Am. J. Hum. Genet. 44:671–678, 1989

Identification and Characterization of 23 RFLP Loci by Screening Random Cosmid Genomic Clones

Donald W. Bowden, Hans Müller-Kahle, Thomas C. Gravius, Cynthia Helms, Diane Watt-Morgan, Philip Green, and Helen Donis-Keller

Human Genetics Department, Collaborative Research, Incorporated, Waltham, MA

Summary

As part of our search for polymorphic DNA probes, we have screened cosmids from a human genomic DNA library for their ability to reveal RFLPs. A total of 101 randomly isolated cosmid clones were tested in Southern hybridizations for polymorphic band patterns. Fifty-four of these clones revealed RFLPs with one or more of nine restriction enzymes. Twenty-three of these clones have been further characterized and assigned to 10 different chromosomes by linkage analysis or by hybridization to panels of human-hamster hybrid cell lines. Fifteen of the probes have heterozygosities ≥.5. The relative efficiency of *Rsa*I and *Pst*I restriction enzymes in detecting polymorphism was different from results obtained with libraries constructed in bacteriophage vectors. Screening randomly selected cosmid probes is an efficient method for detecting RFLPs.

Introduction

It is now well documented that the construction of a detailed linkage map of the human genome will require large numbers of DNA polymorphisms. For example, data for 404 polymorphic loci, including 391 RFLPs, were recently used to construct a linkage map of the human genome (Donis-Keller et al. 1987) with an average spacing of approximately 10 cM. If a map with average spacing of 1 cM is to be constructed, a minimum of 3,300 RFLPs will be required. For RFLP markers to be useful in determining the genetic location of disease genes, heterozygosity of the markers should exceed 40% (PIC ≥ .375), and the alleles should be readily determined by Southern blot hybridization.

Descriptions of two large-scale screening studies have been published recently (Nakamura et al. 1987; Schumm et al. 1988). In the earlier study, clones detecting variable-number-of-tandem-repeat (VNTR) RFLPs were isolated after screening a cosmid genomic library with oligonucleotide probes. Over 350 RFLPs were identified. In the later study over 1,000 lambda bacteriophage clones containing single-copy genomic sequences and over 500 randomly isolated bacteriophage clones were screened for polymorphism, yielding approximately 500 RFLPs. Even with these resources, it is apparent that more RFLPs and more-efficient strategies would be helpful for map construction.

In addition to screening genomic libraries constructed in bacteriophage lambda (Schumm et al. 1988), we have investigated the efficiency of finding polymorphic loci by screening randomly selected cosmids containing human genomic DNA. We were interested in comparing this approach with those described above, for purposes of determining the efficiency with which probes could be isolated, the PICs of such probes, and the chromosome distribution of the RFLP loci.

Material and Methods

Probe Preparation

Both construction of the cosmid library by using the vector c2RB (Bates and Swift 1983) and plating conditions have been described in detail elsewhere (Bowden et al. 1988). Random cosmids were picked from selective plates, inoculated into LB broth (10 g tryptone/li-

Received August 8, 1988; final revision received January 24, 1989.

Address for correspondence and reprints: Dr. Donald W. Bowden, Collaborative Research, Inc., 1365 Main Street, Waltham, MA 02154.

0002-9297/89/4405-0007$02.00

ter, 5 g yeast extract/liter, 5 g NaCl/liter [pH7.0]) containing 250 μg ampicillin/ml, 0.2% (w/v) glucose, and 0.01 M Tris-HCl (pH 7.0), and grown with modest agitation overnight at 37°C. Cosmid DNA was isolated by the alkaline lysis method of Ish-Horowicz and Burke (1981), with the addition of a phenol-extraction step prior to the isopranol precipitation. In some cases cosmid DNA was further purified by centrifugation in gradients of CsCl containing ethidium bromide. Cosmids CRI-C1–CRI-C86 were chosen from platings on *Escherichia coli* 1046 (Bates and Swift 1983), and cosmids CRI-C87–CRI-C101 were chosen from platings on χ2613, a strain defective in *mcr* (modified cytosine restriction) (Raleigh and Wilson 1986; Raleigh et al. 1988). Strain χ2613 was obtained from E. Raleigh (New England Biolabs).

