Volume 15 Number 11 1987 Nucleic Acids Research

Identification of 28 DNA fragments that detect RFLPs in 13 distinct physical regions of the short arm of chromosome 5

Joan Overhauser, Jeannette McMahan and John J.Wasmuth

Department of Biological Chemistry, California College of Medicine, University of California (Irvine), Irvine, CA 92717, USA

Received March 2, 1987; Revised and Accepted May 15, 1987

ABSTRACT

A series of 175 lambdaphage carrying human inserts isolated from a library that is specific for the short arm of human chromosome 5 (5p) have been regionally mapped on 5p using a deletion mapping panel of 16 human-hamster cell hybrids, each of which contains a chromosome 5 with a different deletion in the short arm. Seventy-five single copy DNA fragments were screened with 12 restriction enzymes for their ability to detect restriction fragment length polymorphisms (RFLPs). Twenty-eight of these DNA fragments, which are located in 13 distinct physical regions of 5p, were found to detect RFLPs. These DNA markers make it possible to construct a linkage map that will span the entire length of 5p and will allow the relationship between genetic and physical distance for this region of the genome to be examined at a high level of resolution.

INTRODUCTION

The application of recombinant DNA techniques and restriction enzyme analysis to detect DNA sequence variation in the human genome in the form of restriction fragment length polymorphism (RFLP) provides a means to identify an almost unlimited supply of informative genetic markers throughout the human genome (1). Using DNA markers, it should be possible to construct a linkage map spanning the entire human genome in the forseeable future. The development of a complete linkage map of the human genome will allow a systematic approach to identifying genetic markers linked to virtually any locus responsible for an inherited disease. This could facilitate prenatal diagnosis and carrier detection for hundreds of monogenic disorders of unknown etiology and allow the disease genes to be localized to specific regions of the genome. The ability to focus on specific regions of the genome, especially ones currently barren for polymorphic loci, and saturate them with informative genetic markers that span the regions will undoubtedly hasten the completion of a complete human genetic map. The short arm of human chromosome 5 (5p) is one region where relatively few genetic markers, and no genes of known function, have been mapped.

Using unique sequence DNA fragments from a library specific for 5p (2) in conjunction with a panel of 16 cell hybrids, each of which retains a chromosome 5 missing a different portion of 5p, we have localized 175 DNA fragments to 17 distinct physical region of 5p. Seventy-five of these DNA fragments were screened to identify ones that detect RFLPs and 28, located in 13 different regions of 5p, were found to reveal polymorphic loci.

In addition to providing the means to construct a linkage map that spans the short arm of chromosome 5, these DNA markers, together with the physical map of 5p, will enable us to examine the relationship between genetic and physical distance over a large region. Information about the variability in the physical-genetic distance relationship for as much of the genome as possible is especially important for the use of the "reverse genetics" strategy to isolate genes involved in heritable disorders of unknown etiology where assumptions must be made concerning what recombination frequency means in terms of Mb pairs of DNA.

MATERIALS AND METHODS

Cell Hybrids and Cell Lines

UCW56 is a Chinese hamster ovary cell line which has a mutation in the LARS gene which renders leucyl-tRNA synthetase thermolabile and the cell line nonviable at 39°C. The human gene that can complement the temperature-sensitive phenotype of UCW56 is located on the long arm of chromosome 5 (3). UCW56 cells were fused to fresh peripheral leukocytes or lymphoblastoid cells from individuals with deletions or other rearrangements of 5p. Cell hybrids that retained a human chromosome 5 were selected for their ability to grow at 39°C and were then screened for the retention of an abnormal chromosome 5 as previously described (4). Most of the cell hybrids used for the present study and the rearranged chromosomes 5 they retain have been previously described (5) except for the following: HHW788 was derived from an individual carrying a balanced translocation t(5;12)(p13;q24) and retains the derivative chromosome 5 (5qter-5p13::12q24-12qter). HHW792 contains the derivative chromosome 5 from an individual with the cri du chat syndrome who has a terminal rearrangement [45XX, ter rea (5;13)(5qter-cen-5p15.3::13p12-13qter)] (Dr. K.-L. Ying, Childrens Hospital of Los Angeles). HHW848 was derived from an individual with a reciprocal translocation [t(4,5)(q21;p15.3)] and contains the derivative chromosome 5 (5qter-5p15.3::4q21-4qter). The family in which this translocation is segregating has been described previously (6). HHW909

