*Am J Hum Genet* 36:1159–1171, 1984

# A Locus on Chromosome 11p with Multiple Restriction Site Polymorphisms

DAVID BARKER,[1,2] TOM HOLM,[1] AND RAY WHITE[1]

SUMMARY

We have discovered and characterized a new polymorphic locus on chromosome 11p, *D11S12*, defined by an arbitrary genomic DNA segment cloned in the plasmid pADJ762. Four different polymorphic restriction sites with minor allele frequencies greater than 5% are revealed by Southern hybridization of this probe and its derivatives to digests of human DNAs. These include two *Msp*I sites, a *Taq*I site, and a *Bcl*I site. The frequencies of the common haplotypes at this locus have been determined in a Utah population. Significant linkage disequilibrium has been demonstrated to exist between some pairs of polymorphic sites. A molecular map of this region has been determined, and the polymorphic sites have been localized. Comparison of physical separation with degree of linkage disequilibrium reveals an interesting case where an *Msp*I site and a *Taq*I site that are separated by 6.8 kilobases (kb) show a greater degree of disequilibrium with each other than they do with two polymorphic sites located between them. One of the two interior sites is a *Bcl*I site that is approximately 0.2 kb away from the *Taq*I site but shows the same degree of disequilibrium with the *Taq*I site as with the *Msp*I site 6.7 kb away. Although there is significant linkage disequilibrium at this locus, there are four major haplotypes with frequencies of 5% or greater, and the polymorphic information content (PIC) of this locus is .64.

## INTRODUCTION

The development of a reliable and useful human linkage map depends on th identification of about 300 highly polymorphic genetic markers evenly space throughout the genome [1]. A significant number of DNA probes revealing re

Received January 13, 1984; revised July 25, 1984.

[1] Howard Hughes Medical Institute, University of Utah Medical School, Salt Lake City, UT 84132.
[2] Present address: Collaborative Research Inc., 128 Spring Street, Lexington, MA 02173.

striction fragment length polymorphisms have been identified [2] since this approach was proposed [3] for generating a comprehensive set of genetic markers. The most common form of restriction length polymorphism is caused by base-pair variation resulting in presence or absence of a recognition site for a particular enzyme. We have demonstrated that variation at restriction sites containing the CpG dinucleotide is significantly greater than at other sites [4]. Variation resulting in a series of alleles of different lengths such as has been demonstrated at the insulin [5], zeta-globin [6–8], Ha-ras-1 [9, 10], and immunoglobulin lambda light-chain [11] loci and one arbitrary locus [12] appears to be relatively uncommon. This result is unfortunate from the standpoint of development of genetic markers for linkage studies since the maximum polymorphic information content (PIC) as defined in [3] is 0.375 for a two-allele system, while the range of PICs for the known multiple-allele length polymorphisms mentioned above is from 0.36 to 0.85.

In principle, small DNA regions including several polymorphic sites may serve as highly polymorphic loci if the alleles are not in complete linkage disequilibrium [13]. In this case, the sites define a series of haplotypic combinations that may be considered as multiple alleles, since recombination over distances of the order of 10 kb is expected to be negligible. An excellent example of such a system in human DNA has been provided by the study of polymorphic restriction sites in the vicinity of the beta-globin gene cluster [14–16]. At that locus, although the number of different haplotypes observed in several different human populations was significantly smaller than expected, reflecting substantial linkage disequilibrium within two clusters of adjacent polymorphic sites, a much smaller degree of linkage disequilibrium was observed between the two clusters [14]. The result, in Mediterranean populations, is three major haplotypes with frequencies of .05 or greater. Another example is the immunoglobulin mu chain switch-like region that has been shown to contain five different polymorphic *Sst*I sites, defining at least 25 different haplotypes [17]. Very little linkage disequilibrium was observed at this locus; however, the relative physical location of the polymorphic sites is currently unknown. In our present study, a locus with four dimorphic restriction sites has been characterized with respect to the relation between the physical and population genetic association of the sites. We demonstrate that the degree of linkage disequilibrium is not always a monotonic function of physical separation. Our results support the view that loci with multiple restriction site polymorphisms can be valuable genetic markers since polymorphic sites that are physically very close can show a number of haplotypes.

## MATERIALS AND METHODS

### *Southern Transfers and Hybridization*

At early stages of this work, Southern transfers [18] of human DNAs were performed as described in [19] and [20] with DBM paper. Subsequently, we have used Genescreen™ (NEN), Genatran (D&L Filter, Woburn, Mass.), or 0.45 μm, pharmaceutical grade Zeta-pore™ from AMF-CUNO (Meriden, Conn.) to produce reusable Southern blots, employing a modification of the protocol recommended by NEN for Genescreen™. Agarose gels containing 2.5–5 μg of restricted human DNA per lane are washed for 30 min in 0.2 N

