The EMBO Journal vol.4 no.9 pp.2257–2260, 1985

A linkage map of three anonymous human DNA fragments and SOD-1 on chromosome 21

Smita D.Kittur[1], Stylianos E.Antonarakis[1], Rudolph E.Tanzi[6], Deborah A.Meyers[2], Aravinda Chakravarti[3], Yoram Groner[4], John A.Phillips[1,7], Paul C.Watkins[5], James F.Gusella[6] and Haig H.Kazazian,Jr.[1]

[1]Genetics Unit, Department of Pediatrics, [2]Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, [3]Human Genetic Program, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA, [4]Weizmann Institute, Rehovot, Israel, [5]Integrated Genetics, Inc., Framingham, MA 01701, and [6]Neurogenetics Laboratory, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

[7]Present Address: Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

Communicated by Y.Groner

Using DNA polymorphisms adjacent to single-copy genomic fragments derived from human chromosome 21, we initiated the construction of a linkage map of human chromosome 21. The probes were genomic *Eco*RI fragments pW228C, pW236B, pW231C and a portion of the superoxide dismutase gene (SOD-1). DNA polymorphisms adjacent to each of the probes were used as markers in informative families to perform classical linkage analysis. No crossing-over was observed between the polymorphic sites adjacent to genomic fragments pW228C and pW236B in 31 chances for recombination. Therefore, these fragments are closely linked to one another ($\hat{\theta}$ = 0.00, lod score = 6.91, 95% confidence limits = 0–10 cM) and can be treated as one 'locus' with four high-frequency markers. There is a high degree of non-random association of markers adjacent to each of these two probes which suggests that they are physically very close to one another in the genome. The pW228C - pW236B 'locus' was also linked to the SOD-1 gene ($\hat{\theta}$ = 0.07, lod score = 4.33, 95% confidence limits = 1–20 cM). On the other hand, no evidence for linkage was found between the pW228C-pW236B 'locus' and the genomic fragment pW231C ($\hat{\theta}$ = 0.5, lod score = 0.00). Based on the fact that pW231C maps to 21q22.3 and SOD-1 to 21q22.1, we suggest that the pW228C-pW236B 'locus' lies in the proximal long arm of chromosome 21. These data provide the outline of a linkage map for the long arm of chromosome 21, and indicate that the pW228C-pW236B 'locus' is a useful marker system to differentiate various chromosome 21s in a population.

Key words: DNA polymorphisms/linkage analysis/SOD-1 gene/linkage disequilibrium

Introduction

The mapping of genes on human chromosomes is important for many reasons, among which is the development of a system of markers which will aid in the genomic localization of disease genes. Genes can be mapped to chromosomes by use of rodent-human somatic-cell hybrid lines and by *in situ* hybridization. The ordering of the genes and the recombination distance between them are obtained from linkage studies in informative families using allelic differences observed in the genes of interest. Such studies have now been carried out for the short arm of chromosome 11 (Lebo *et al.*, 1983; Antonarakis *et al.*, 1983; Fearon *et al.*, 1984; Gerhard *et al.*, 1984), the long arm of chromosome 13 (Cavenee *et al.*, 1984), and the X chromosome (Drayna *et al.*, 1984).

Chromosome 21 is an acrocentric chromosome of the G group which has been estimated to contain 1.8% of the human genome (Maynard-Smith *et al.*, 1961). Using DNA polymorphisms adjacent to single-copy DNA fragments which map to human chromosome 21, we initiated the construction of a linkage map for this chromosome. For this purpose, we have used genomic *Eco*RI fragments pW228C, pW236B, pW231C and a genomic superoxide dismutase (SOD-1) gene fragment (see Materials and methods for probe origin). Linkage analysis showed that fragments pW236B and pW228C are very closely linked to one another and can be considered as one locus with four high-frequency polymorphic markers. These fragments are also linked to the SOD-1 gene, but very loosely linked to pW231C.

