Volume 15 Number 4 1987 Nucleic Acids Research

Isolation of polymorphic DNA fragments from human chromosome 4

T.Conrad Gilliam, Steven T.Healey, Marcy E.MacDonald, Gordon D.Stewart, John J.Wasmuth[1], Rudolph E.Tanzi, Jerry C.Roy and James F.Gusella*

Neurogenetics Laboratory, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston MA 02114 and [1]Department of Biological Chemistry, University of California, Irvine, CA 92717, USA

Received December 1, 1986; Accepted January 9, 1987

ABSTRACT

We have identified and characterized 40 DNA probes detecting restriction fragment length polymorphism (RFLP) on human chromosome 4. Single copy human clones were isolated from a bacteriophage library enriched for chromosome 4 sequences. Each clone was hybridized to somatic cell hybrid DNAs for verification of its species and chromosomal origin and for regional localization. Sequences specific for chromosome 4 were tested for their ability to detect RFLPs in human DNA and their potential utility as genetic markers was assessed. Approximately 263,000 base pairs or 0.13 % of the chromosome was screened for sequence variation. The estimate of heterozygosity calculated from this large body of data, H = 0.0021, indicates that the degree of sequence variation on chromosome 4 is comparable to other autosomes. The characterization of these 40 markers has tripled the number of polymorphic loci available for linkage studies on chromosome 4, making it feasible to begin construction of a detailed linkage map that will span the entire chromosome.

INTRODUCTION

The development of recombinant DNA technology has heralded a new era in human genetics. Application of molecular techniques permits the direct detection of DNA sequence variation as restriction fragment length polymorphism (RFLP) (1,2). This has provided an abundant supply of highly informative genetic markers throughout the genome. The defects in a number of autosomal inherited disorders including Huntington's disease (3,4), polycystic kidney disease (5) and cystic fibrosis (6-8) have recently been localized by linkage to polymorphic DNA loci. RFLP markers offer the promise of a complete linkage map of the human genome that will permit a systematic approach to mapping the several thousand monogenic disorders.

In 1983, we reported the assignment of Huntington's disease (HD) to chromosome 4 by linkage to the DNA marker locus D4S10 detected by probe G8 (3). This provided an impetus to generate additional RFLP markers for chromosome 4, some of which should be in the vicinity of the defect. Our search has resulted in the characterization of 40 polymorphic loci detected by

© IRL Press Limited, Oxford, England.

DNA probes scattered on both the long and short arms of this autosome. In combination with the "expressed" polymorphisms GC and MN, a few anonymous DNA probes and several cloned genes that detect RFLPs (9,10), these new DNA markers virtually ensure the construction of a complete linkage map that will permit efficient investigation of any inherited disorders mapping to this autosome.

MATERIALS AND METHODS

Cell Lines

DNA was prepared from cell hybrid lines for assignment of loci to chromosome 4 and regional localization as described (11). Three different cell hybrids containing only chromosome 4 were used. W4-3A/SB5 supplied by Dr H. Willard and HHW416 (constructed by John Wasmuth) are subclones of two different human-hamster somatic cell hybrid lines that have segregated all other human chromosomes except chromosome 4. JS4/A9.1 supplied by Dr E. Stanbridge (Irvine) has a human chromosome 4 with an integrated selectable marker on a mouse A9 background (12). Probes were regionally localized using two additional cell hybrids, CF26/20, a human-hamster somatic cell hybrid containing the region 4pter-q11 as part of a 4/15 translocation chromosome (13) and HHW693, a human-hamster hybrid containing 4pter-p15.1 as part of a 4/5 translocation chromosome (14). Family DNAs were prepared using lymphoblastoid cell lines from members of nuclear families within a large Venezuela Huntington's disease kindred (3).

Chromosome 4 Library

The library used in this study was prepared at the Lawrence Livermore National Laboratory from chromosome 4 DNA obtained by flow-sorting of chromosomes from a human-hamster somatic cell hybrid line containing human chromosomes 4, 8, and 21. DNA from the chromosome 4 peak was digested to completion with HindIII and cloned in the bacteriophage vector Charon 21A. After packaging, 23,000 independent recombinants were obtained. The library was amplified and stored in phage dilution buffer over chloroform.

Probe Preparation

DNA was prepared from single phage plaques using a conventional plate lysate miniprep procedure (15). Approximately 2 ug of DNA was digested with HindIII and the insert restriction fragment was separated electrophoretically on low gelling temperature agarose (Seaplaque, FMC Bioproducts). After ethidium bromide staining, the insert fragment was excised as an agarose block and either stored or used immediately as template for synthesis of 32P

labelled probe by the oligonucleotide priming procedure of Feinberg and Vogelstein (16).