Screening for Polymorphism

Labeled cosmid probes were screened for their ability to reveal RFLPs by hybridization to DNAs from five unrelated individuals, each DNA type being digested with nine restriction enzymes (*Bam*HI, *Bgl*II, *Eco*RI, *Hin*cII, *Hin*dIII, *Msp*I, *Pst*I, *Rsa*I, and *Taq*I) (Schumm et al. 1988). ^{32}P-labeled probes were prepared by nick-translation as described by Schumm et al. (1988). Human genomic DNA isolation, restriction-enzyme digestion, agarose gel electrophoresis, preparation of Southern blot filters, prehybridization, hybridization, washing, and exposure of Southern blot filters were carried out as described elsewhere (Donis-Keller et al. 1987; Schumm et al. 1988). Hybridizations were carried out in the presence of excess, unlabeled human DNA, in order to block repeat-containing sequences from hybridizing to filter-bound genomic DNA (Litt and White 1985).

Pedigree Analysis

Pedigree analysis was carried out on human DNA samples from 3-generation families provided by Centre d'Etude du Polymorphisme Humain (CEPH) or from DNA prepared from transformed lymphoblastoid lines which were purchased from the Camden Cell Repository or were a gift from R. White. These families were K884, K13291, K13292, K13293, K13294, K1331, K1332, K1333, K1334, K1340, K1341, K1344, K1345, K1346, K1349, K1350, K1362, K1413, K1416, K1418, and K1421. Cosmid probes that detected polymorphism were labeled and initially hybridized to Southern blots containing genomic DNA from parents of these families to confirm that the probe revealed polymorphism, to find which families were informative, and to provide an estimate of PIC. Genotypic data were gathered by subsequent hybridization to Southern blots for these informative families.

Linkage Analysis and Chromosome Assignment

Genotypic data were analyzed, as described elsewhere (Donis-Keller et al. 1987), for maximum-likelihood multipoint linkage (Lander and Green 1987) with a data set containing genotypic information on over 400 polymorphic loci which included mapping data from 23 cosmid probes. Two-point linkage analysis was also carried out. These loci have been assigned to 22 autosomes and the X chromosome (Donis-Keller et al. 1987). Cosmid probes were assigned to linkage groups—and thus to specific chromosomes—by using the criteria of lod 4 as proof of linkage (Donis-Keller et al. 1987). In one case, where two-point lod scores were not ⩾4.0, chromosome assignment was made by hybridization to panels of rodent-human hybrid cell lines containing varying human chromosome complements as described by Donis-Keller et al. (1987).

χ^2 *Analysis*

The significance of differences between sets of data was tested using the χ^2 test for independence (Sokal and Rohlf 1969, pp. 589–607). Differences were considered significant if $P \leqslant .05$.

Results

Screening and Identification of Polymorphic Cosmid Clones

One rationale for using cosmids to screen for polymorphism is that they contain relatively large DNA inserts (averaging 38 kb for the library used in the present study) and consequently would screen a large expanse of human DNA for polymorphic sites. We have screened 101 randomly picked cosmid clones from a human genomic DNA library for their ability to detect RFLPs. A summary of these screening results is shown in table 1. Fifty-four (53%) of the probes revealed RFLPs with one or more of the nine enzymes used in our study (see Material and Methods). Nineteen (19%) of the probes revealed more than one allelic system. These took the form of more than one unrelated RFLP revealed with the same enzyme or of more than one unrelated RFLP revealed with different enzymes. Of these 19 probes over half (10 probes) revealed three or more allelic systems, suggesting clustering of polymorphic sites. Seven of the polymorphic loci had more than two alleles.

Table 1

Summary of Screening for Polymorphic Cosmid Probes

Total no. of probes screened	101
Probes revealing polymorphism	54
Probes revealing more than one allelic system	19
Probes revealing three or more allelic systems	10
Patterns suggesting variable-length or insertion/deletion polymorphisms	3
Probes revealing more than two alleles in a single polymorphic system	7

Only three of the probes revealed RFLP patterns consistent with their being insertion/detection-type polymorphisms or VNTR probes (Nakamura et al. 1987), i.e., the same DNA samples show similar polymorphic patterns with several enzymes. It is interesting that all three RFLPs had low (<.3) PIC, and are consequently not useful in mapping studies.

As figure 1 shows, 19 of the probes have PIC values >.3, and 15 have heterozygosities ≥.5. For probes with more than one allelic system, the combined heterozygosities are given.