Results

All four probes map to chromosome 21q

The four probes (pW228C, pW236B, pW231C, SOD-1) were hybridized to DNA from Chinese hamster ovary (CHO)-human somatic cell hybrid lines containing all or part of chromosome 21 as the only human chromosome. The hybrids used were 153E9a and 72532X-6 containing the entire human chromosome 21 and 2FUrl and 153E7b containing 21q (kindly provided by David Patterson, University of Colorado Health Sciences Center, Denver, Colorado [Patterson *et al.*, 1983]). Each probe showed the same hybridization signal with all four hybrid cell lines suggesting that all these DNA fragments are localized on chromosome 21q (data not shown). No hybridization was observed with DNA from the parental Chinese hamster ovary cell line.

Linkage between pW228C and pW236B

Seven large nuclear families of different ethnic backgrounds were used for the analysis. Using probe pW228C, polymorphic *Msp*I and *Bam*HI sites were studied. With pW236B, polymorphic *Eco*RI and *Taq*I sites were analyzed (Watkins *et al.*, 1984).

Segregation patterns in all families examined were completely consistent with Mendelian inheritance for each of the four polymorphic sites. Since these sites are highly polymorphic, 80% of the total families examined were informative. Close linkage was observed between these probes since no recombinants were found in 31 chances ($\hat{\theta}$ = 0.00, lod score = 6.91, 95% confidence limits = 0–10 cM) (Table III).

Linkage disequilibrium among the four markers (BamHI-pW228C, MspI-pW228C, EcoRI-pW236B, and TaqI-pW236B

The study of linkage disequilibrium among these four polymorphic sites was performed in 37 Greek nuclear families. The frequency of the presence and absence of each polymorphic site in this population is shown in Table I. The expected probability

Table I. Frequency of the presence of polymorphic restriction sites for pW228C, pW236B, SOD-1, and pW231C in the Greek population

Polymorphic site	No. of chromosomes examined	Site present	Frequency
*Bam*HI, pW228C	108	80	0.74
*Msp*I, pW228C	108	77	0.71
*Taq*I, pW236B	108	74	0.69
*Eco*RI, pW236B	108	76	0.70
*Msp*I, SOD-1	102	5	0.05
*Taq*I, pW231C	103	90	0.87

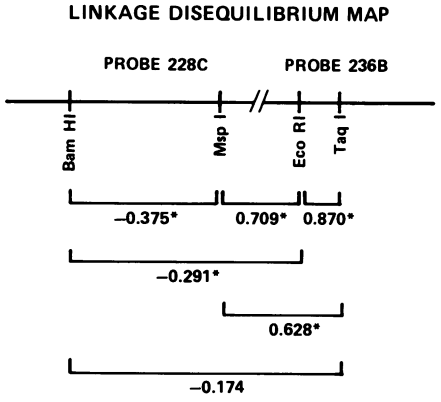

Fig. 1. Linkage disequilibrium map showing Δ values between *Bam*HI, *Msp*I, sites of probe pW228C and *Eco*RI, *Taq*I sites of probe pW236B. *Represents statistically significant values at the 1% level. + and − represent excess or deficiency of ++ and −− haplotypes relative to +− and −+ haplotypes respectively.

that, for example, any chromosome 21 is marked with the absence of polymorphic sites *Eco*RI-pW236B and *Taq*I-pW236B is 0.30 × 0.31 = 0.093. The observed value was 0.278 which differs significantly from the expected. The strength of the linkage disequilibrium between two loci can be measured with the use of the delta value (Hill *et al.*, 1968; Morton, 1982). Δ is the correlation coefficient between the uniting gametes at the two loci and is independent of the actual frequencies of the polymorphic sites (Hill *et al.*, 1968). When Δ = 0 association between two loci is random and when Δ = ±1 association is complete with a positive or negative sign implying an excess or deficiency of ++ and −− haplotypes relative to +− and −+ haplotypes, respectively (Chakravarti *et al.*, 1984a). Δ values for each pair of sites were *Bam*HI/*Msp*I = −0.38, *Bam*HI/*Taq*I = −0.17, *Bam*HI/*Eco*RI = −0.29, *Msp*I/*Taq*I = +0.63, *Msp*I/*Eco*RI = +0.71, *Taq*I/*Eco*RI = +0.87 (Figure 1). The most striking finding is the extent of linkage disequilibrium between the *Msp*I site discovered by probe pW228C and the *Eco*RI site discovered by probe pW236B. The Δ value for these two sites is 0.71. This degree of linkage disequilibrium is extremely strong and suggests that pW228C and pW236B are within 50 kb of each other (Chakravarti *et al.*, 1984a).

