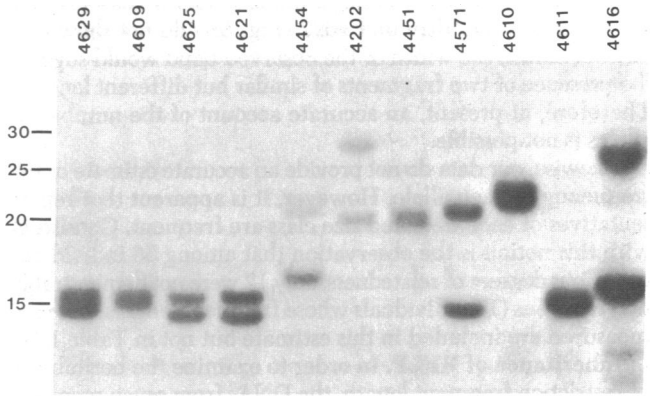

FIG. 2. Polymorphisms in human DNA restriction fragments homologous to pAW101. Five micrograms of DNA from each individual was digested with *Eco*RI, subjected to electrophoresis through a 0.4% agarose gel, transferred to diazobenzyloxymethyl-paper, and hybridized with a nick-translated probe prepared from the recombinant plasmid pAW101. Numbers above each lane indicate the individual within a large Mormon pedigree (17); numbers to the left indicate length of DNA in kb.

Table 1. Lengths (in kb) of *Eco*RI fragments homologous to pAW101

Individual	Fragment 1	Fragment 2	Individual	Fragment 1	Fragment 2
4142	14.8	20.2	4617	15.7	
4202	20.0	28.8	4618	14.8	16.0
4214	22.2	28.2	4620	15.2	22.4
4218	14.8		4621	14.6	15.7
4224	14.9	20.2	4622	15.1	
4225	14.8	20.2	4623	16.1	22.4
4229	14.9	19.7	4624	16.1	22.5
4230	14.7	28.5	4625	14.6	15.7
4451	20.1		4626	21.7	
4452	14.8	20.5	4627	14.3	14.9
4453	14.9		4628	15.0	20.8
4454	16.3	20.8	4629	14.9	15.7
4516	20.1	27.6	4873	16.1	
4571	14.7	20.9	4874	14.6	22.4
4606	16.0	22.8	4875	14.7	18.5
4608	15.6		4876	14.8	22.5
4609	15.3	15.6	H2	15.0	15.5
4610	21.9	23.4	H3	14.4	15.1
4611	15.1		H4	14.4	15.2
4612	16.3	19.7	902	15.7	19.0
4613	20.9	28.2	903	14.8	23.0
4616	15.9	28.0			

DNAs from 43 individuals were digested, transferred, and hybridized as described in the legend to Fig. 2. Fragments 1 and 2 correspond to shorter and longer fragments, respectively. Their lengths were determined by using λCH4A digested with *Eco*RI and λCH4A digested with *Eco*RI and *Bam*HI as molecular size markers. The mobilities of these markers had been revealed by prior hybridization with ^{32}P-labeled phage λCH4A-rHs18. Lengths reported are the averages of two independent measurements on the same autoradiogram, except for individuals 902 and 903, where they are the averages of two independent measurements on each of three autoradiograms. Lengths for individual 4620 were determined by comparison with other bands in Fig. 3*B*.

some alleles differed by only a few hundred base pairs, even with an agarose gel system optimized for resolution of these long fragments it was not always possible to determine whether two fragments had the same or different lengths unless they were run in adjacent lanes. Furthermore, with improved resolution, some individuals who appeared to be homozygous might in fact be heterozygous. For example, although individuals 4451 and 4609 (Fig. 2) and 4626 and 4622 (Fig. 3*B*) do not show two distinct bands, the width of the observed band would suggest the presence of two fragments of similar but different lengths. Therefore, at present, an accurate account of the number of alleles is not possible.

Likewise, our data do not provide an accurate estimate of the frequency of each allele. However, it is apparent that representatives of each detected size class are frequent. Consistent with this notion is the observation that among 56 individuals of various degrees of relatedness, only 12 were not demonstrably heterozygous (13 individuals whose fragment lengths were not measured are included in this estimate but not in Table 1).