Southern Transfer and Hybridization

Transfer of phage plaques to nitrocellulose filters and hybridization to human or hamster genomic DNA were carried out as previously described (17). Southern blotting and hybridization were carried out as described using Genatran nylon membrane (Plasco Inc., Woburn MA) (3). When necessary, blots were stripped by washing for 2 hours in distilled water at 65C and rehybridized with different probes up to 10-15 times.

RESULTS

Isolation of Phage Containing Chromosome 4 DNA

The probes used in our search for RFLP were selected from a chromosome 4 enriched library constructed at the Lawrence Livermore National Laboratories by flow-cytometry of intact chromosomes derived from a human-hamster hybrid cell line containing human chromosomes 4, 8 and 21. The library consisted of HindIII fragments from the peak fraction of human chromosome 4-containing material cloned in the vector Charon 21A. Our efforts provide a characterization of the high standard and utility of this library which is generally available to interested investigators through the National Laboratory Gene Library Project. A summary of these results is provided in Figure 1.

We first chose 2315 phage clones at random and placed them in gridded arrays. Duplicate nitrocellulose replicas of these were hybridized to 32P labeled human or hamster genomic DNA to eliminate those clones containing highly repeated sequences. Phage clones that hybridized most intensely to the human DNA contain repeat sequences of human origin while those hybridizing primarily to hamster DNA contain hamster repetitive elements (17). The two occurred in a ratio of approximately 3.3:1 indicating that inserts in the majority of phage in the library would be DNA segments of human origin. The 1111 (48%) clones that failed to hybridize detectably to either probe were chosen for further characterization.

Miniprep DNA from each phage clone was isolated, digested with HindIII and fractionated by agarose gel electrophoresis. Ethidium bromide staining material obscured fragments up to 400 bp preventing an assessment of how many phage did not have an insert. Only phage containing inserts greater than 600 bp were chosen for probe preparation. The inserts were excised from low-melt agarose gels, labeled and used to probe genomic blots similar to those in

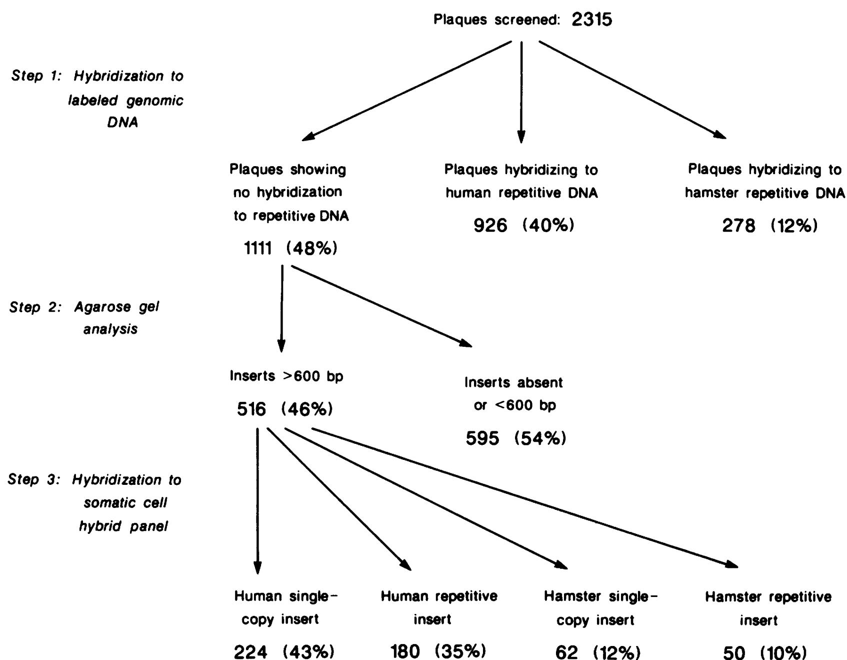

Figure 1-- Flow Chart of Chromosome 4 Library Characterization

Figure 2. Lanes of HindIII digested human and hamster genomic DNA were included on the blots to distinguish the species of origin of the probes, and to monitor whether they were entirely free of repetitive elements. Other lanes on the gels contained HindIII digested DNAs from selected somatic cell hybrid lines to verify the chromosome 4 origin and to assign a regional location to each probe. The analysis in Figure 2 assigns probe 16MG7 to the region 4pter-p15.1 and demonstrates cross-hybridization of the sequence to two hamster fragments.