Relative Efficiencies of Restriction Endonucleases in Detecting Polymorphism

The relative efficiency of different restriction enzymes in detecting polymorphism is summarized in table 2, where the relative number of RFLPs (i.e., allelic systems) is expressed per 100 probes. Results with the cosmid probes are compared with the results from our study of bacteriophage lambda genomic clones (Schumm et al. 1988). For *Bam*HI, *Bgl*II, and *Msp*I there is no significant difference between the two studies in terms of relative efficiency. *Pst*I polymorphisms were detected at a higher rate (twofold), and the rate of *Eco*RI polymorphism is somewhat higher in our cosmid study. On the other hand, the rate at which *Rsa*I polymorphisms were found in our cosmid study was reduced over twofold, and the rate at which *Taq*I polymorphisms were found was also reduced. It is interesting to note that the total number of allelic systems/100 probes was essentially the same for cosmid and phage libraries (see Discussion).

Table 2

Restriction-Enzyme Efficiency in Detecting Polymorphism with Cosmid Probes

Enzyme	RFLPs/100 Cosmid Probes	RFLPs/100 Bacteriophage Lambda Probes[a]
*Bam*HI	6	8
*Bgl*II	8	7
*Eco*RI	12	8
*Hin*cII	2	ND
*Hin*dIII	7	8
*Msp*I	20	20
*Pst*I	18	9
*Rsa*I	6	15
*Taq*I	11	17.5
Total no. of allelic systems/100 probes	90	92.5

NOTE.—ND = not done.

[a] SOURCE: Schumm et al. (1988).

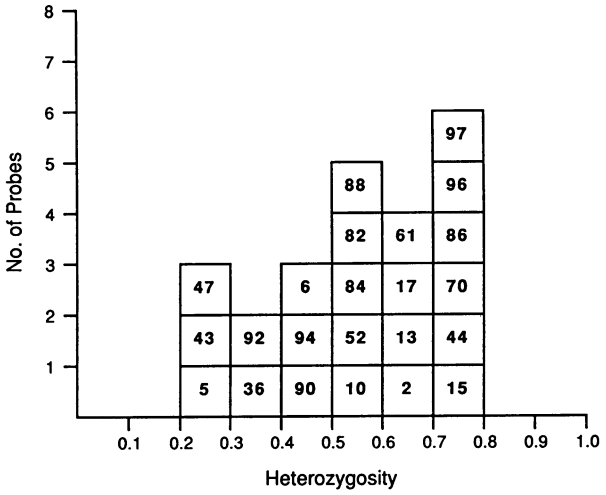

Figure 1 Histogram of heterozygosities vs. the number of probes in each class, based on the 23 mapped polymorphisms. Probe numbers are abbreviated to the numeral and are shown in the appropriate box.

Pedigree Analysis, Linkage, and Chromosome Assignment

Genotypic information was gathered by hybridization of 23 RFLP probes (totaling 34 allelic systems) to Southern blots containing CEPH family DNA. In addition to contributing information for linkage analysis, pedigree analysis was frequently essential to define allelic fragments for complex polymorphisms. Data from these studies are summarized in table 3, which lists the probe number, the enzyme(s) for which the probe is polymorphic, alleles and their fragment sizes, PIC and heterozygosity, and chromosome assignment based on either linkage analysis or hybridization to rodent-human hybrid cell lines (see Materials and Methods).