NaOH, 0.6 M NaCl to denature the DNA. The pH is neutralized by subsequent washing in 0.5 M Tris, pH 7.5, and 1.5 M NaCl for 30 min. The gel is then placed on a sponge immersed in a reservoir of 25 mM sodium phosphate buffer, pH 6.5, with two layers of Whatman 3MM paper between the surface of the sponge and the gel. A piece of membrane cut to the size of the gel and prewet in the phosphate buffer is placed on top of the gel, and this is covered with two pieces of Whatman 3MM paper followed by a stack of absorbent pads. After at least 12 hrs of transfer, the membrane is removed, washed briefly in 2 × SSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate), and baked at 80°C for at least 1 hr. The membrane is then washed at 65°C in 0.1 × SSC, 0.5% sodium dodecyl sulfate (SDS) for 1 hr. This washing step is important for the elimination of nonspecific binding of radioactive DNA in the subsequent hybridization. The washed filter is subsequently prehybridized for 2 or more hrs in 5 × SSC, 10 × Denhardt's solution (Denhardt's solution is 0.02% bovine serum albumen, 0.02% Ficoll 400, and 0.02% polyvinylpyrollidone 360), 50 mM sodium phosphate buffer, pH 6.7, 500 μg/ml sonicated denatured salmon DNA (Sigma, St. Louis, Mo.), 5% dextran sulfate, and 50% formamide at 42°C. Probes were labeled by nick-translation as described [4] and hybridized with the filters in 5 × SSC, 1 × Denhardt's, 0.02 M sodium phosphate, pH 6.7, 100 μg/ml sonicated denatured salmon DNA, 10% dextran sulfate, and 50% formamide for 20–30 hrs at 42°C. The filters were washed twice at a maximum stringency of 65°C in 0.1 × SSC, 0.1% SDS for 30 min, blotted dry, and exposed to Kodak X-AR5 film backed by a Dupont Cronex Lightning-Plus Intensifier. Following exposure, radioactive probe bound to the filter was removed by washing in 0.4 N NaOH at 45°C for 30 min, followed by neutralization in 0.1 × SSC, 0.5% SDS, 0.2 M Tris, pH 7.5, at 45°C for 30 min and then prehybridized before reuse.

*Subcloning*

The 5.5-kb and 3.3-kb *Eco*RI fragments of phage JA76 [21] were isolated from low-melting-point agarose and ligated with *Eco*RI-digested and phosphatased pBR325 [22] DNA to produce the subclones pADJ762 and pADJ765, respectively. These segments were further dissected to obtain probes appropriate for the mapping of polymorphic sites. The plasmids 762PP3 and 762BB3 were produced by complete digestion of pADJ762 with *Pst*I or *Bam*HI, respectively, followed by religation. A 550-bp *Pst*I fragment of pADJ762 was ligated with *Pst*I-cut pBR322 to form plasmid 762PSC5-24. This fragment is distinguished from a second *Pst*I fragment from pADJ762 of similar length in that it contains both an *Msp*I and a *Taq*I site as predicted by the partial digest mapping results described below. Probe 765BB5 was constructed by double digestion of pADJ765 with *Bam*HI plus *Bgl*II followed by religation. Probes 765PR1 and 765PR2 were isolated from the products of religation of pADJ765 DNA partially digested with *Pst*I. The collinearity of these segments with pADJ765 was confirmed by partial digest mapping of the *Pst*I and *Bam*HI sites in pADJ765 and the two derivatives. The probe mp762MRS carries both terminal *Msp*I-*Eco*RI fragments of pADJ762 cloned into the *Acc*I site of the M13 single-stranded phage vector mp8 [23]. This clone was constructed by circularizing the *Eco*RI insert fragment of pADJ762, digesting the ligated product with *Msp*I, isolating the appropriate junction fragment, and ligating it with phosphatased *Acc*I-digested mp8.

For all of the constructions described, restriction enzymes and T4 DNA ligase were obtained from commercial suppliers and used under recommended conditions. Alkaline phosphatase from calf intestine was used to remove fragment-terminal phosphates. Cloning in the M13mp8 vector was carried out with a system from BRL.

*Restriction Mapping*

The relative order of restriction sites in pADJ762 and pADJ765 was determined by partial digest mapping essentially as described in [24], with a modification described previously [25]. The *Cla*I-*Pvu*I fragments spanning the inserts in these plasmids were end-labeled at the *Cla*I site and subjected to partial digestion with each of the relevant

enzymes, including *Msp*I, *Taq*I, *Bam*HI, *Pst*I, and *Bcl*I. The products were electrophoresed in 0.7% agarose gels, which were then dried and exposed to X-ray film. The order of close sites produced by two different enzymes was unambiguously determined by co-electrophoresis in adjacent lanes of the partial products generated by the two enzymes and direct comparison of the relative mobilities. In many cases, the results were confirmed by analysis of single- and double-digest products that often also provided more accurate estimates of distances between sites. To map the *Bcl*I sites of pADJ762 and pADJ765, labeled fragments were prepared from plasmid DNAs grown in the *dam*$^-$ strain SK383.

On the map of JA76 presented by Adams et al. [21], the position of one *Eco*RI site in the region from which pADJ762 and pADJ765 are derived is ambiguous. By comparing the results of single- and double-digests of DNA from JA76 and CH4A [26] with *Bgl*II, *Bam*HI, and *Eco*RI with the map of these sites in the subclones derived from JA76, it is possible to infer that there is a *Bam*HI site within 500 base pairs (bp) of the *Eco*RI site joining the left arm of the CH4A vector to the genomic segment and a *Bgl*II site at least 6 kb from this junction. These results are consistent with the order of fragments presented in the RESULTS. The length of the 12.2-kb *Bam*HI fragment of JA76 that spans the junction of the CH4A vector right arm with the insert DNA indicates that the orientation of the fragment cloned in pADJ765 is as shown in figure 2. The orientation of the 5.5-kb fragment with respect to the 3.3-kb fragment has been determined as described in RESULTS.