Some of the phage clones yielded more than one HindIII fragment. The vast majority of these represented contiguous inserts presumably resulting from partial digestion of the genomic DNA at the initial cloning step. Less than 4% of the phage had multiple inserts than were noncontiguous in the genome. 79% of the phage characterized contained human segments, but approximately half of these contained repetitive sequences not detected in our initial screen. Ultimately, about 10% of the original phage from the library yielded single copy human inserts that were large enough to be used effectively as probes on genomic blots. All of these detected an appropriate sized HindIII fragment

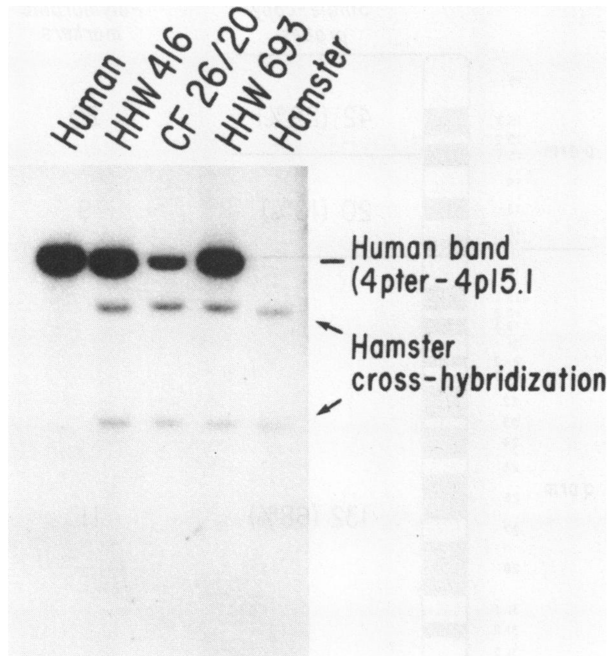

Figure 2-- Regional Localization of Chromosome 4 Probes

A typical mapping blot for regional mapping on chromosome 4 is shown. Each lane contains 5 ug of HindIII-digested genomic DNA that had been fractionated by agarose gel electrophoresis, transferred to nylon filter and hybridized to probe 16MG7. Hybrid cell lines containing chromosome 4 or portions thereof are described in Materials and Methods. HHW416 contains only human chromosome 4, CF26/20 contains 4pter-q11 and HHW693 has the region 4pter-p15.1. Human DNA was isolated from normal lymphoblast lines while hamster DNA was prepared from cell lines Wg3H and UCW3, the parent lines of the somatic cell hybrids. The probe 16MG7 detects a human segment assigned to the region 4pter-p15.1 but also displays crosshybridization with two hamster fragments.

mapping to human chromosome 4 when they were used to probe filters similar to that used in Figure 2. In a few instances, the probe also revealed additional human fragments not mapping to chromosome 4. For several probes, crosshybridization to hamster sequences was also detected indicating that they may contain conserved coding sequences.

In some cases, it was not possible to regionally localize the human segment on chromosome 4 because of weak signal in one of the hybrid cell line DNAs, particularly CF26/20. Of the 194 DNA segments that were regionally assigned (Figure 3), the proportion mapping to the long arm (68%, 4q11-qter),

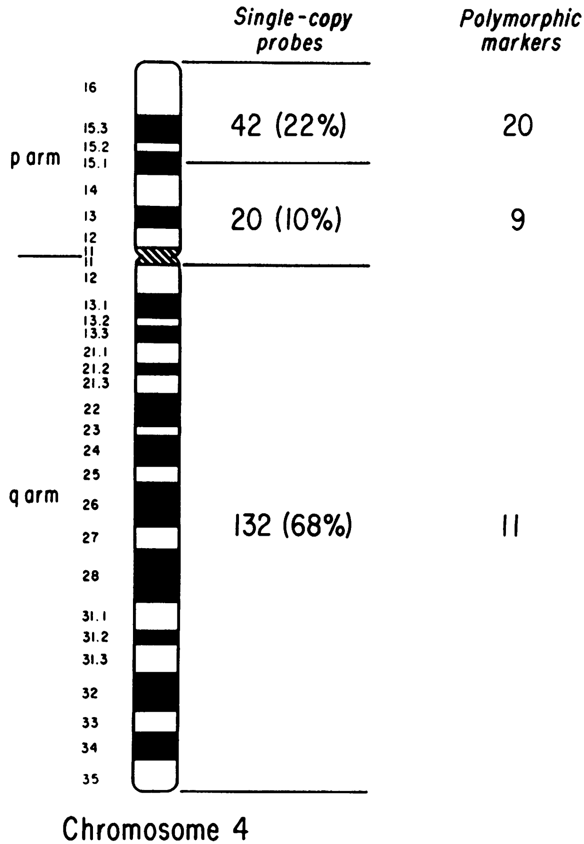

Figure 3-- Regional Distribution of Chromosome 4 Probes.
A summary of the total number of randomly chosen probes and probes detecting RFLPs mapping to each of the three regions defined by our panel of somatic cell hybrid lines is shown. The number of probes detecting RFLPs is highest in the telomeric portion of the short arm since more of these were included in the screen.

and short arm (32%, 4pter-q11) is similar to that expected (72% and 28%, respectively) based on cytogenetic estimates of the physical sizes of these chromosome segments. On the short arm, however, the clones did not distribute equally to the 4pter-p15.1 and 4p15.1-q11 regions. Instead, 68%, a significantly higher proportion than expected, mapped above the breakpoint of the HHW693 hybrid in the pter-p15.1 segment. There are several potential explanations for this significant deviation ($P<0.01$), including nonrandomness in the library or in the distribution of HindIII fragments on the chromosome, the inherent difficulty of assigning exact breakpoints in cytogenetic analysis

and the possibility that the pericentromeric region contains a lower relative content of single copy sequence.