Table 3

Description of Polymorphic Cosmid Probes

Cosmid	Enzyme	Polymorphism(s)	Fragment Size	PIC (HET.)	Chromosome
CRI-C2[a]	*Pst*I	A1	6.5	.55 (.67)	12
		A2	5.0		
		A3	3.0		
		A4	6.5 + 3.0		
		A5	5.0 + 3.0		
	*Eco*RI	A1	13.7		
		A2	9.4 + 4.3		
CRI-C5	*Eco*RI	A1	6.3	.15 (.21)	X
		A2	5.9		
		B1	3.5		
		B2	3.2		
CRI-C6	*Pst*I	A1	7.8	.37 (.40)	X
		A2	6.9		
CRI-C10	*Pst*I	A1	2.3	.34 (.54)	11
		A2	2.0		
CRI-C13[a]	*Pst*I	A1	27	.30 (.30)	2
		A2	26		
		B1	18.0		
		B2	8.7		
		C1	14.0		
		C2	12.5		
CRI-C15	*Eco*RI	A1	12.3 + 5.6	. . . (.62)[b]	14
		A2	8.1 + 3.2		
		A3	8.1 + 2.9		
		A4	8.1 + 2.7		
		A5		2.7	
CRI-C17[a]	*Msp*I	A1	13.5	.57 (.69)	3
		A2	11.0		
		B1	6.7		
		B2	5.0		
CRI-C36[a]	*Eco*RI	A1	13.3	.29 (.33)	2
		A2	12.0		
CRI-C43[a]	*Eco*RI	A1	8.6	.21 (.21)	2
		A2	6.9		
CRI-C44[a]	*Msp*I	A1	6.2 + 2.7	.69 (.74)	5
		A2	6.2 + 2.2		
		A3	4.5 + 3.5 + 2.6 + 2.2		
		A4	3.3 + 2.6		
		A5	4.5 + 3.3 + 2.6 + 2.2		
		A6	4.5 + 3.3		
CRI-C47[a]	*Bam*HI	A1	10.5	.19 (.26)	4
		A2	6.5 + 5.5		
CRI-C52[a]	*Eco*RI	A1	12.0	.50 (.57)	
		A2	9.0 + 7.6		
		A3	8.5 + 7.6		
CRI-C61[a]	*Taq*I	A1	7.8	.57 (.64)	5
		A2	6.3		
		A3	4.9		
		A4	4.5		
CRI-C70[a]	*Msp*I	A1	10.0	.67 (.71)	14
		A2	9.6		
		B1	6.8		
		B2	6.7		
		C1	4.4		
		C2	4.2		

(continued)

Table 3 (continued)

Cosmid	Enzyme	Polymorphism(s)	Fragment Size	PIC (HET.)	Chromosome
CRI-C82[a]	*Msp*I	A1	9.4	.48 (.52)	4
		A2	8.3		
		B1	4.3		
		B2	2.8		
CRI-C84[a]	*Eco*RI	A1	8.5	.40 (.50)	2
		A2	7.6		
	*Hin*dIII	A1	4.7	.40 (.50)	
		A2	1.8 + 1.1		
CRI-C86[a]	*Pst*I	A1	10.0	.38 (.66)	12
		A2	8.1		
		B1	3.9		
		B2	3.7		
CRI-C88	*Msp*I	A1	10.1	.34 (.50)	X
		A2	8.1		
CRI-C90	*Hin*dIII	A1	12.3	.30 (.40)	1
		A2	6.7		
CRI-C92	*Pst*I	A1	18.0	.36 (.31)	11
		A2	14.0 + 5.8		
CRI-C94	*Hin*dIII	A1	9.4	.25 (.40)	8
		A2	8.2		
		A3	7.0		
CRI-C96[a]	*Bgl*II	A1	17.5	.62 (.70)	8
		A2	14.5 + 3.6		
		A3	10.2 + 4.3		
		A4	13.0 + 4.3		
		A5	10.2 + 4.3 + 3.6		
CRI-C97[a]	*Taq*I	A1	15.5	. . . (>.6)[b]	5
		A2	9.3		
		B1	6.8		
		B2	4.7		
		C1	7.0		
		C2	3.4		

[a] Map position shown by Donis-Keller et al. (1987).
[b] All allele systems not completely characterized.

For some probes (e.g., CRI-C2), data from individual allelic systems have been combined to generate haplotypes, and the PIC and heterozygosity were calculated as the sum of the systems. Two probes, CRI-C84 (chromosome 2) and CRI-C86 (chromosome 12), revealed two different allelic systems, but only two haplotypes were seen in the CEPH families, indicating a high degree of linkage disequilibrium at each locus.

Figure 2 shows autoradiograms from hybridizations of two high-PIC probes, CRI-C44 and CRI-C96, to Southern blots containing CEPH family DNAs. CRI-C44 and CRI-C96 show heterozygosities >70% when used as probes in the CEPH families. CRI-C44 shows some lane background, which we ascribe to repetitive DNA in the probe, while CRI-C96 gives a strong, clear hybridization signal. Both of these probes have relatively complex fragment patterns, with alleles frequently consisting of many fragments. Pedigree analysis in these 3-generation families with a large number of children facilitated characterizing the alleles.