The four DNA polymorphisms were then used to construct haplotypes for the pW228C-pW236B region of chromosome 21. Because of the strong linkage disequilibrium among these four polymorphisms, three common haplotypes were observed, ++++, +−−−, and −+++ (+ stands for presence of site, − denotes absence of site). The order of polymorphic sites in the haplotypes is given in Table II. Haplotypes 1, 2, and 3 were observed in 43, 24, and 23 chromosomes of only 108 chromosomes analyzed, compared to the expected number of of 28, 2, and 10 chromosomes, respectively. The overall linkage disequilibrium for all haplotypes is very high [x^2 = 130.09, p <0.0001] (Table II).

Table II. Haplotypes for the pW228C-pW236B region of chromosome 21

Haplotype	Observed	Expected
++++	43	27.50
+−−−	24	2.14
−+++	23	9.63
+−+−	6	11.07
++−−		
+++−	6	29.54
++−+		
Other	6	28.12
	108	108

x^2 = 130.09, df = 1, p <0.0001

The data presented represent haplotypes oberved in 108 chromosome 21s of 27 couples in a single ethnic group, Greeks. The order of the sites is *Bam*HI-pW228C, *Msp*I-pW228C, *Eco*RI-pW236B, and *Taq*I-pW236B.

Linkage between pW228C/pW236B, and the SOD-1 gene

A polymorphic *Msp*I restriction site was found in the SOD-1 gene region. After genomic DNA was digested with *Msp*I and hybridized to the SOD-1 genomic probe, we observed SOD-1 sequences in various individuals in 3.2-kb or 5-kb fragments. Three types of individuals were observed; homozygotes for the 3.2-kb fragment, homozygotes for the 5-kb fragment and heterozygotes for the 5-kb and 3.2-kb fragments. The frequency of the presence of the *Msp*I site in the Greek population was 0.05 among 102 chromosomes examined. Segregation patterns were completely consistent with Mendelian inheritance of this polymorphic site. Six families were informative for SOD-1 and the pW228C-pW236B haplotypes and close linkage between the pW228C-pW236B 'locus' and SOD-1 was observed ($\hat{\theta}$ = 0.07, lod score = 4.33, 95% confidence limits = 1−20 cM).

Linkage between the pW228C/pW236B 'locus' and pW231C

A polymorphic *Taq*I restriction site was found after restriction analysis using the pW231C DNA fragment. After genomic DNA was digested with *Taq*I and hybridized to the pW231C probe, we observed pW231C sequences in various individuals in 5.1-kb or 4.2-kb fragments. Three types of individuals were observed; homozygotes for the 5.1-kb fragment, homozygotes for the 4.2-kb fragment, and heterozygotes for the 5.1-kb and 4.2-kb fragments. The frequency of the presence of this site was 0.87 in the Greek population among 103 chromosomes examined. Family studies using this polymorphic site again demonstrated Mendelian inheritance of the site. No evidence for linkage was found between the pW228C-pW236B locus and pW231C ($\hat{\theta}$ = 0.5).

Construction of a linkage map

The human SOD-1 gene has been mapped to 21q22.1 using somatic cell hybrids (Sinet *et al.*, 1976). The DNA fragment pW231C has been mapped at the lower end of 21q very close to qter (Antonarakis *et al.*, 1984 and D. Patterson, personal communication). Using the data presented here, along with these other studies, we propose the outline of a map of chromosome 21q as centromere [pW228C/pW236B − 7 cM − SOD-1] − >30 cM − pW231C (Figure 2).

Discussion

Several genes have been assigned to chromosome 21 by *in-situ*

Table III. Linkage analysis between the pW228C, pW236B, SOD-1 and pW231C DNA fragments

DNA fragments	Families examined	R/T	θ 0.05	0.10	0.20	0.30	0.40	$\hat{\theta}$	lod score	Max 95% confidence limits
pW228C-pW236B	7	0/31	6.54	5.81	4.30	2.66	1.01	0.00	6.91	0−0.10
pW228C/pW236B-SOD-1	6	2/31	4.31	4.25	3.44	2.25	0.87	0.07	4.33	0.01−0.20
pW228C/pW236B-pW231C	4	6/17	−3.81	−2.24	−0.90	−0.33	−0.08	0.50	0.00	

For each pair of loci, the lod score for different values of recombination fraction (θ) and the estimated value of θ ($\hat{\theta}$) and its corresponding maximum lod scores are shown.
R/T = Recombinants/Total.