Screening for DNA Polymorphisms

Our screen for polymorphism consisted of hybridizing each probe to genomic DNA from 5 unrelated individuals digested with each of 33 different restriction enzymes. Aldridge et al. (18) have suggested that screening this small number of individuals favors the identification of frequent RFLPs that are most informative for family studies. Our choice of restriction enzymes includes most of those proposed previously to detect frequent variation in addition to several other enzymes producing a range of fragment sizes which can be easily resolved by agarose gel electrophoresis. We biased our search for RFLPs to probes on the short arm of the chromosome since these might be near the HD gene (19-21), but several long arm probes were also tested. Of 57 probes screened, 50 displayed variation with at least one enzyme and many of these revealed multiple variants with several different enzymes.

We have further characterized approximately 75% of the variants detected in the initial screen. Our choices were based on intensity of hybridization signal, ease of resolution of the variant fragments under our standard gel conditions and cost and frequency of use of the particular restriction enzyme within the lab. Table 1 summarizes the results of this characterization. For each putative RFLP reported, segregation of the alleles was traced in several nuclear families in order to confirm Mendelian transmission. A typical result is shown in Figure 4 which displays the pattern of fragments observed with probe B3D hybridized to Southern blots containing PstI-digested DNA from a typical nuclear family. One parent is heterozygous and the other homozygous for the marker. The phenotype observed for each child is consistent with Mendelian inheritance. RFLPs were typed in several families of this type to calculate the odds in favor of Mendelian inheritance versus random distribution of the different sized fragments as suggested by Barker et al. (22). Mendelian segregation was overwhelmingly favored (>100,000:1) in every instance.

Characterization of Chromosome 4 Markers

Polymorphism information content (PIC) has been proposed as a useful measure of the quality of DNA marker loci for genetic linkage studies (1). It is a function of the number and individual frequency of alleles at a marker locus. An expanded panel of 34 unrelated individuals was used to assess the frequency of each restriction fragment variant in the general North American population. Ten of the 50 DNA markers that detected RFLPs possessed PIC

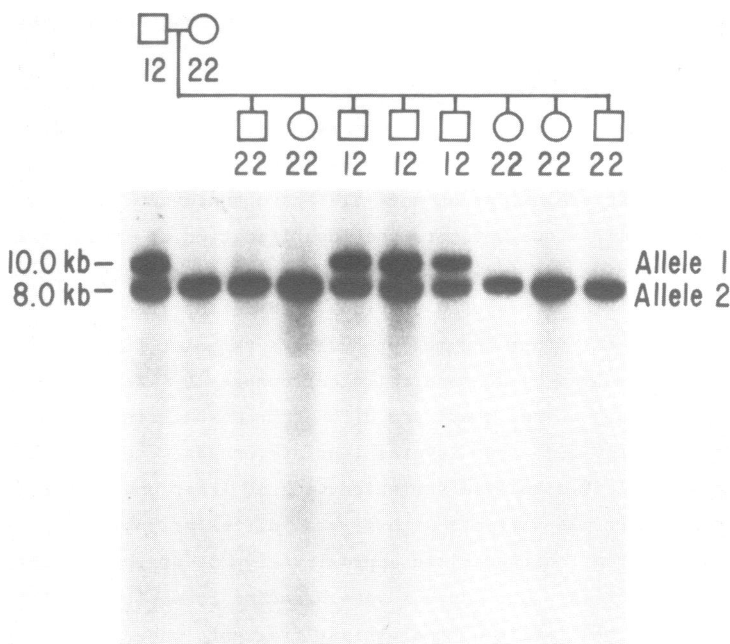

Figure 4-- Mendelian Inheritance of a Typical RFLP
A typical RFLP tested for Mendelian transmission in a nuclear family is shown. Each lane contains approximately 5 ug of PstI-digested DNA isolated from lymphoblastoid cell lines generated from each member of the family, fractionated by agarose gel electrophoresis (0.8% agarose, 90 V, 18 hours), transferred to a nylon filter and hybridized to probe B3D.

values less than 0.1 indicating they would be useful in less than 10% of families studied. The calculated PIC values for the other 40 markers are listed in Table 1. These loci all revealed sufficient polymorphism to be useful for linkage mapping given a reasonable source of pedigrees.