### *Calculation of Disequilibrium Measures*

Values of D were calculated according to Hartl [27], employing the directly determined haplotype and gene frequency values. D is equal to the difference $P_{11}P_{22}$-$P_{12}P_{21}$, where $P_{11}$ is the frequency of the haplotype with allele 1 at both loci, etc. D is a measure of the deviation of the observed haplotype frequencies from the values expected if the alleles of the two loci were randomly associated. Values of the correlation coefficient, r, were calculated as $r = D/(p_1p_2q_1q_2)1/2$ [27].

## RESULTS

### *Origin of Probe pADJ762 and Characterization of Site Polymorphisms*

In the course of experiments designed to examine possible polymorphism at the boundaries of members of the Kpn family of repeated sequences, single-copy fragments adjacent to five different Kpn sequences were identified and subcloned. The genomic sequences including these repeat elements were initially isolated [21] from a Charon 4A library of partial *Eco*RI products. The repeats were subsequently characterized as being related to the KpnI family [28]. The single-copy subclones were hybridized to Southern transfers of panels of human DNAs digested with a number of different restriction endonucleases to screen for polymorphism. None of the probes revealed significant variation at the boundary of the repeated sequence; however, probe pADJ762, from phage JA76, was found to reveal frequent polymorphism at four different restriction sites.

The four common dimorphic sites observed with pADJ762 are a *Taq*I site, a *Bcl*I site, and two *Msp*I sites. The fragment patterns revealed by pADJ762 in human DNAs digested with these enzymes are shown in figure 1. The common *Taq*I patterns consist of allelic fragments of 8.3 kb and 3.2 kb and a constant 3.9-kb fragment (fig. 1*A*). The *Bcl*I alleles are an 11.6-kb fragment or a 4.3-kb fragment. A constant 7.0-kb band is also seen in *Bcl*I digests (fig. 1*B*). The *Msp*I fragments are invariant 1.3-kb and 1.1-kb fragments, allelic fragments of 2.1 kb and 1.7 kb as shown in figure 1*C*, and a second set of allelic fragments, one of

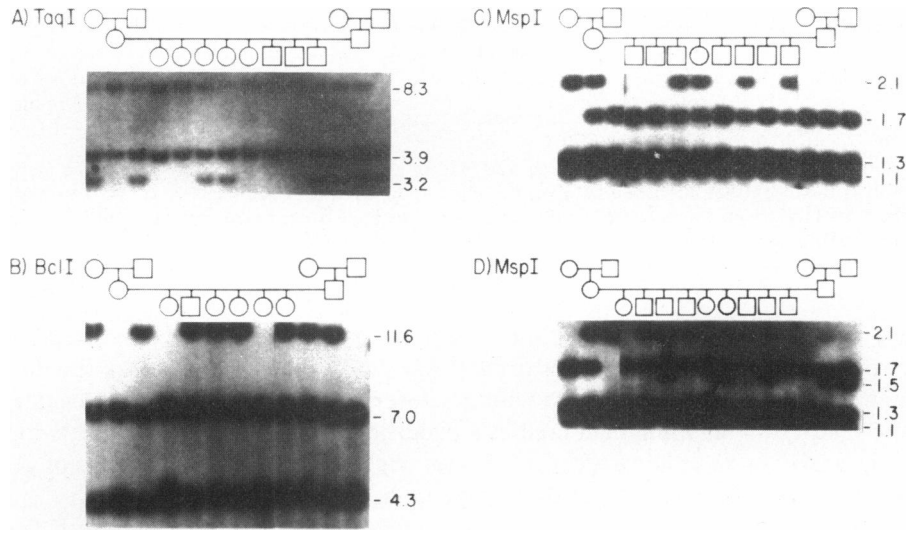


FIG. 1.—Examples of the common polymorphic fragment patterns detected by probe pADJ762 and their inheritance in nuclear families. The fragment patterns are those revealed by *A*) *Bcl*I, *B*) *Taq*I, *C*) *Msp*I (2.1-kb and 1.7-kb alleles), and *D*) *Msp*I (1.3- and 1.5-kb alleles). The 1.3-kb allele of the latter *Msp*I polymorphic system comigrates with an invariant 1.3-kb *Msp*I fragment.

which is seen as an additional band at 1.5 kb in figure 1*D*. This fragment is allelic with a 1.3-kb fragment that comigrates with the invariant 1.3-kb fragment as demonstrated below.