has the deleted chromosome 5 (5qter-5p15.2:) from an individual with the cri du chat syndrome (Dr. Eric Niebuhr, Institute of Medical Genetics, University of Copenhagen). Figure 1 depicts the regions of 5p that are missing from each of the abnormal chromosomes 5 isolated in the various cell hybrids.

Isolation of Low Copy DNA Fragments from 5p

HHW213 is a human-Chinese hamster cell hybrid in which the only human DNA present is a chromosome 5 with an intact p arm but which is missing about 95% of the q arm (3). The preparation of a genomic DNA library from this cell hybrid in the lambdaphage vector EMBL-4, the isolation of recombinant phage with human DNA inserts, the preparation of phage DNA and the isolation of unique copy DNA fragments have all been previously described (2,4,5).

Southern Blot Hybridization

High molecular weight DNA extracted from lymphoblastoid cells or human-hamster cell hybrids was digested to completion with different restriction enzymes and subjected to electrophoresis through 0.8% agarose gels. Gels were soaked in .25N HCl for 7 min., rinsed with H_2O, then placed on top of 3MM Whatman paper that was saturated with .5N NaOH. Nylon membranes (Genetran 45, Plasco Inc., Woburn, MA) that had been prewetted in H_2O were put on top of the gels and 3MM Whatman paper, paper towels and a weight were placed on top. Two-twelve hours later the membranes were removed and rinsed in 3X SSC before use.

DNA fragments were labeled by the random primer method according to Feinberg and Vogelstein (6,7). Normally, 30-90 ng of each DNA fragment was labeled to a specific activity of $\sim 10^8$ cpm/µg for each Southern blot hybridization.

Membranes were prehybridized for 2-4 hours at 65°C in hybridization solution consisting of 10% PEG, 7% SDS, 1% BSA, .25M NaPhosphate pH7.2, .25M NaCl, 1mM EDTA. The membranes were then placed in a seal-a-meal bag with 10 ml hybridization solution, 250 µg sonicated denatured carrier DNA and 10^7 cpm of radiolabeled probe. The membranes were incubated at 65°C for 16-24 hours then washed in 2X SSC for 1/2 hour at 65°C and .5X SSC for 1/2 hour at 65°C before autoradiography.

Screening for Restriction Length Polymorphisms

Lymphoblastoid cultures from members of the six extended family Utah pedigrees were obtained from the Human Genetic Mutant Cell Repository (Cambden, N.J.). Each probe tested was hybridized to filters containing DNA from six unrelated individuals (3 sets of parents) which had each been