Despite their proximity, most pairs of RFLPs detected by a single probe were not in complete linkage disequilibrium. The individual genotypes for these sites can be used to increase the informativeness of the marker locus by construction of haplotypes to produce a higher PIC value (Table 1). Thus, the strategy of screening a single probe with multiple enzymes to detect independent RFLPS can readily increase the value of a particular marker. The estimated PIC for each marker, after combining information from all RFLPs detected, ranged from 0.10 to 0.84. For many of these loci, it should be possible to further increase PIC by analyzing additional variants detected in our initial polymorphism screen but not characterized here.

Table 1--Polymorphic anonymous DNA loci on chromosome 4

Locus Symbol	Probe Name	Insert Size ^	Chromosome Region @	Enzyme	Invariant Fragments ^	Allelic Fragments ^ (Frequency)	PIC	Overall PIC
D4S15	3C1	2.8	A	BanI	.4	6.0(76%), 4.5(24%)	.30	.35
				XmnI		23(95%), 20(5%)	.09	
D4S16	3E5	3.5	B	MspI		14(44%), 11+2.8(56%)	.37	
D4S17	4F6	0.75	B	BamHI		16(38%), 2.7(62%)	.36	.36
				RsaI		.55(62%), .50(38%)	.36	
D4S18	4F2	2.8	A	BamHI		10(72%), 5.0(28%)	.32	
D4S19	2B3	2.1	A	Sau96	4.0	2.0(38%), 1.0(62%)	.36	
D4S20	pGS42	2.7	A	Sau96	.7,.6,.5,.4	.85(52%), .54+.31(48%)	.37	.43
				BclI		20(96%), 19+.80(4%)	.07	
D4S21	G1E5	2.7	A	BamHI		11(45%), 9.0(55%)	.37	.71
				BanII*	.4	3.5(68%), 1.9(32%)	.34	
				BstXI	1.0	3.0(45%), 2.5(24%), 1.5(31%)	.58	
D4S22	G3E7	2.3	B	BclI		18(67%), 10(33%)	.34	
D4S23	GDS5	5.0	A	BglII	.6	10(22%), 8.0(78%)	.28	.30
				HincII		7.5(97%), 3.0+4.5(3%)	.06	
D4S24	G3D6	2.0	C	HinfI*		2.2(6%), 2.0(60%), 1.8(34%)	.43	.70
				StuI	10	18(59%), 12(41%)	.37	
				PvuII		8.0(63%), 6.5(37%)	.36	
D4S25	G3D5	3.2	B	EcoRI		4.7(26%), 4.0(74%)	.31	
D4S26	4H4	1.8	B	PstI*		7.0(6%), 6.5(94%)	.11	
D4S27	4E8	2.8	C	MspI		12(54%), 7.2(46%)	.37	
D4S30	pGS41	6.0	C	BstN1		2.0(77%), 1.7(23%)	.29	
D4S31	G1G12	3.0	C	PstI		9.0(50%), 7.0(50%)	.37	.47
				Bsp1286	1.2	3.8(50%), 3.6(50%)	.37	
				BglI		18(91%), 6.8(9%)	.15	
D4S32	GDS1	4.2	C	Bsp1286		3.9(81%), 3.7(19%)	.26	
D4S33	GDS8	4.5	C	EcoRV	1.0	6.5(40%), 6.0(60%)	.37	.43
				BanII	5.0	14(95%), 10(5%)	.09	
D4S34	3C5	2.0	C	BstXI		9.5(18%), 7.5(82%)	.24	
D4S38	4F4	0.8	C	BanII*		5.0(6%), 3.5(94%)	.11	
D4S39	A4C	1.5	B	BanI	5.0	6.0(63%), 3.5(37%)	.36	
D4S40	B5A	5.1	C	Bsp1286	2.2	5.2(25%), 5.0(75%)	.31	.84
				PstI	6.2	16(71%), 5.5(29%)	.33	
				XbaI		12(42%), 10(27%), 7.5(25%), 6.5(6%)	.63	
D4S41	6MC1	4.2	A	SacI	5.5	23(90%), 18(10%)	.16	
D4S42	C11H	3.5	A	PvuII		23(8%), 10(92%)	.14	
D4S44	B3D	4.5	C	PstI		10(32%), 8.5(68%)	.34	
D4S45	G1G9	4.2	C	Bsp1286		2.2(78%), 2.1(22%)	.28	
D4S46	5MG4	3.8	A	MspI	1.7	4.2(66%), 2.2+2.0(34%)	.35	
D4S47	16MD4	3.0	B	HincII		6.8(62%), 3.0(38%)	.36	
D4S48	2MG7	1.8	A	HinfI	.5,.6	1.3(24%), 1.0(76%)	.30	
D4S49	6MA7	1.9	A	TaqI		7.0(54%), 3.9(46%)	.37	
D4S50	16MH2	2.1	A	BstXI		3.8(24%), 3.5(76%)	.30	.43
				MspI		19(10%), 16+2.2(90%)	.16	
D4S51	16MG7	6.5	A	TaqI	1.0,2.0	4.7(78%), 4.5(22%)	.28	.45
				TaqI		1.7(80%), 1.4(20%)	.27	
D4S52	17MD1	1.8	A	PvuII		7.5(60%), 5.0(40%)	.37	.66
				PvuII		3.4(7%), 3.0(93%)	.12	
				BstXI		18(42%), 17(58%)	.37	
D4S53	19MC12	4.2	A	EcoRV	3.5	10(21%), 6.6+3.7(79%)	.28	
D4S54	1MC11	2.4	A	Bsp1286	10,.5	1.5(24%), 1.3(76%)	.30	
D4S55	D7G	2.1	B	TaqI	6.0	23(63%), 20(37%)	.36	
D4S56	17MA3	2.4	B	Sau96	1.0	2.5(27%), 7.5(73%)	.32	.57
				MboI	1.1,1.0	.8(28%), .45+.35(72%)	.32	
D4S57	B7C	0.8	A	BglII		7.5(76%), 6.5(24%)	.30	
D4S59	C7F	0.9	A	HindIII		1.5(10%), 0.9(90%)	.16	
D4S60	B4D	2.5	A	BstXI	2.5	9.4(9%), 3.7(91%)	.15	.29
				BanI		5.4(90%), 2.8+2.6(10%)	.16	
D4S61	7MG3	2.5	A	EcoRI*		4.8(85%), 4.5(15%)	.22	.30
				EcoRI		1.5(94%), 1.4(6%)	.11	