*Localization of Polymorphic Sites*

To determine the relative physical location of the polymorphic sites revealed by probe pADJ762, restriction maps of the cloned segment and an adjacent segment subcloned from the phage JA76 were determined. Figure 2 shows the origin of these subclones relative to the map of JA76. The locations of *Msp*I, *Taq*I, *Bcl*I, *Bam*HI, *Pst*I, and *Bgl*II sites within the two subclones were determined as described in MATERIALS AND METHODS, and the resulting map is shown in figure 3. To facilitate the analysis of the positions of the polymorphic sites, segments of pADJ762 and pADJ765 were further subcloned as described in MATERIALS AND METHODS. The locations of these segments are indicated as open bars above the map of restriction sites shown in figure 3.

Hybridization of probe 765PR2 to *Bcl*I-digested DNAs reveals alleles of either an 11.6-kb fragment or a 7.3-kb and a 4.3-kb fragment, indicating that this probe spans a polymorphic site. The probe 765BB5 detects allelic 7.3-kb and 11.6-kb fragments, showing that it lies within the 7.3-kb genomic *Bcl*I fragment that flanks the polymorphic site. Since pADJ762 detects the allelic fragments of 11.6 and 4.3 kb and an invariant 7.0-kb fragment, the map order of *Bcl*I fragments must be 7.0 kb-4.3 kb-7.3 kb, with the polymorphic site between the 4.3- and 7.3-kb fragments. The *Bcl*I site found in the segment cloned in pADJ765 must correspond to this polymorphic site since the genomic blotting data predict only

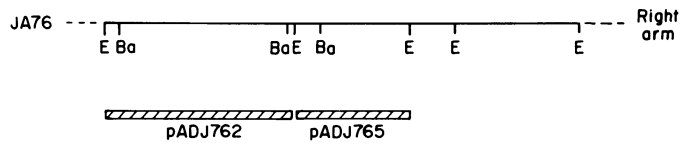


FIG. 2.—Derivation of the subclones pADJ762 and pADJ765 from the phage JA76. The map of JA76 is based on that described [21] with modifications as described in MATERIALS AND METHODS. The *Eco*RI fragments represented by the *two hatched bars* have been cloned separately into the *Eco*RI site of pBR325.

one site, the polymorphic one, in this region. In addition, the orientation of the pADJ762 fragment with respect to the pADJ765 fragment must be as shown in figure 3 in order to be consistent with the observed genomic *Bcl*I fragment sizes. The restriction mapping data predict an additional 0.2-kb *Bcl*I fragment, which is shown in figure 3; however, this was not observed in the genomic blots of gels designed to resolve the larger *Bcl*I fragments.

The probe 765PR2 detects *Taq*I allelic states defined by either an 8.3-kb *Taq*I fragment or fragments of 3.2 and 5.1 kb, indicating that it spans a polymorphic *Taq*I site. The probe 765BB5 detects the allelic 8.3- and 5.1-kb fragments, while pADJ762 detects allelic 3.2- and 8.3-kb fragments and a constant 3.9-kb fragment. The inferred order of *Taq*I fragments is 3.9 kb-3.2 kb-5.1 kb, with the polymorphic site between the 3.2-kb and 5.1-kb fragments. The allelic patterns indicate that the *Taq*I site found in pADJ765 between the points defined by the right end of pADJ762 and the left end of 765BB5 must represent the polymorphic site, since the patterns are consistent only with the presence of a single, polymorphic site in this region. The map of *Taq*I sites in pADJ762 predicts an additional 0.4-kb fragment between the 3.9- and 3.2-kb fragments. This fragment has occasionally been observed in Southern transfers and is also inferred from the allele fragment sizes of a rare polymorphism at the *Taq*I(R) site (fig. 3) described below.

The variant *Msp*I site that accounts for the allelic 2.1- and 1.7-kb fragments is located 0.4 kb beyond the left end of pADJ762. This is inferred from probing *Msp*I and *Msp*I plus *Eco*RI digests of DNAs of each genotypic constitution with the 762PP3 probe. In *Msp*I-digested DNAs, this probe detects the allelic 1.7-kb and 2.1-kb bands and a constant band of 1.3 kb. However, in *Eco*RI plus *Msp*I digests of the same DNAs, only a monomorphic 1.3-kb doublet is observed, showing that the *Eco*RI site at the left end of the pADJ762 cloned segment separates the 762PP3 probe from the polymorphic *Msp*I site. The lengths of the genomic *Msp*I and *Msp*I plus *Eco*RI fragments indicate that the polymorphic site lies 0.4-kb distal to the *Eco*RI site, with another *Msp*I site 0.4 kb distal to the variant site (fig. 3).

The location of the site designated *Msp*I(E) is defined first by the failure to detect with probe 762PP3 the characteristic 1.5-kb band in *Msp*I digests of DNAs from individuals heterozygous for the *Msp*I(E) site. This result places the site to the right of the *Msp*I site at coordinate 2.7 on the map of figure 3. The probe 762BB1 detects only the allelic 1.3-kb and 1.5-kb fragments, showing that it lies within the 1.3-kb allelic fragment. The positions of *Msp*I sites in this region