Inheritance of RFLP. In order to examine the heritability of restriction fragment length, the DNAs from seven members of a family spanning three generations were examined. A pedigree (17) showing the familial relationships is shown in Fig. 3*A* and the DNA restriction fragment patterns in Fig. 3*B*. In every case the fragment lengths are consistent with their inheritance as Mendelian alleles through three generations.

Specifically, assuming that individuals 4629, 4626, and 4622 are actually incompletely resolved heterozygotes for fragments

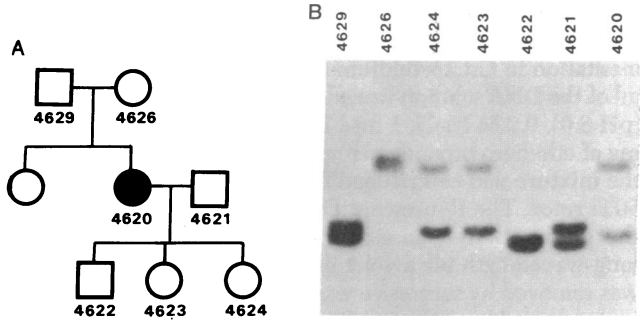

FIG. 3. (*A*) Pedigree (17) showing familial relationships of seven individuals. (*B*) Restriction fragment patterns of seven individuals. DNA was digested, transferred, and hybridized as described in the legend to Fig. 2.

of similar lengths, the fragment pattern of individual 4620 can be accounted for by the inheritance of the lower molecular size fragment from 4626 and the lower molecular size fragment from 4629. The fragment pattern of individual 4622 is consistent with the inheritance of the lower molecular size fragments from both 4620 and 4621. The fragment pattern of 4623 and 4624 indicate inheritance of the higher molecular size fragments of 4620 and 4621.

Alternative models of random distribution of fragment lengths among individuals or of mitochondrial (maternal) inheritance can be ruled out. Because each member of the family appears to be heterozygous, each progeny individual is potentially useful for detecting linkage.

Basis for the DNA Polymorphism. RFLP can result from the creation or elimination of a restriction site by base-pair substitution or modification or from the movement of sites by rearrangements of DNA segments (deletions, insertions, or inversions). The large number of alleles at this locus suggested to us that rearrangements involving DNA segments of various length were likely responsible for the *Eco*RI fragment polymorphism. This was tested by probing the human DNA after digestion with a different enzyme, *Bam*HI. Polymorphisms due to rearrangements could be revealed in *Bam*HI-digested DNA, whereas polymorphisms due to base-pair differences at *Eco*RI sites would not. Table 2 shows the fragment lengths for DNA from individuals 902 and 903 digested with either *Eco*RI or *Bam*HI. Digestion with either enzyme gives two fragments in both cases. DNA from individual 902 shows a 3.3-kb difference in length both between its *Eco*RI fragments and between its *Bam*HI fragments. DNA from individual 903 shows an 8.2-kb difference in length between the *Eco*RI fragments and between the *Bam*HI fragments. Because the differences in length between the *Eco*RI fragment pairs and *Bam*HI fragment pairs are the same for both enzymes, these data are consistent with

Table 2. Lengths (in kb) of *Eco*RI and *Bam*HI fragments homologous to pAW101

Individual	Enzyme	Fragment 1	Fragment 2	Difference (2 − 1)
902	*Eco*RI	15.7	19.0	3.3
	*Bam*HI	17.8	21.1	3.3
903	*Eco*RI	14.8	23.0	8.2
	*Bam*HI	16.9	25.1	8.2

DNAs from individuals 902 and 903 were digested with *Eco*RI or with *Bam*HI and transferred and hybridized as described in the legend to Fig. 2. Lengths of *Eco*RI fragments are from Table 1. Lengths of *Bam*HI fragments are the averages of two independent measurements on the same autoradiogram.

the notion that the variability in restriction fragment lengths at this locus is due to rearrangements involving DNA segments of variable lengths.