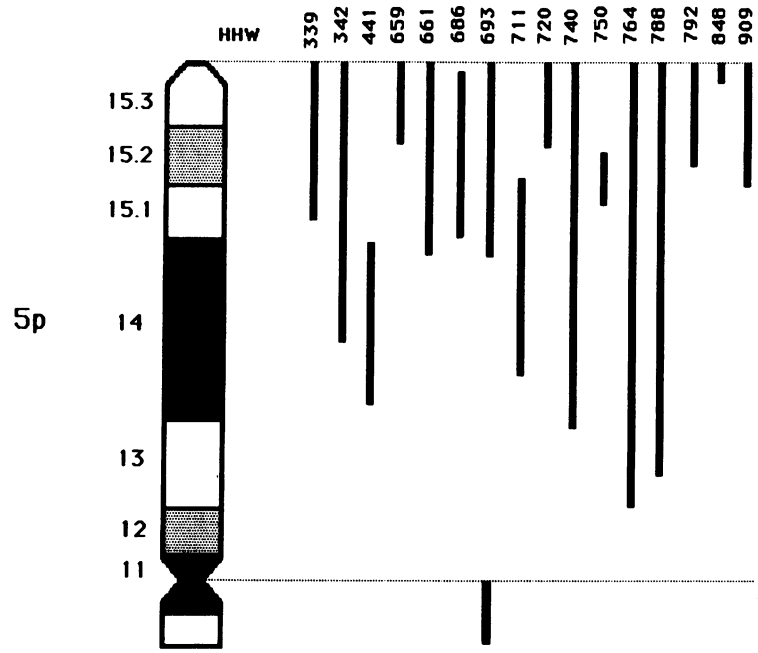

Figure 1. Location of deletions of the human chromosomes 5 present in human-hamster cell hybrids. An ideogram depicting the portion of chromosome 5 present in HHW213 is shown on the left. The exent of the deletions of the chromosomes 5 present in the other human-hamster hybrids are indicated by the black bars. All of the cell hybrids with the exception of HHW213 and HHW693 contain an intact long arm of chromosome 5.

digested with 12 different restriction endonucleases. Identification of an RFLP resulted in the screening of an additional set of six unrelated individuals.

RESULTS

Isolation and Physical Mapping of Single Copy DNA Fragments from 5p

Over 300 low or unique copy DNA fragments were subcloned from the 175 independent lambdaphage carrying human inserts from the HHW213 library. These DNA fragments were screened by Southern blot hybridization against the hybrid cell mapping panel depicted in Figure 1 in order to localize the fragments to specific regions of 5p (or the proximal 5-10% of 5q) and to determine which were completely free of repetitive sequences. Figure 2 shows autoradiograms of blot hybridizations using four different probes located in different regions of 5p. Based on the chromosome breakpoints of the chromosomes 5 in

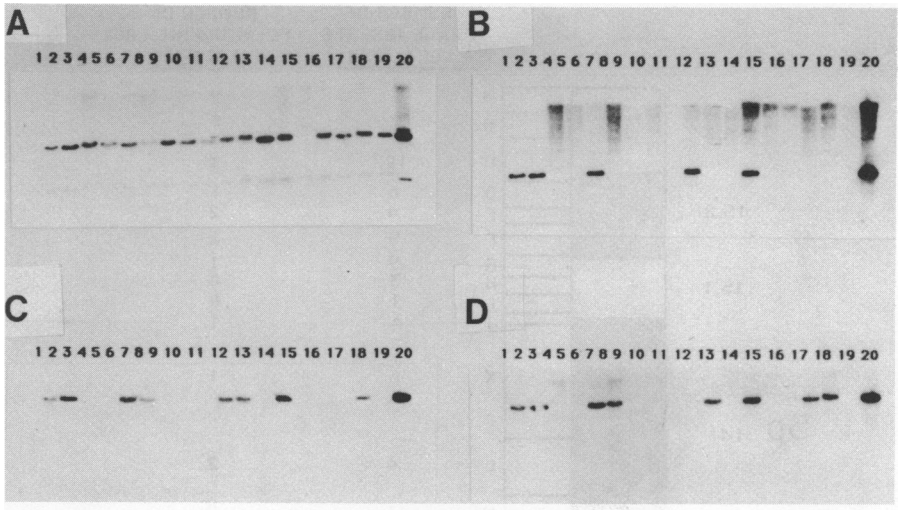

Figure 2. Southern blot analysis using the hybrid cell deletion panels. Genomic DNA was digested with HindIII. Lanes 1-20 are as follows: CHO, HHW105, HHW213, HHW339, HHW342, HHW441, HHW659, HHW661, HHW686, HHW693, HHW711, HHW720, HHW740, HHW750, HHW764, HHW788, HHW792, HHW848, HHW909, human. Panels A-D were hybridized with the following DNA fragments, A-J028E-B; B-J0125H-B; C-J0209E-D; D-J0D194H-A/B.

the different cell hybrids and the presence or absence of hybridization signals from each probe to DNA from each cell hybrid, the probes could be localized to 17 distinct physical regions of 5p, as shown in Figure 3. The resolution of the physical map of 5p, especially above 5p14, is considerably better than one we previously reported (5) which is primarily due to the addition to the mapping panel of three new cell hybrids with different breakpoints in 5p15. It should be noted that relatively few of the DNA fragments above 5cen (16/145 or ~ 11%) are located within 5p14 even though this region encompasses about 30% of the cytogenetic length of the short arm of chromosome 5.