@ A=4pter-p15.1; B=4p15.1-q11; C=4q11-qter

* denotes an insertion/deletion type polymorphisms also detectable with several other enzymes

^ Fragment sizes are given in kilobase pairs if > 10 kb are rounded to the nearest integer.

Only three probes detected three or four allele patterns with PIC values above 0.375, the maximum for a two-allele system. Five probes displayed insertion/deletion type variation that could be detected with several different enzymes, compared to 71 apparent single site variations. Unfortunately, none of the insertion/deletion polymorphisms generated the large numbers of alleles seen for certain markers of this type (23-25) supporting the view that the latter are relatively rare and that selective cloning procedures may be needed to identify them more efficiently.

It has been proposed that restriction enzymes containing either CpG or TpG dinucleotides or purine/pyrimidine strings within their recognition sequences are more likely to detect RFLPs (18,22). Like others, we find that MspI and TaqI, whose recognition sequences both contain CpG, frequently detect polymorphism. These two enzymes, together with PvuII and BstXI which have recognition sequences containing TpG, account for 17 of the 55 putative single base change polymorphisms in Table 1. There was no indication, however, that contiguous stretches of purine or pyrimidine bases were associated with frequent polymorphism.

Heterozygosity on Chromosome 4

There is some question whether polymorphisms occur uniformly in the genome or whether there are regional and/or chromosomal differences in their frequency of occurrence. Systematic studies of two segments, the beta globin locus (26) and the serum albumin gene (27) have yielded heterozygosity estimates of 0.0041 and 0.0025, respectively. This indicates that approximately 1 in every 250 to 400 base pairs differs between any two chromosomes in the region surrounding these coding sequences. Cooper et al.(28,29) screened about 27,000 bp on the X chromosome and chromosomes 7, 15, 21 and 22 with anonymous DNA probes. Their estimate of heterozygosity on the autosomes, 0.0039, was comparable to the previous studies, but considerably higher than the figure of 0.0006 calculated for the X chromosome.

We have used our more extensive data set to estimate the level of heterozygosity detected on chromosome 4. Heterozygosity was calculated by the method of Nei (30) as described for RFLPs by Cooper et al. (28,29), taking into account all variants observed among the 10 chromosomes (5 individuals) tested. About 25% of the variants included were differences that were not pursued for the generation of markers and are therefore not listed in Table 1. Overall, we screened approximately 263,000 bp or 0.13% of the chromosome for sequence variation and calculated an average heterozygosity for chromosome 4 of 0.0021 over all bases scanned. There is a degree of uncertainty in our estimate introduced by the possibility of overlap in the sites detected by

different enzymes, and the consequent potential for observation of the same variation with more than one enzyme. The level of heterozygosity we observed agrees most closely with the estimate of Murray et al. (27) for the albumin locus which is also on chromosome 4, but does not differ strikingly from the estimates of Jeffreys (26) and Cooper et al. (28,29). On average, about 1 base in 500 differs between any two chromosome 4's.