A Second Polymorphism at this Locus. Human DNA digested with *Hin*dIII, transferred to diazobenzyloxymethyl-paper, and probed with DNA from λCH4A-rHs18 revealed three invariant fragments of lengths 1.5, 4.2, and 5.5 kb and one fragment of variable length, either 10.3 or 11.0 kb. These results suggest that the *Hin*dIII restriction fragment c from the phage (Fig. 1) has homology with the 1.5-kb human fragment, that fragments a and d have homology with the 4.2- and 5.5-kb human fragments, and that fragment b has homology with the variable fragment. Table 3 gives the lengths of the variable fragments for 47 individuals. Fragments observed for individuals in the small pedigree (Fig. 3*A*) are consistent with their inheritance as Mendelian alleles. Within the population examined, the 10.3-kb allele is more common than the 11.0-kb allele. Of 47 individuals, 31 are homozygous for the 10.3-kb allele, 3 are homozygous for the 11.0-kb allele, and 13 are heterozygous.

Comparison of the variable fragments produced with either *Eco*RI (Table 1) or *Hin*dIII (Table 3) indicates that the *Hin*dIII RFLP is separable from the *Eco*RI RFLP. Individuals heterozygous for *Eco*RI fragments (for example, individual 903) can be homozygous for the *Hin*dIII fragments. Likewise, an individual, 4608, can be apparently homozygous for *Eco*RI fragments but heterozygous for *Hin*dIII fragments. The precise location and structural basis for the *Hin*dIII polymorphism are under investigation. The *Hin*dIII polymorphism increases the usefulness of this locus in linkage studies.

Table 3. *Hin*dIII fragments homologous to λCH4A-rHs18

	Fragment			Fragment	
Individual	10.3-kb	11.0-kb	Individual	10.3-kb	11.0-kb
4126	+		4613	+	
4142	+		4615	+	
4202	+		4616	+	+
4214	+	+	4617	+	
4218	+		4618	+	
4224	+		4620	+	
4225	+		4621	+	+
4227	+	+	4622	+	
4228		+	4623	+	+
4229	+		4624	+	+
4230	+	+	4625	+	
4334		+	4626	+	
4451	+		4627	+	
4452	+		4628	+	
4453	+		4629	+	
4454	+	+	R1	+	
4516	+		R3	+	
4571	+	+	R4		+
4606	+	+	R6	+	
4608	+	+	R7	+	
4609	+		R9	+	
4610	+		902	+	+
4611	+		903	+	
4612	+	+			

DNAs from 47 individuals were digested with *Hin*dIII, transferred, and hybridized as described in the legend to Fig. 2, except the nick-translated probes were prepared either from λCH4A-rHs18 or from pAW101. In addition to the variable fragments, probe from λCH4A-rHs18 revealed the nonpolymorphic fragments 1.5, 4.2, and 5.5 kb, and probe from pAW101 revealed nonpolymorphic fragments 1.5 and 5.5 kb, described in the text.

DISCUSSION

The cloned human DNA segment found in recombinant phage λCH4A-rHs18 defines a locus in human DNA exhibiting a high degree of DNA sequence polymorphism. We have shown that there exist at least eight and possibly more variants at this locus, which are revealed as a series of allelic *Eco*RI fragments ranging in size from 14 to 29 kb as well as two alleles revealed as variant *Hin*dIII fragments. This yields at least 16 distinct allelic combinations per chromosome for the marker system defined by these enzymes.

This level of polymorphism will permit effective application of this genetic marker system to family linkage studies. The observation that among 56 individuals only 12 could not clearly be shown to be heterozygous for *Eco*RI fragments supports this notion. Moreover, at least one of these individuals was seen to be heterozygous for *Hin*dIII fragments. Further resolution of the alleles at this locus can best be achieved by reducing the size of the restriction fragments containing the polymorphic region because the smallest difference we can resolve with complete confidence between restriction fragments in the 15- to 20-kb size range is perhaps 1 kb. If we can reduce the size of the smallest fragment containing the polymorphic region to 2 or 3 kb by cutting with additional restriction enzymes, we should be able to resolve with confidence differences of 0.1 kb between allelic fragments.

However, the resolution of fragments run in adjacent lanes is much better than 1 kb. In fact, length differences of as little as 0.3 kb are readily discriminated, as seen in Fig. 2. The importance of this approach for the use of the locus in its present stage of development was illustrated in Fig. 2, showing a small pedigree which is almost completely informative even though some of the allelic *Eco*RI fragments are similar in length. Often the need of the geneticist is to discriminate among several alleles with respect to their inheritance within a small family group. In this instance the DNA of parents and children would all be run on the same gel, permitting the high level of resolution demonstrated in Fig. 2.