RFLP Analysis

DNA fragments that gave signals indicative of single copy sequences upon hybridization to the deletion panels were used to screen for RFLPs. To help ensure that polymorphic loci spanning the entire length of 5p were identified, several probes (if more than one was available) from each of the 17 different physical regions were examined. Seventy-five DNA fragments were initially screened with genomic DNA from 6 unrelated individuals that had each been digested with 12 different restriction enzymes (BglII, EcoRI, HincII, HindIII,

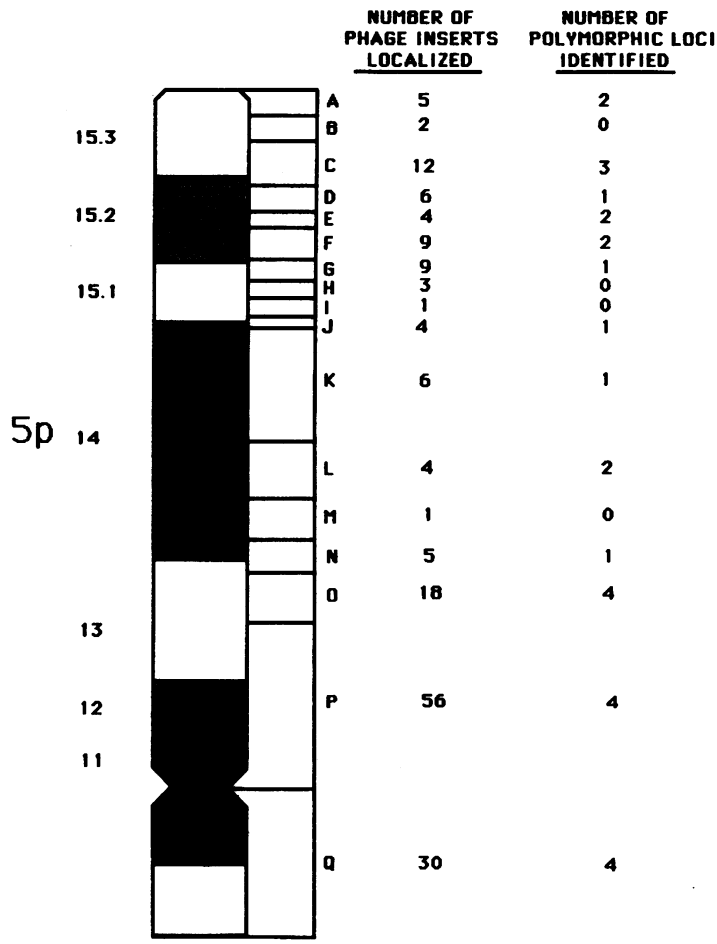

Figure 3. Physical location of DNA fragments and polymorphic loci on 5p. An ideogram of the relevant portion of chromosome 5 is shown. The letters A-Q refer to the 17 distinct regions that can be distinguished by hybridization patterns of DNA fragments. The number of phage inserts and RFLPs that have been assigned to each region is indicated on the right.

MboI, MspI, StuI, TaqI, XbaI, RsaI, SphI, SstI). Any DNA fragment that detected an RFLP in this set of individuals was rescreened with DNA from 6 additional unrelated individuals. The polymorphisms revealed by three probes are shown in Figure 4. For each polymorphism detected, Mendelian inheritance was confirmed by tracing the segregation of alleles in one or more informative nuclear families. Examples of the segregation of alleles for two of the polymorphic loci are shown in Figure 5.