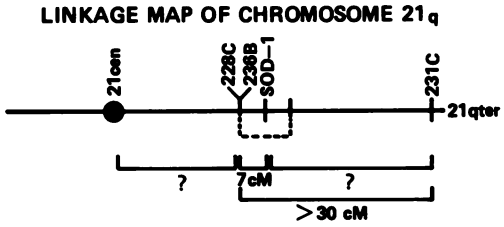

Fig. 2. Linkage map for human chromosome 21. The genetic distance in cM found between pW228C, pW236B, SOD-1, and pW231C is shown. One cM is defined as the genetic distance equivalent to a recombination fraction, θ, of 0.01. The location of pW228C and pW236B relative to the SOD-1 locus is indeterminate from our data; therefore, the pW228C/pW236B 'locus' is shown on either side of the SOD-1 locus.

hybridization and analysis of somatic cell hybrid panels. Among these are ribosomal rRNA genes to the short arm (Schmickel and Knoller, 1977), superoxide dismutase-1 (SOD-1) to 21q22.1 (Tan *et al.*, 1973), interferon receptor to 21q21-qter (Tan *et al.*, 1973), phosphoribosylglycinamide synthetase to 21q22 (Patterson *et al.*, 1981; Moore *et al.*, 1977), liver phosphofructokinase to 21q22 (Vora and Francke, 1981), and cystathionine β synthase to 21q (Skovby *et al.*, 1984). Mutation of this latter gene is responsible for the most common form of homocystinuria. In this study, we have constructed a linkage map for SOD-1 and three random single-copy DNA fragments (pW228C, pW236B and pW231C) which were derived from a chromosome 21-specific genomic library (Watkins *et al.*, 1984).

Using DNA polymorphisms in family studies, we found that probes pW228C and pW236B are closely linked to each other. In fact, polymorphisms detected using one probe show a significantly high degree of linkage disequilibrium with polymorphisms detected by the other. This degree of linkage disequilibrium is comparable to the most striking association observed between polymorphisms in the same gene cluster, such as the β globin cluster (Antonarakis *et al.*, 1982; Chakravarti *et al.*, 1984a) or the growth hormone cluster (Chakravarti *et al.*, 1984b). This observation suggests that these probes are in all likelihood within 100 kb, and perhaps even 20−50 kb, of each other. However, the exact physical distance between pW228C and pW236B is yet to be determined. Since pW228C and pW236B are DNA fragments picked at random from a chromosome 21-specific library, their tight linkage is unexpected and truly extraordinary. The four DNA polymorphisms associated with these two probes provide an excellent set of markers for the long arm of chromosome 21 and can be used for the identification of different chromosome 21s in various studies.

The maximum likelihood estimate from our family data places pW228C and pW236B at a distance of 7 cM from the SOD-1 gene. Thus, these probes lie on the proximal long arm of chromosome 21. We have observed a number of crossing-over events between these probes and pW231C such that the best estimate of the recombination distance between pW228C/pW236B and pW231C is greater perhaps than 40 cM. Patterson and coworkers have used somatic cell hybrids containing various regions of chromosome 21 to map pW231C to the terminal region of the long arm of chromosome 21. In addition, pW231C appears to lie at the long-arm breakpoint (q23.2) of a ring 21 chromosome (Kazazian *et al.*, 1985).

Since we did not have any informative families for both SOD-1 and pW231C, no estimate of the recombination fraction between these probes is available. Although we know that the pW228C-pW236B complex is close to SOD-1, we do not know whether this complex is proximal or distal to SOD-1 on 21q (Figure 2). Families in which both pW231C and SOD-1 are informative are necessary to answer this question. Continued construction of a linkage map for chromosome 21 will facilitate the study of diseases associated with this chromosome.

Materials and methods

Fifteen large nuclear families of different ethnic backgrounds (Venezuelans, Utah Caucasians, other Caucasians, Asian Indians, and Blacks) were used for the linkage analysis, of which 10 were informative for different DNA polymorphism markers.