The observation that the DNA structural basis for the polymorphism is due to a rearrangement series within the allelic restriction fragments suggests, by analogy with sites containing transposable elements found in *E. coli* (18), yeast (19), and *Drosophila* (20), that an active DNA element may reside at this locus. If so, and if that element can be cloned, we may be able to develop a hybridization probe that will detect other loci at which the active element resides, permitting the ready identification of other highly polymorphic loci.

The procedures outlined here for testing phages that lack repetitious sequences will eventually produce the large number of marker loci required for a human linkage map. Furthermore, new protocols should make the development of additional probe/restriction enzyme pairs that reveal polymorphism significantly more efficient. As markers become available, their linkage relationships will be determined by following segregation in a large Mormon pedigree. The development of a linkage map could permit antenatal diagnosis by linked markers for most suspected genetic diseases by testing of fetal cells from amniotic fluid samples (4, 21). As markers linked to disease loci are defined, the resolution of the genetic models will become straightforward (1) because the genotypes will be known even for cases of incomplete penetrance. The development of markers for disease loci whose chromosomal location is known will be aided by this construction of recombinant DNA probes specific for known chromosomal segments (22).

Finally, although genetic research has historically required the construction of strains of specific genotype for the testing of specific genetic hypotheses, detailed marker data from large

human pedigrees should often permit the identification of individuals of the required test genotype as the result of natural matings. This should make the human substantially more tractable as a system for the study of genetics.

We gratefully acknowledge Mark Skolnick, David Botstein, and Ron Davis for providing discussion, good advice, and ample enthusiasm. The generous efforts of Tim Helentjaris and Rob Hadley provided the Mormon DNAs for these experiments. We thank Geoff Wahl for valuable assistance with diazobenzyloxymethyl-paper and Tom Maniatis for the human library. Jon Berkowitz, Mark Weber, Annelie James, and Alena Raidl provided expert technical assistance. Bill Petri offered many useful suggestions on the manuscript. This work was supported by the Wills Foundation, the American Cancer Society (NP-193B), and the National Institutes of Health (GM-27313).

1. Kravitz, K., Skolnick, M., Cannings, C., Carmelli, D., Baty, B., Amos, B., Johnson, A., Mendell, N., Edwards, C. & Cartwright, G. (1979) *Am. J. Hum. Genet.* **31,** 601–619.
2. Botstein, D., White, R. L. Skolnick, M. & Davis, R. W. (1980) *Am. J. Hum. Genet.* **32,** 314–331.
3. Bishop, T. & Skolnick, M. H. (1980) in *Cancer Incidence in Defined Populations* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Banbury Report No. 4, pp. 421–426.
4. Kan, Y. W. & Dozy, A. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5631–5635.
5. Jeffreys, A. J. (1979) *Cell* **18,** 1–10.
6. Schmidt, C. W. & Deininger, P. L. (1975) *Cell* **6,** 345–358.
7. Lawn, R. W., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15,** 1157–1174.
8. Steffen, D. & Weinberg, R. A. (1978) *Cell* **15,** 1003–1010.
9. Hayward, G. S. & Smith, M. G. (1972) *J. Mol. Biol.* **63,** 383–395.
10. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–254.
11. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5350–5354.
12. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
13. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. & Goodman, H. M. (1977) *Science* **196,** 1313–1319.
14. Beckingham, K. & White, R. (1980) *J. Mol. Biol.* **137,** 349–373.
15. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
16. Sato, S., Hutchinson, C. A. & Harris, J. I. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 542–546.
17. Skolnick, M. H. (1980) in *Cancer Incidence in Defined Populations* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Banbury Report No. 4, pp. 285–296.
18. Kleckner, N. (1977) *Cell* **11,** 11–23.
19. Cameron, J. R., Loh, E. Y. & Davis, R. W. (1979) *Cell* **16,** 739–751.
20. Potter, S. S., Brorein, W. J., Dunsmuir, P. & Rubin, G. M. (1979) *Cell* **17,** 415–427.
21. Orkin, S., Alter, B., Altay, C., Mahoney, M., Lazarus, H., Hobbins, J. & Nathan, D. (1978) *N. Engl. J. Med.* **299,** 166–172.
22. Gusella, J., Keys, C., Varsanyi-Breiner, A., Kao, F., Jones, C., Puck, T. T. & Housman, D. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 2829–2833.