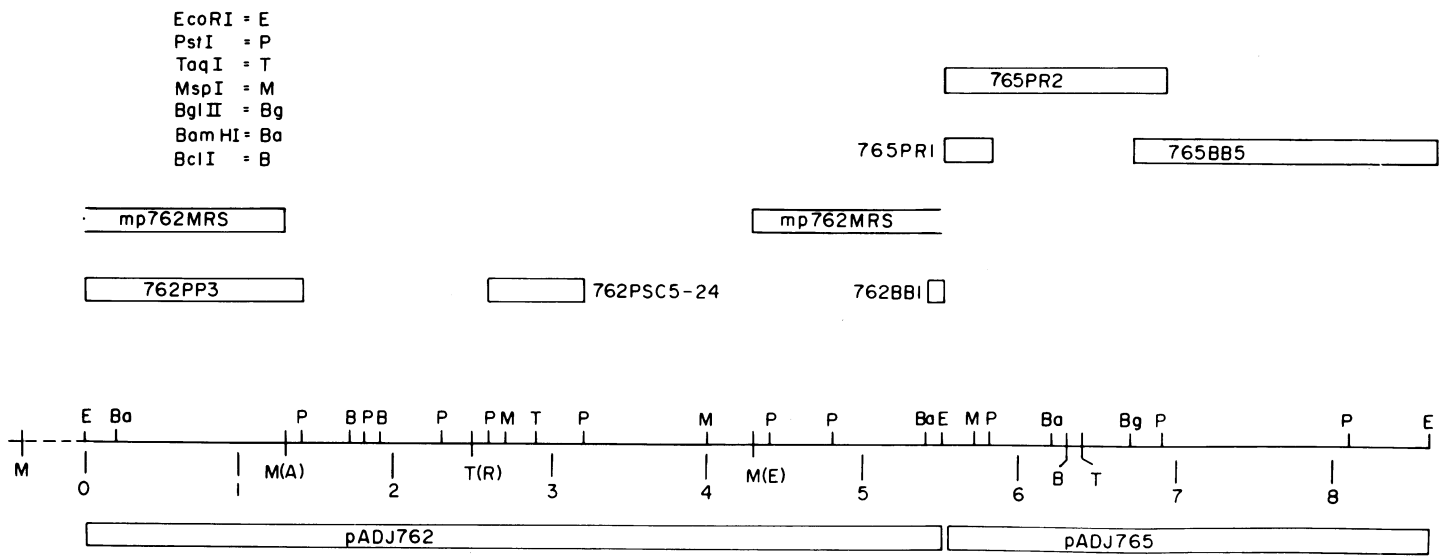


FIG. 3.—Map of restriction sites in pADJ762 and PADJ765. The locations of *Msp*I, *Taq*I, *Bcl*I, *Pst*I, *Bam*HI, *Eco*RI, and *Bgl*II sites in pADJ762 and pADJ765 were determined as described in MATERIALS AND METHODS, and the relative orientation of these two fragments was determined as described in the text. *The open bars below the map* indicate the extent of these two clones. *The open bars above the map* represent the subcloned segments used to determine the locations of polymorphic sites. Those sites that have been observed to vary among the individuals studied are *marked below the map*. All of these sites are present in the cloned segments, indicating that the JA76 clone represents a haplotype found at a relatively high frequency (.33) in the Utah population examined. *The intermittent line at the left end of the map* indicates an uncloned genomic region where the marked *Msp*I site is inferred to lie.

are consistent with the identity of the variant *Msp*I site with the site detected at coordinate 4.3 on the map of figure 3, accounting well for the 1.3-kb and 1.5-kb allelic sizes. The possibility that the *Msp*I site at coordinate 5.7 or any site to the right of the *Eco*RI junction site is variant is ruled out by the finding that probe 765PR2 detects the 1.3-kb and 1.5-kb alleles in *Msp*I digests but detects neither of these fragments nor any derivative variant fragments in *Msp*I plus *Eco*RI digests. The comigration of the 1.3-kb allele with an invariant fragment precludes the direct discrimination of heterozygotes from homozygotes for the 1.5-kb allele with the pADJ762 probe. For 11 of the 16 individuals scored with this allele, heterozygosity is inferred from the fact that one or more offspring does not carry the 1.5-kb allele. For the other five individuals, *Msp*I digests of the DNAs were examined with the probe mp762MRS (fig. 3), which does not include homology to the invariant 1.3-kb band, and the genotype was determined directly. One individual homozygous for the 1.5-kb allele was found, confirming the absence of detectable homology between the comigrating 1.3-kb fragments.

*Characterization and Localization of Two Rare Polymorphic Sites*

One rare *Msp*I site variant and one rare *Taq*I variant were observed in the families shown in figure 4. The rare *Msp*I allele (fig. 4*A*) is a 3.1-kb fragment that is allelic with the 2.1- and 1.7-kb fragments. DNAs from individuals with this allele were digested with *Msp*I or *Msp*I plus *Eco*RI and probed with 762PP3. In the *Eco*RI plus *Msp*I digest, an allelic band of 2.7 kb is detected in addition to the expected 1.3-kb band, indicating that the 3.1-kb *Msp*I fragment spans the *Eco*RI site at the left end of 762PP3 and that the variant *Msp*I site is that at coordinate 1.3, which is marked as *Msp*I(A). The *Msp*I site at the left end of the 3.1-kb fragment is also variant, as described in the previous section; however, the 0.4-kb difference between the *Msp*I and *Msp*I-*Eco*RI fragments detected by