Linkage Analysis and Chromosome Assignment

Cosmid probes were assigned to chromosomes by linkage (lod $\geq$ 4.0) to existing linkage groups that in another study (Donis-Keller et al. 1987) had been assigned to chromosomes. All but one (CRI-C10) of the probes for which we collected genotypic data were linked to previously mapped probes and consequently could be assigned to chromosomes (see table 3). As noted in table 3, the chromosomal location of 15 of these probes has been published previously (Donis-Keller et al. 1987). Linkage of the additional eight cosmids to

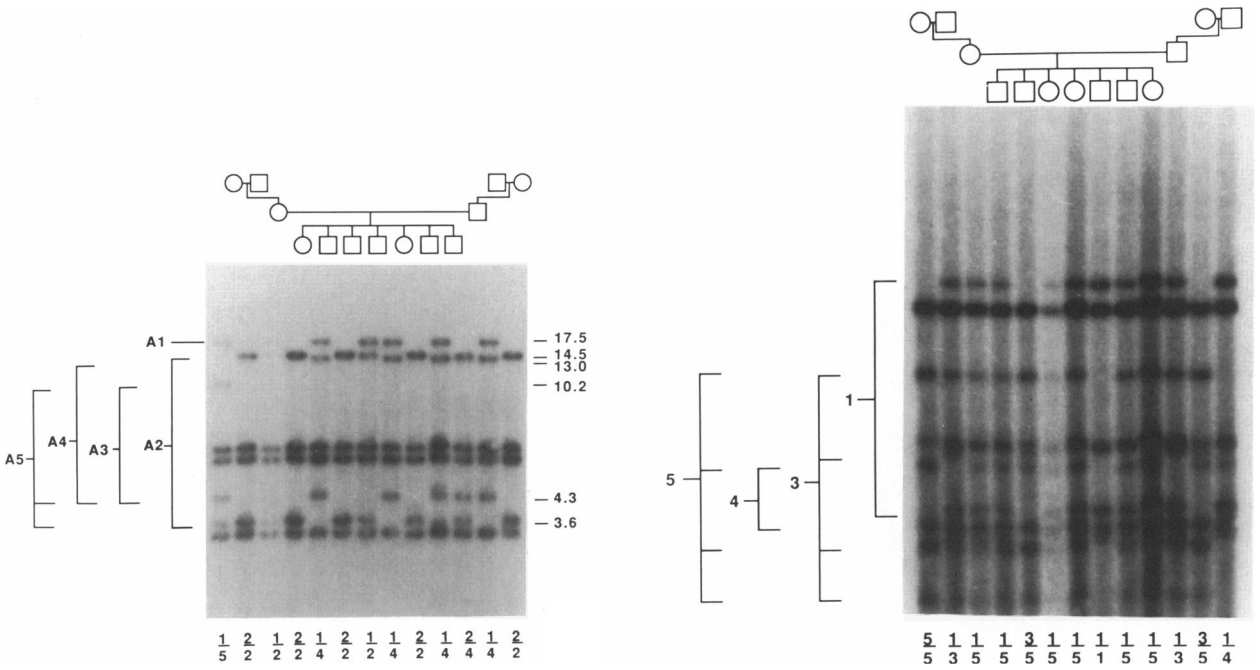

Figure 2 Autoradiograms of Southern blots containing CEPH-family DNA hybridized with cosmid probes. Allelic fragments are shown by brackets on the left of each panel. *A*, Hybridization of *Bgl*II-digested K1345 with nick-translated CRI-C96. Allele fragment sizes are designated on the right. *B*, Hybridization of *Msp*I-digested K13291 with CRI-C44.

previously mapped markers is demonstrated in table 4, which shows the two-point linkage with RFLPs that have been previously mapped. Probe CRI-C10 had two-point lod scores >2.0 with two linked probes on chromosome 11, a result that did not fulfill our criteria for linkage. The placement of CRI-C10 on chromosome 11 was confirmed by its pattern of hybridization to panels of human-hamster hybrid cell lines. Two of the 23 RFLP loci (CRI-C94 and CRI-C96) show pairwise linkage, with lod > 3. The RFLP loci lie on 10 different

Table 4

Two-Point Linkages

Cosmid	Chromosome	Assignment	Linked Probe[a]	LOD	Recombination Fraction
CRI-C5	X	Sex linkage	p43-15[b]	4.21	.00
CRI-C6	X	Sex linkage	CRI-R393	18.05	.00
CRI-C10	11	Hybrid panel			
	11	Linkage	CRI-L834	2.12	.08
			D11S16	2.11	.00
CRI-C15	14	Linkage	pAW101	5.55	.20
CRI-C88	X	Sex linkage	pDP34[c]	4.48	.12
CRI-C90	1	Linkage	CRI-R275	5.42	.00
CRI-C92	11	Linkage	CAT	4.25	.14
CRI-C94	8	Linkage	CRI-C96	4.60	.08

[a] Except where noted map position is shown by Donis-Keller et al. (1987).