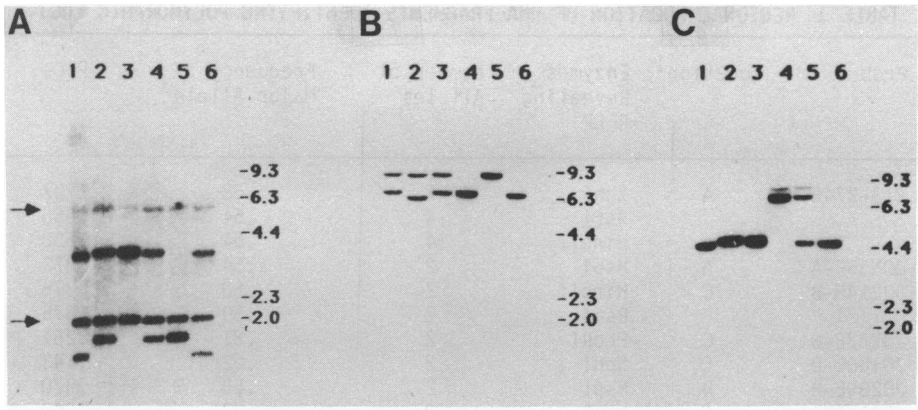

Figure 4. Polymorphisms revealed by three DNA fragments. Genomic DNA from 6 unrelated individuals was digested with the appropriate restriction enzymes and subjected to blot hybridization. A, DNAs digested with SphI and hybridized with 213-274E-C. (Arrows indicate invariant bands.) B, DNAs digested with MspI and hybridized with DNA fragment J0205E-D. C, DNAs digested with MspI and hybridized with DNA fragment J0110H-C. The numbers to the right of each panel indicate the position of migration of size markers.

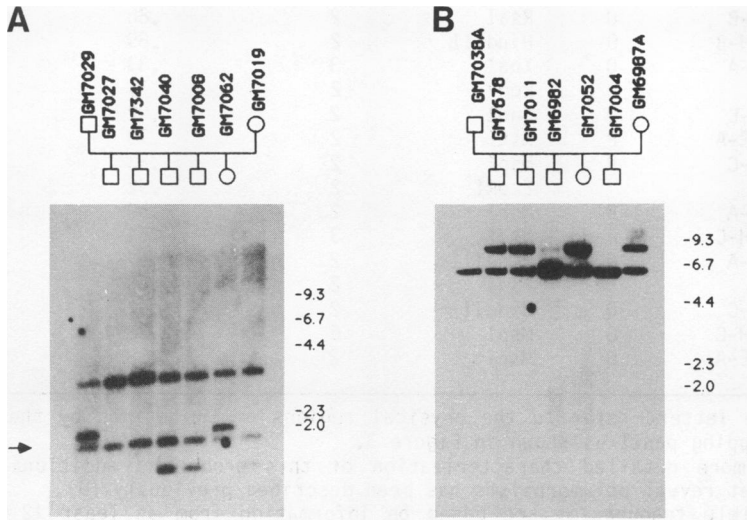

Figure 5. Mendelian inheritance of RFLPs. A, DNAs from the family shown were digested with SphI and hybridized with DNA fragment 213-274E-C. (Arrows indicate invariant bands.) B, DNAs from the family shown were digested with MspI and hybridized with DNA fragment J0205E-D. The numbers to the right of each panel indicate the position of migration of size markers.