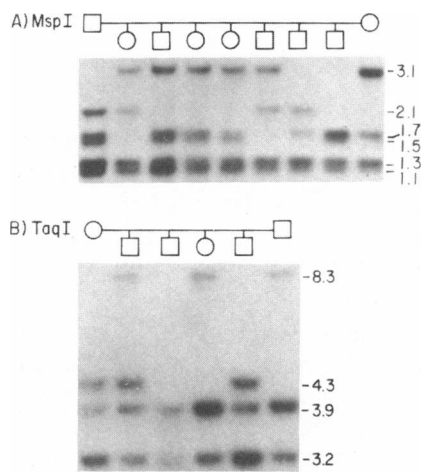


FIG. 4.—Examples of rare polymorphic patterns detected by the probe pADJ762 and their inheritance in nuclear families. The fragment patterns are revealed by *A*) *Msp*I and *B*) *Taq*I.

TABLE 1

OCCURRENCES OF THE DIFFERENT HAPLOTYPES AT THE *D11S12* LOCUS AMONG 100 UNRELATED INDIVIDUALS

| Haplotype | *Msp*I | *Msp*I(A) | *Taq*I(R) | *Msp*I(E) | *Bcl*I | *Taq*I | No. |
|---|---|---|---|---|---|---|---|
| I | + | + | + | + | − | + | 80 |
| II | + | + | + | + | + | + | 67 |
| III | − | + | + | + | − | − | 29 |
| IV | + | + | + | − | − | + | 17 |
| V | − | + | + | + | − | + | 1 |
| VI | + | − | + | + | − | + | 1 |
| VII | + | + | + | + | + | − | 1 |
| VIII | − | + | + | + | + | + | 1 |
| IX | − | + | + | + | + | − | 1 |
| X | + | + | + | + | − | − | 1 |
| XI | + | + | − | + | + | + | 1 |

NOTE: The variable restriction sites indicated are arranged in the same order as they are found in the genome. "+" indicates presence of the site and "−" indicates its absence.

762PP3 in individuals with the 3.1-kb *Msp*I allele indicates that the site is present on the chromosome that carries the 3.1-kb fragment allele. When probed with pADJ762, DNAs from individuals with the rare *Taq*I polymorphism show alleles consisting of a 4.3-kb fragment or both a 3.9-kb and a 0.4-kb fragment (the latter is not shown in fig. 4*B*). The probe 762PSC5-24 hybridizes to the 4.3-kb and 0.4-kb *Taq*I alleles but not to the 3.9-kb allele. This probe also hybridizes to the 3.3-kb *Taq*I fragment, indicating that the order of fragments is 3.9 kb-0.4 kb-3.3 kb, with the polymorphic site between the 3.9-kb and 0.4-kb fragments at the site labeled *Taq*I(R) in figure 3.

*Population Genetic Characterization of Site Polymorphisms*

*Taq*I, *Msp*I, and *Bcl*I digests of DNAs from approximately 400 individuals have been probed with pADJ762 or the appropriate derivative. These include 170 individuals from one segment of a single kindred, K1085, and more than 225 individuals from 20 three-generation family units where samples have been donated by all four grandparents, both parents, and a number of their offspring [29]. The grandparents from the latter group and spouses from the K1085 group who were not in the direct line of descent are the unrelated individuals in the analyses presented below. The observation of proper inheritance of the restriction fragment alleles in these large family studies confirms their Mendelian character and rules out sporadic partial digestion or DNA contamination as the cause of the apparent polymorphisms.

The haplotypes of 200 chromosomes from 100 unrelated individuals have been determined by inspection in individuals where no more than one site is heterozygous or by inference from inheritance of haplotypes by the descendants of scored individuals. Among the 100 randomly chosen individuals, no instances of ambiguous haplotypes occurred. The number of occurrences of each observed type is presented in table 1. Only four of the possible 16 haplotypic combinations of the four different common polymorphisms are found at a frequency greater than

5%. This is due partly to the low minor allele frequency of 3 of the polymorphisms and partly to the significant degree of disequilibrium between pairs of polymorphic sites. Table 2 presents two different measures of disequilibrium between each pair of the four common polymorphic sites. By both of these measures, the correlation coefficient r and the value of D [27], it is clear that the *Msp*I and *Taq*I sites, which are the farthest separated pair of polymorphic sites at this locus, show a highly significant degree of association. Interestingly, both sites show a lesser degree of association with polymorphic sites located between them. This is particularly striking for the *Msp*I(E) site, where both of the flanking *Msp*I and *Taq*I sites show insignificant values of r and D with the *Msp*I(E) site. These flanking sites also share the same degree of disequilibrium with the *Bcl*I site polymorphism. In this case, the D and r values are significant, indicating partial disequilibrium with the *Bcl*I site. The very similar degree of correlation of the internal *Bcl*I and *Msp*I(E) sites with the external *Taq*I and *Msp*I sites is expected because of the high degree of association of the external sites with each other. The interesting feature of this locus is that the observed degree of disequilibrium between two sites is not always a decreasing function of their physical separation, suggesting that recombination may be relatively insignificant compared to the processes of mutation or genetic drift or possibly selection in the formation of the constellation of haplotypes found in this population.