[b] Locus DXS42 map position given by Drayna et al. (1985); segregation data derived from the CEPH data base.

[c] Locus DXYS1 map position given by Page et al. (1984); segregation data derived from the CEPH data base.

chromosomes: chromosome 2 (four loci), chromosome 5 and X (three loci each), chromosomes 4, 8, 11, 12, and 14, (two loci each), and chromosomes 1 and 3 (one locus for each).

Discussion

As part of our search for probes revealing DNA polymorphisms, we have evaluated the practicality of finding RFLPs when using randomly selected human genomic cosmid clones as probes. The use of cosmid clones to search for RFLPs was originally demonstrated by Litt and White (1984) in a study in which they screened 43 cosmid probes for *Msp*I and *Taq*I polymorphisms. Twelve of these cosmids revealed multiple *Msp*I or *Taq*I RFLPs. Here we have summarized our screening of 101 randomly selected cosmid probes for their ability to reveal RFLPs with nine restriction enzymes, characterized the RFLPs, and genetically mapped 23 of the RFLP loci. We have found this approach to be efficient and relatively easy to perform. Colonies picked at random can be used directly without prescreening, and "miniprep" lysates are satisfactory for preparing nick-translated probes. The presence of repetitive sequences is not a hindrance, since unlabeled genomic DNA blocks these sequences from hybridizing to filter-bound DNA (Litt and White 1985). In only two cases (CRI-C13 and CRI-C97) was it necessary to subclone fragments in order to identify allelic fragments. Therefore, cosmid DNA can usually be used as a probe. CRI-C13 appears to detect low-copy-number repeats on chromosome 2 and Y and possibly on other chromosomes (D. W. Bowden, unpublished data), a finding that accounts for its complex fragment pattern.

The utility of cosmids for screening larger segments of genomic DNA suggests that they would be more efficient than phage probes at revealing RFLPs. This appears to be true—but only partially so. In our study, 54% of the cosmids revealed polymorphism. Nakamura et al. (1987) screened for cosmids homologous to several oligonucleotide probes and found that 21% (77 of 372) of these probes revealed VNTR RFLPs and that 47% (174 of 372) revealed simple site polymorphisms. Schumm et al. (1988) found that 31% of 1,664 phage genomic clones revealed RFLPs. The number of allelic systems detected per probe by cosmid and phage clones was essentially the same (see table 2), however, indicating that phage clones had a greater probability of detecting multiple allele systems.

It is probable that many cosmid clones reveal polymorphisms that are obscured by constant fragments. We can infer this from experiments in which subclones from cosmids revealed polymorphisms that were not evident by hybridization of the parent cosmid to Southern blots (data not shown). Cosmids CRI-C87–CRI-C101 were selected from platings on host strains defective for the *Escherichia coli* methylcytosine restriction system (Raleigh and Wilson 1986; Raleigh et al. 1988), and 40% of these RFLPs had PIC values $\geq .3$ (table 3). This frequency is significantly higher ($P \leq .04$) as determined by χ^2 analysis than that for the cosmids isolated from platings on the host strain 1046, suggesting the possibility that different plating strains will generate a different spectrum or frequency of polymorphic probes.

It is interesting to note that, for our cosmid study, the efficiency with which enzymes revealed polymorphism was different than that seen in the results of Schumm et al. (1988) and Barker et al. (1984). In particular, *Pst*I RFLPs (our second most common type; table 2) were found at a significantly higher rate ($P \leq .002$) with cosmid probes and *Rsa*I RFLPs were found less frequently ($P \leq .015$), as determined by χ^2 analysis. The *Pst*I recognition site does not contain the CpG sequence that other studies had associated with high levels of polymorphism (Barker et al. 1984; Schumm et al. 1988). One reason for this might be the biased representation of sequences in our cosmid library (Bowden et al. 1988), probably introduced at the step of library construction. The source of this bias is unclear but may be due to selective elimination of methylated sequences (including CpG) present in the human genomic DNA by the modified cytosine restriction systems of the *E. coli* host (Raleigh and Wilson 1986; Raleigh et al. 1988). *Escherichia coli* 1046, the host of 86 of the 101 cosmids that we screened (see Material and Methods), apparently has an active *mcr* system.