TABLE 1 REGIONAL LOCATION OF DNA FRAGMENTS IDENTIFYING POLYMORPHIC LOCI

Probe	Location[1]	Enzymes Revealing RFLP	Number of Alleles	Frequency of Major Allele[3]	PIC
213-274E-C[2]	A	SphI	>7	.33	.707
		SstI	4	.54	.558
		HincII	4	.54	.558
J0N35E-A	A	MspI	2	.54	.372
J0214H-B	C	HincII	2	.50	.375
		RsaI	2	.50	.375
J0183E-B	C	EcoRI	2	.81	.261
J0180E-D	C	SphI	2	.92	.143
J0209E-B	D	MspI	2	.58	.370
J0D194H-A/B	E	MspI	2	.83	.240
J094E-C	E	RsaI	2	.58	.370
J0120H-B	F	MspI	2	.71	.329
J023H-B	F	SphI	2	.96	.075
J056E-F	G	XbaI	3	.58	.462
J099H-C	J	SstI	2	.58	.370
J081H-A	K	TaqI	2	.96	.075
		SphI	2	.96	.075
J0109E-C	L	EcoRI	2	.55	.372
J044E-B	L	EcoRI	2	.58	.370
		MspI	2	.63	.361
J0N44H-D	N	HincII	2	.75	.305
J028E-B	O	RsaI	2	.88	.195
J0162H-B	O	HindIII	2	.89	.184
J071H-A	O	XbaI	3	.33	.603
		EcoRI	2	.71	.329
J018E-C	O	SphI	2	.92	.143
J0N22E-A	P	SstI	2	.79	.277
J092H-C	P	MspI	2	.83	.240
		EcoRI	2	.92	.143
J035H-A	P	MspI	2	.50	.375
J0110H-C	P	MspI	3	.54	.429
J030H-A	Q	HindIII	2	.67	.342
		RsaI	2	.83	.240
J016E-C	Q	HincII	2	.92	.143
J0205H-C	Q	MspI	6	.33	.708
J0157E-A	Q	MspI	2	.58	.370

1. The letters refer to the physical regions of 5p defined by the deletion mapping panel as shown in Figure 3.
2. A more detailed characterization of this probe and additional enzymes that reveal polymorphisms has been described previously (9).
3. Allele frequencies are based on information from at least 12 unrelated individuals.

Of the 75 probes examined, 28 revealed RFLPs with one or more enzymes. Table 1 contains a characterization of the 28 DNA fragments and the polymorphisms they reveal. As depicted in Figure 3, the polymorphic loci span

the entire length of the short arm of chromosome 5. The polymorphisms revealed by five fragments (213-274E-C, J081H-A, J0214H-B, J0209E-B, and J0180E-D) were described previously (5). The polymorphism information content (PIC) of 19 of the 28 loci defined by these probes is greater than 0.3 indicating they will be generally useful genetic markers. Those probes that revealed infrequent polymorphisms with the 12 enzyme initially tested are being rescreened with an additional 10 enzymes in an effort to increase their informativeness. Two probes in particular revealed a high degree of polymorphism, J0205H-C, which reveals 6 different alleles with MspI and 213-274E-C which, as described previously, reveals a highly polymorphic locus near the terminus of 5p (9). This probe reveals RFLPs with at least 11 different enzymes and appears to be detecting rearrangements associated with a hypervariable sequence (10,11). Polymorphic loci have now been identified in 13 of the 17 different physical regions of 5p defined by the deletion mapping panel, as shown in Figure 3 and Table 1.

DISCUSSION

Here we describe an approach for identifying a large number of DNA fragments that detect polymorphic loci which have been accurately localized to specific regions of a chromosome. Using this approach, 28 polymorphic loci have been localized to 13 different regions of the short arm of chromosome 5, each region defined by a distinct hybridization pattern to the deletion mapping panel. As shown in Table 1 and Figure 3, this approach resulted in the identification of RFLPs spanning the entire length of the short arm of chromosome 5. This number of polymorphic loci, together with a knowledge of their physical location, should provide the means to compile a complete linkage map for this region of the genome. The ability to accurately physically map many DNA fragments will allow us to concentrate on screening fragments from specific regions where few or no polymorphic loci have been found in case gaps in a linkage map of 5p are encountered.

The accurate physical mapping of many polymorphic loci to specific regions of 5p will also enable us to address two basic questions related to physical vs. genetic distances. 1. Is the relationship between cytogenetic distance and DNA content uniform over the length of a chromosome? That is, does the DNA content per unit of cytogenetic distance vary drastically from one defined cytogenetic region to another? 2. How uniform is the relationship between physical distances (measured in Mb pairs of DNA) and genetic distances over the length of a large chromosomal region? With respect