## DISCUSSION

We have described a new highly polymorphic locus in human DNA with four major haplotypic combinations of three polymorphic sites: a *Taq*I site, an *Msp*I site, and a *Bcl*I site. A fourth polymorphism at a second *Msp*I site does not contribute significant additional haplotypic diversity because its alleles are very highly correlated with the alleles of the *Taq*I site polymorphism. We have shown that the strongly correlated *Msp*I and *Taq*I site polymorphisms have a significantly weaker correlation with the polymorphic *Bcl*I and *Msp*I sites that lie between them; that is, they show a smaller degree of linkage disequilibrium with sites

TABLE 2

MEASURES OF DISEQUILIBRIUM BETWEEN PAIRWISE COMBINATIONS OF THE POLYMORPHIC SITES OBSERVED AT *D11S12*

| A. | *Msp*I | *Msp*I(E) | *Bcl*I |
|---|---|---|---|
| *Msp*I(E) | −0.133 | ... | ... |
| *Bcl*I | 0.267 | 0.226 | ... |
| *Taq*I | 0.926 | −0.133 | 0.267 |

| B. | *Msp*I | *Msp*I(E) | *Bcl*I |
|---|---|---|---|
| *Msp*I(E) | −0.014 | ... | ... |
| *Bcl*I | 0.047 | 0.030 | ... |
| *Taq*I | 0.124 | −0.014 | 0.047 |

NOTE: Calculation of these values is described in MATERIALS AND METHODS. Values in (A) are of the value of the correlation coefficient r. Values in (B) are of the disequilibrium measure D.

that lie in the DNA between them than they do with each other. The *Taq*I site polymorphism shows a significantly weaker correlation with a *Bcl*I site less than 0.2 kb away than it does with the *Msp*I site that is 6.8 kb distant. This result indicates that the degree of linkage disequilibrium is not a reliable indicator of relative physical separation over distances of the order of tens of kilobases. The simplest explanation for this situation is that the processes of mutation and genetic drift are primarily responsible for the observed associations. The current estimate of recombination frequency as 1 centimorgan per million bp is consistent with the supposition that recombination in a 7-kb interval is less than 0.01 centimorgan and, hence, potentially less significant than mutation and drift in the formation of new haplotypes [16].

Questions concerning the relative effects of selection, recombination, mutation, and genetic drift have arisen in similar studies of the haplotypes defined by restriction site polymorphisms at the β-globin locus [14, 16]. At that locus, where more than eight different polymorphic sites are found in one 60-kb region, there is no simple correlation of disequilibrium with physical distance between polymorphic sites in any of the Mediterranean, American black, or Southeast Asian populations examined. Moreover, in each population, there are only two or three major haplotypes, and these are not related to one another by single mutational or recombinational events. This observation suggests that intermediate or ancestral haplotypes have been lost through selection or drift. In contrast, three of the four major haplotypes reported here for *D11S12* are similar, with haplotypes II and IV differing at only a single site from haplotype I (table 1). Only haplotype III differs from each of the other three major haplotypes at two or more restriction sites. Even in this case, however, two of the minor haplotypes, V and X, represent possible intermediates between haplotypes III and I. Whether the apparent difference in the degree of relatedness between major haplotypes reflects a significant difference in the forces affecting the evolution of the β-globin and *D11S12* loci may be further assessed by characterization of additional polymorphic sites at the *D11S12* locus, examination of the *D11S12* haplotype patterns and frequencies in other population groups and by population studies of additional loci with multiple restriction site polymorphisms.

The comparison of population genetic properties of the site polymorphisms at the *D11S12* locus with the molecular map of this region also provides additional confirmation that DNA regions with multiple site polymorphisms can be valuable genetic markers. The finding that two polymorphisms at sites that are about 0.2 kb distant show a significant degree of random association suggests that valuable genetic markers can be "built" by detection of variable sites within a fairly small DNA segment.

The probe pADJ762 has been mapped to chromosome 11p by use of the human-rodent hybrid cell mapping method [30]. This assignment has been confirmed by in situ hybridization [31]. The latter technique indicates that the *D11S12* locus is on the distal end of 11p, beyond 11p5.1. Genetic linkage studies show that the locus is between and closely linked to the β-globin and insulin loci [29], which have also been physically localized to the distal end of 11p [2]. The polymorphisms at the *D11S12* locus have proven useful in the definition of the

genetic map of this region and should prove valuable in continuing studies of the relationship of physical and genetic distances in this region of chromosome 11. The allele and haplotype frequencies that we report here should be typical of other Northern European populations [32]; however, other human populations are as yet uncharacterized with respect to their degree of heterozygosity at this locus.

## ACKNOWLEDGMENTS

We are indebted to Jesse Adams for making available his genomic clones containing KpnI repeats. We are grateful to Arnold Oliphant, Charles Aquadro, Bruce Weir, and Lyn Jorde for helpful discussions and to Rick Kelley for providing the SK383 bacterial strain.