Deviation from Expected Restriction Fragment Sizes

The large data set gathered in our RFLP screen has also permitted us to examine the distribution of fragment sizes from a single autosome produced by each restriction enzyme relative to that expected based on the recognition sequence. A similar analysis by Aldridge et al. (18) for the X chromosome previously suggested a substantial deviation from the expected distribution for the enzymes MspI and TaqI. Our data indicate that a similar phenomenon occurs on chromosome 4. We observed deviations from the expected mean restriction fragment size for both enzymes. The mean size of 88 MspI fragments was 9.2kb compared to a predicted 2.7kb. Furthermore 107 TaqI fragments averaged 5.9kb compared to a predicted 2.5 kb. Although the significance of this deviation is not obvious, it is clearly not confined to the sex chromosomes. Since both TaqI and MspI have CpG dinucleotides in their respective recognition sequence, the larger than expected fragment sizes observed with each might relate to the clustering of CpG rich sequences in "HTF islands" combined with our inability to monitor very small fragments (31).

DISCUSSION

Chromosome 4 comprises approximately 6.5% of the human genome and probably encodes 5,000 to 10,000 genes (10). A linkage map of this autosome would likely span at least 200 to 300 cM. To date, there have been insufficient polymorphic loci assigned to chromosome 4 to permit construction of a complete map. A number of cloned genes and pseudogenes detecting RFLPs have been described (9,10). The former include albumin (ALB), alpha-fetoprotein (AFP), several alcohol dehydrogenase genes (ADH1, ADH2, ADH3), epidermal growth factor (EGF), interleukin 2 (IL2) and the fibrinogen cluster (FGA, FGB, FGG), all of which reside on the long arm. This region also contains two useful expressed polymorphic loci, GC or vitamin D binding protein, and MN, the MNS blood group system. A number of anonymous DNA probes detecting RFLP have been described, a few of which map to the short arm (32). Among these is D4S10, the highly informative locus linked to HD in band 4p16.

Only two other disease genes have been assigned to chromosome 4 by family studies: Dentinogenesis imperfecta (DGI1), a defect in dentition and Sclerotylosis (TYS), a skin disorder. This small number is surprising in view of the presence on the long arm of the two expressed markers, GC and MN, both commonly used in linkage studies.

In this report, we have presented 29 new RFLP loci on the short arm that overcome the paucity of markers assigned to this region. The most terminally located of these are potential candidates for linkage to the HD gene. In addition to RFLPs on the short arm, we have contributed 11 new DNA markers to the long arm of the chromosome. With this report, there are more characterized polymorphic DNA markers for chromosome 4 than for any other autosome.

The body of data generated in this study is the largest analyzed to estimate heterozygosity for a single autosome. Although it should be realized that our calculation is not based on an unbiased screen of the entire chromosome since testing of short arm probes was favored, the level of heterozygosity detected on chromosome 4, at 0.0021, is comparable to other autosomal regions. Thus, given the high quality of the chromosome 4 library used for these investigations, additional markers could be readily generated for this chromosome should they be needed. Our data, taken together with previous descriptions of long arm chromosome 4 markers, imply that it is now feasible to construct a detailed linkage map of human chromosome 4. The map will not only contribute to more precise localization of the HD gene, but will permit more efficient linkage studies using multipoint methods to detect other inherited disorders caused by defects on chromosome 4. Furthermore, the .pa individual marker loci in the linkage group will be excellent starting points for the construction of a fine structure physical map of the chromosome using pulsed field or inversion field gel electrophoresis (33,34). The existence of a set of genetically and physically ordered DNA fragments spanning all regions of the chromosome will permit a detailed comparison of the two types of maps and should greatly facilitate efforts to determine the complete DNA sequence of chromosome 4.

ACKNOWLEDGMENTS

This work was supported by NINCDS grants NS16367 (Huntington's Disease Center Without Walls) and NS20012, and by grants form the McKnight Foundation, Hereditary Disease Foundation and Julieanne Dorn Fund for Neurological Research. JFG is a Searle Scholar of the Chicago Community Trust. TCG and MEM were supported by the Anna Mitchell Fellowship and Grace E. Neuman Fellowship

respectively, from the Hereditary Disease Foundation. GDS was supported by a fellowship from the Wills Foundation. We thank Drs M. Smith, E. Stanbridge and H. Willard for providing hybrid cell lines and Dr M. VanDilla for providing the chromosome 4 library.