to the first question, the distribution of the 145 DNA fragments mapped on 5p is quite interesting. If one divides 5p into three regions of approximately the same cytogenetic length, 38% of the inserts (55/145) are located in 5pter-5p15 (A-J in Figure 3, ~35% of 5p); 51% of the inserts are located in 5p13-5cen (O-P in Figure 3, ~35% of 5p) but only 11% (16/45) are located within 5p14 (K-N in Figure 3) even though this region represents about 30% of the cytogenetic length of 5p. The disparity between the cytogenetic size of 5p14 and the small proportion of inserts found to be in this region could be due to any one of several factors including: 1. The method used to isolate phage with human DNA inserts (selection for the presence of human repetitive sequences). Thus, repetitive sequences may be sparse within 5p14. 2. There is an underrepresentation of DNA fragments from 5p14 in the library, possibly due to the presence of a high proportion of sequences that do not replicate well in the E. coli host strain. 3. The region 5p14 may in fact contain much less DNA than would be expected based on its cytogenetic size. Using the high resolution probe map of 5p and pulsed-field gel electrophoresis for long range restriction mapping, we can examine this latter possibility directly. In fact, it is now feasible to construct a complete long range restriction map using rare cutting restriction enzymes and pulsed-field electrophoresis for the entire length of 5p. This will allow an examination of the relationship between cytogenetic length and Mb pairs of DNA for this whole region. We are beginning the long range restriction mapping, initially concentrating on the region 5p15-5pter, where we have the physical resolution and because this region contains the gene or genes responsible for the phenotype of the 5p deletion syndrome, cri du chat (5). In addition, in collaboration with Dr. K. Kidd, we are using the RFLPs to compile a linkage map of 5p. Development of both a linkage map and a long range restriction map will allow us to investigate the relationship between physical distance (measured in both cytogenetic length and Mb pairs of DNA) and the genetic distance over the length of 5p at a very high level of resolution.

ACKNOWLEDGEMENTS

The excellent technical assistance of Ulla Bengtsson is gratefully acknowledged. We thank R. Apostel, Vivegen, Orange, CA; Dr. K.-L. Ying, Children's Hospital, Los Angeles, CA; Dr. M.R. Hayden, Health Sciences Centre Hospital, Univ. British Columbia, Vancouver, Canada; and Dr. E. Niebuhr, Institute of Medical Genetics, Univ. Copenhagen, Denmark for their assistance in providing blood samples or lymphoblastoid cultures from individuals with

rearrangements of 5p. This work was supported by U.S. Public Health Service grant HD 18642. J.O. was supported by NIH postdoctoral fellowship HD06892-01A1.

REFERENCES

1. Botstein, D., White, R.L., Skolnick, M., and Davis, R. (1980) Am. J. Hum. Genet. 32, 314-331.
2. Carlock, L.R., and Wasmuth, J.J. (1985) Somat. Cell. Mol. Genet. 11, 267-276.
3. Dana, S., and Wasmuth, J.J. (1982) Somat. Cell. Mol. Genet. 8, 245-264.
4. Overhauser, J., Golbus, M.S., Schonberg, S. A., and Wasmuth, J.J. (1986) Am. J. Hum. Genet. 39, 1-10.
5. Overhauser, J., Beaudet, A.L., and Wasmuth, J.J. (1986) Am. J. Hum. Genet. 39, 562-572.
6. Froster-Iskenius, U.G., Hayden, M.R., Wang, H.S., Kalousek, D.K., Horsman, D., Pfeiffer, R.A., Schottky, A., and Schwinger, E. (1986) Am. J. Hum. Genet. 38, 759-767.
7. Feinberg, A.P., Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
8. Feinberg, A.P., Vogelstein, B. (1984) Anal. Biochem. 137, 266-267.
9. Overhauser, J., Beaudet, A.L., and Wasmuth, J.J. (1987) Nucl. Acid. Res. (in press).
10. Wyman, A.R., and White, R. (1980) Proc. Natl. Acad. Sci. USA 77, 6754.
11. Migone, N., Feder, J., Cann, H., van West, B., Hwang, J., Takahashi, N., Honjo, N., Piazza, A., and Cavalli-Sforza, L.L. (1983) Proc. Natl. Acad. Sci. USA 80, 467.