## REFERENCES

1. BISHOP DT, CANNINGS C, SKOLNICK M, WILLIAMSON JA: The number of polymorphic DNA clones required to map the human genome, in *Statistical Analysis of DNA Sequence Data,* edited by WEIR BS, New York, Marcel Dekker, 1983
2. SKOLNICK MH, WILLARD HF, MENLOVE LA: Report of the Committee on Human Gene Mapping by Recombinant DNA Techniques. Human Gene Mapping (7) (1984): Seventh International Workshop on Human Gene Mapping. *Cytogenet Cell Genet* 37(1–4):1–666, 1984. Simultaneous publication in *Birth Defects: Orig Art Ser* 20(2):210–273, 1984
3. 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
4. BARKER D, SCHAFER M, WHITE R: Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. *Cell* 36:131–138, 1984
5. BELL GI, KARAM MH, RUTTER WJ: Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. *Proc Natl Acad Sci USA* 78:5759–5763, 1981
6. CHAPMAN BS, VINCENT KA, WILSON AC: Length polymorphisms, gene correction and recombination in the zeta globin region of humans. Submitted for publication
7. 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
8. 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
9. GOLDFARB M, SHIMUZU K, PERUCHO M, WIGLER M: Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. *Nature* 296:404–409, 1982
10. TAPAROWSKY E, SUARD Y, FASANO O, SHIMUZU K, GOLDFARB M, WIGLER M: Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. *Nature* 300:762–765, 1983
11. TAUB RA, HOLLIS GF, HIETER PA, KORSMEYER S, WALDMANN TA, LEDER P: Variable amplification of immunoglobulin λ light chain genes in human populations. *Nature* 304:172–174, 1983
12. WYMAN AW, WHITE R: A highly polymorphic locus in human DNA. *Proc Natl Acad Sci USA* 77:6754–6758, 1980
13. LANGE K, BOEHNKE M: Some combinatorial problems of DNA restriction fragment length polymorphisms. *Am J Hum Genet* 35:177–192, 1983
14. ANTONARAKIS SE, BOEHM CD, GIARDINA PJV, KAZAZIAN HH JR: Nonrandom association of polymorphic restriction sites in the β-globin gene cluster. *Proc Natl Acad Sci USA* 79:137–141, 1982

15. ORKIN SH, KAZAZIAN HH JR, ANTONARAKIS SE, ET AL.: Linkage of β-thalassaemia mutations and β-globin gene polymorphisms in human β-globin gene cluster. *Nature* 296:627–631, 1982
16. ANTONARAKIS SE, ORKIN SH, KAZAZIAN HH JR, ET AL.: Evidence for multiple origin of the $B^E$-globin gene in Southeast Asia. *Proc Natl Acad Sci USA* 79:6608–6611, 1982
17. MIGONE N, FEDER J, CANN H, ET AL.: Multiple DNA fragment polymorphisms associated with immunoglobulin μ chain switch-like regions in man. *Proc Natl Acad Sci USA* 80:467–471, 1983
18. SOUTHERN EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503–517, 1975
19. ALWINE JC, KEMP DJ, PARKER BA, ET AL.: Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. *Methods Enzymol* 68:220–242, 1979
20. WAHL GM, STERN M, STARK GR: Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization by using dextran sulfate. *Proc Natl Acad Sci USA* 76:3683–3687, 1979
21. ADAMS JW, KAUFMAN RE, HARRISON M, NIENHUIS, AW: A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. *Nucleic Acids Res* 8:6113–6128, 1980
22. BOLIVAR F: Construction and characterization of new cloning vehicles, III. Derivatives of plasmid pBR322 carrying unique EcoRI sites for selection of EcoRI generated recombinant DNA molecules. *Gene* 4:121–136, 1978
23. MESSING J: New M13 vectors for cloning. *Methods Enzymol* 101:20–78, 1983
24. SMITH HO, BIRNSTIEL ML: A simple method for DNA restriction site mapping. *Nucleic Acids Res* 3:2387–2398, 1976
25. BARKER D, HOFF M, OLIPHANT A, WHITE R: Sequence specificity of a second type II restriction endonuclease from thermus aquaticus. *Nucleic Acids Res* 12:5567–5581, 1984
26. DEWET JR, DANIELS DL, SCHROEDER JL, ET AL.: Restriction maps for twenty-one Charon vector phages. *J Virol* 33:401–410, 1980
27. HARTL DL: *Principles of Population Genetics*. Sunderland, Mass., Sinauer, 1980, pp 110–111
28. SHAFIT-ZALGARDO B, BROWN FL, MAIO JJ, ADAMS, JW: KpnI families of long, interspersed repetitive DNAs associated with the human β-globin gene cluster. *Gene* 20:397–407, 1982
29. WHITE R, LEPPERT M, BISHOP DT, ET AL.: High resolution linkage mapping in the human. Submitted for publication
30. NAYLOR SL, SAKAGUCHI AY, BARKER D, WHITE R, SHOWS TB: DNA polymorphic loci mapped to human chromosomes 3, 5, 9, 11, 17, 18 and 22. *Proc Natl Acad Sci USA* 81:2447–2451, 1984
31. HARPER ME: Chromosome mapping of single-copy sequences by *in situ* hybridization, in *Banbury Report 14: Recombinant DNA Applications to Human Diseases*, edited by CASKEY CT, WHITE RL, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, 1983, pp 159–165
32. MCCLELLAN T, JORDE LB, SKOLNICK MH: Genetic distances between the Utah Mormons and related populations. Submitted for publication