*To whom correspondence should be addressed

REFERENCES

1. Botstein, D., White, R.L., Skolnick, M., and Davis, R. (1980) Am. J. Hum. Genet. 32,314-331.
2. Gusella, J.F. (1986) Ann. Rev. Biochem. 55,831-54.
3. Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y., Young, A.B., Shoulson, I., Bonilla, E., and Martin, J.B. (1983) Nature 306,234-238.
4. Gusella, J.F., Tanzi, R.E., Anderson, M.A., Hobbs, W., Gibbons, K., Raschtchian, R., Gilliam, T.C., Wallace, M.R., Wexler, N.S., and Conneally, P.M. (1984) Science 225,1320-1326.
5. Reeders, S.T., Breuning, M.H., Davies, K.E., Nicholls, R.D., Jarman, A.P., Higgs, D.R., Pearson, P.L. and Weatherall, D.J. (1985) Nature 317,542-544.
6. Tsui, L., Buchwald, M., Barker, D., Braman, J., Knowlton, R., Schumm, J., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Helms, C., Gravius, T., Parker, C., Rediker, K., and Donis-Keller, H. (1985) Science 230,1054-1057.
7. White, R. L., Woodward, S., Leppert, M., O'Connell, P., Hoff, M., Herbst, J., Lalouel, J.M., Dean, M., and Van de Woude, G. (1985) Nature 318,382-384.
8. Wainwright, B., Scambler, P.J., Schmidtke, J., Watson, E.A., Law, H.Y., Farrall, M., Cooke, H.J., Eiberg, H., and Williamson, R. (1985) Nature 318,384-386.
9. Kidd, K.K., and Gusella J.F. (1985) Human Gene Mapping 8, Cytogenet Cell Genet 40,107-127.
10. Gusella, J.F., Gilliam, T.C., MacDonald, M.E., Cheng, S.V., and Tanzi, R.E. (1986) J. Med. Genet. 23,193-99.
11. Gusella, J., Varanyi-Breiner, A., Kao, F., Jones, C, Puck, T.T., Keys, C., Orkin, S., and Housman, D. (1979) Proc. Natl. Acad. Sci. USA 76,5239-5243.
12. Saxon, P.J., Srivatsan, E.S., Leipzig, G.V., Sameshima, J.H. and Stanbridge, E.J. (1985) Mol. Cell. Biol. 5,140-146.
13. Bocian, M., Mohandas, T., Sparkes, R.S., and Funderbunk., S. (1979). Cytogenet. Cell Genet. 25,138.
14. Wasmuth, J.J., Carlock, L.R., Smith, B., and Immken, L.L. (1986) Am. J. Hum. Genet. 39, 397-403.
15. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular cloning, A laboratory manual. Cold Spring Harbor Press, New York.
16. Feinberg, A.P., and Vogelstein, B. (1983) Anal. Biochem. 137,266-267.
17. Gusella, J.F., Keys, C., Varsanyi-Breiner, A., Kao, F., Jones, C., Puck, T.T., and Housman, D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2829-2833.
18. Aldridge, J., Kunkel, L., Bruns, G., Tantravahi, U., Lalande, M., Brewster, T., Moreau, E., Wilson, M., Bromley, W., Roderick, T., and Latt, S.A. (1984). Am. J. Hum. Genet. 36,546-564.
19. Magenis, R.E., Gusella, J., Weliky, K., Olson, S., Haight, G., Toth-Fejel, S., and Sheehy, R. (1986) Am. J. Hum. Genet., 39, 383-391.

20. Zabel, B.U., Naylor, S.L., Sakaguchi A.Y., and Gusella J.F. (1986) Cytogenet Cell Genet, 42, 187-191.
21. Gusella, J.F., Tanzi, R.E., Bader, P.I., Phelan, M.C., Stevenson, R., Hayden, M.R., Hofman, K.J., Faryniarz, A.G., and Gibbons, K. (1985) Nature 318,75-78.
22. Barker, D., Schafer, M., and White, R. (1984) Cell 36,131-138.
23. Wyman, A. and White, R. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 6754-6758.
24. Barker, D., McCoy, M., Weinberg, R., Goldfarb, M., Wigler, M., Burt, R., Gardner, E., and White, R. 1983. Mol. Biol. Med. 1,199-206.
25. Bell, G.I., Karam, J.H., and Rutter, W.J. 1981. Proc. Natl. Acad. Sci. U.S.A. 1981. 78,5759-5763.
26. Jeffreys, A.J. (1979). Cell 18,1-10.
27. Murray, J.C., Mills, K.A., Demopolous, C.M., Hornung, S., and Motulsky, A. (1984) Proc. Nat. Acad. Sci. USA 81, 3486-3490.
28. Cooper, D.N. and Schmidtke, J. (1984). Hum. Genet. 66,1-16.
29. Cooper, D.N., Smith, B.A., Cooke, H.J., Niemann, S. and Schmidtke, J. (1985). Hum. Genet. 69,201-205.
30. Nei, M. (1975) Molecular Population Genetics and Evolution (North Holland, Amsterdam).
31. Bird, A.P. (1986) Nature 321,209-213
32. Gilliam, T.C., Scambler, P., Robbins, T., Ingle, C., Williamson, R., and Davies, K.E. (1984). Hum Genet. 68,154-158.
33. Smith, C.L., Lawrence, S.K., Gillespie, G.A., Cantor, C.R., Weissman, S.M., and Collins, F.S. (1986) In "Methods in Enzymology", in press.
34. Carle G.F., Frank, M., and Olson, M.V. (1986) Science 232,65-68.