The quality of the RFLPs revealed, as judged by the PIC/heterozygosity of the loci is also an important criterion for judging a screening system. Nineteen percent of the loci have PIC values $\geq .3$, which is comparable to the results obtained in the study by Schumm et al. (1988). Perhaps more important is the number of high-PIC (defined as PIC $\geq .7$) probes found. Schumm et al. (1988) found 17—or about 1 (1.7%) in 60—such probes in their survey of 1,025 phage prescreened to select for a population carrying single-copy genomic sequences, and they found 3 (0.47%) in 639 random clone probes (about 1 in 200). We did not find any probes (of 101 tested) with PIC $\geq .7$ (table 3) in our study. In a similar study, which screened a chromosome-specific cosmid library (T. Keith, personal communi-

cation), 3 (1.4%) of 211 cosmid probes isolated had PIC ≥ .7. Nakamura et al. (1987) do not provide a calculated PIC distribution, but at least 17 of 372 (1 in 22 [4.6%]) cosmids selected with oligonucleotide probes have heterozygosities ≥.70. Using the same measure (i.e., heterozygosities), we found 3 of 101 (1 in 33 [3.3%]) cosmids with heterozygosities ≥.70.

We accumulated genotypic data on 23 of the cosmids that reveal RFLPs and appear to have useful PIC values. The map positions of 15 of the loci revealed by these probes have been published (Donis-Keller et al. 1987). The RFLPs map to 10 different chromosomes and are distributed in a pattern that does not vary significantly when probes isolated from two different phage genomic libraries are used (Donis-Keller et al. 1987). Twenty-two of 23 probes were linked to other mapped loci under a criterion of lod ≥ 4.0. This demonstrates the resolving power of a human RFLP genetic map to rapidly assign probes to chromosomes and regions of chromosomes by pedigree analysis on CEPH families. We have found screening for polymorphism by using randomly selected cosmid probes to be simple and convenient. Results are comparable to those reported by Nakamura et al. (1987) and Schumm et al. (1988) in their large screening studies.

References

Barker, D., M. Schafer, and R. White. 1984. Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell **36**:131–138.

Bates, P. F., and R. A. Swift. 1983. Double *cos* site vectors: simplified cosmid cloning. Gene **26**:137–146.

Bowden, D. W., H. Muller-Kahle, T. R. Fulton, T. C. Gravius, D. F. Barker, and H. Donis-Keller. 1988. Studies on locus expansion, library representation, and chromosome walking using an efficient method to screen cosmid libraries. Gene **71**:391–400.

Donis-Keller, H., and 32 co-authors. 1987. A genetic linkage map of the human genome. Cell **51**:319–337.

Drayna, D., K. Davies, D. Hartley, J.-L. Mandel, G. Camerino, R. Williamson, and R. White. 1985. Genetic mapping of the human X chromosome by using restriction fragment length polymorphisms. Proc. Natl. Acad. Sci. USA **81**: 2836–2839.

Ish-Horowicz, D., and J. F. Burke. 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2998.

Lander, E. S., and P. Green. 1987. Construction of multilocus linkage maps in humans. Proc. Natl. Acad. Sci. USA **84**:2363–2367.

Litt, M., and R. White. 1985. A highly polymorphic locus in human DNA revealed by cosmid-derived probes. Proc. Natl. Acad. Sci. USA **82**:6206–6210.

Nakamura, Y., M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin, and R. White. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science **235**: 1616–1622.

Page, D. C., M. E. Harper, J. Love, and D. Botstein. 1984. Occurrence of a transposition from the X chromosome long arm to the Y chromosome short arm during human evolution. Nature **311**:119–123.

Raleigh, E. A., N. E. Murray, H. Revel, R. M. Blumenthal, D. Westaway, A. D. Reith, P. W. J. Rigby, J. Elhai. and D. Hanahan. 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. Nucleic Acids Res. **16**:1563–1575.

Raleigh, E. A., and G. Wilson. 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA **83**:9070–9074.

Schumm, J. W., R. G. Knowlton, J. C. Braman, D. F. Barker, D. Botstein, G. Akots, V. Brown, T. Gravius, C. Helms, K. Hsiao, K. Rediker, J. Thurston, and H. Donis-Keller. 1988. Identification of more than 500 restriction fragment length polymorphisms by screening random genomic clones. Am. J. Hum. Genet. **42**:143–159.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman, San Francisco.