Restriction endonuclease analysis of genomic DNA

Nuclear DNA was isolated from the leukocytes contained in 10−15 ml of EDTA anti-coagulated blood or from the cultured lymphoblastoid cells contained in one 25 cm^3 tissue-culture flask (Kunkel *et al.*, 1978). 5−10 μg of DNA was digested with one of various endonucleases using conditions recommended by the commercial suppliers. Southern transfer analysis of the resulting DNA fragments was performed as described (Southern, 1975; Scott *et al.*, 1979).

Radioactive probes

The following probes were used: (i) Genomic *Eco*RI fragments pW228C (1.5 kb), pW236B (1.85 kb) and pW231C (2.1 kb) cloned in pBR328 (Watkins *et al.*, 1984), and (ii) SOD-1 (4.1-kb *Bgl*II fragment of genomic DNA) in pBR322 (Levanon *et al.*, 1985). The anonymous *Eco*RI fragments were isolated from a chromosome 21-specific DNA library by Watkins *et al.* (Watkins *et al.*, 1984). This library was constructed from DNA of the mouse-human hybrid cell WA17 which contains human chromosome 21 as its only human chromosome.

All fragments were radiolabeled with [^{32}P]dATP and [^{32}P]dCTP by nick-translation function of *E. coli* DNA polymerase I (Maniatis *et al.*, 1976).

Experiments involving recombinant DNA were performed in P_1-EK_1 containment in accordance with the National Institutes of Health guidelines.

Linkage analysis and linkage disequilibrium

Linkage analysis was carried out using the method of maximum likelihood (Morton, 1955) and the computer program LIPED (Ott, 1974). Lod scores were calculated at various recombination fractions where the lod for each recombination fraction represents the logarithm of the odds in favor of linkage versus non-linkage. Non-random association (linkage disequilibrium) between the different polymorphic sites detected by probes pW228C and pW236B was calculated as previously described (Chakravarti *et al.*, 1984a).

Acknowledgements

We would like to thank Emily Pasterfield for preparation of this manuscript. We also thank Drs. D. Patterson and M. Van Keuren for providing us with DNA from hybrid cell lines 153E9a, 72532X-6, 2FUr1 and 153E7b. This work was

supported by NIH grants to HHK and JFG, and a New Investigator Research Award to SEA.

References

Antonarakis,S.E., Boehm,C.D., Giardina,P.J.V. and Kazazian,H.H.,Jr. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 137-141.

Antonarakis,S.E., Gusella,J., Stetten,G., Potter,M., Watkins,P. and Kazazian,H.H.,Jr. (1984) *Pediatr. Res.*, **78**, 219A.

Antonarakis,S.E. Phillips,J.A.,III, Mallonee,R.L., Kazazian,H.H.,Jr., Fearon, E.R., Waber,P.G., Kronenberg,H.M., Ullrich,A. and Meyers,D.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6615-6619.

Cavenee,W., Leach,R., Mohandas,T., Pearson,P., White,R. (1984) *Am. J. Hum. Genet.*, **36**, 10-24.

Chakravarti,A., Buetow,K.H., Antonarakis,S.E., Waber,P.G., Boehm,C.D. and Kazazian,H.H.,Jr. (1984a) *Am. J. Hum. Genet.*, **36**, 1239-1258.

Chakravarti,A., Phillips,J., Mellits,K., Buetow,K. and Seeburg,P. (1984b) *Proc. Natl. Acad. Sci. USA*, **81**, 6085-6089.

Drayna,D., Davis,K., Harley,D., Mandel,J., Camerino,G., Williamson,R. and White,R. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2836-2839.

Fearon,E.R., Antonarakis,S.E., Meyers,D.A. and Levine,M.A. (1984) *Am. J. Hum. Genet.*, **36**, 329-337.

Gerhard,D.S., Kidd,K.K., Kidd,J.K., Gusella,J., Housman,D. (1984) in Ahmad,F., Black,G., Schultz,J., Scott,W.A. and Whelan,W.J. (eds.), *ICSU Short Report*, pp. 170-171.

Hill,W.G. and Robertson,A. (1968) *Theor. Appl. Genet.*, **38**, 226-231.

Kazazian,H.H.,Jr., Antonarakis,S.E., Wong,C., Trusko,S.P., Stetten,G., Oliver,M., Potter,M.J., Gusella,J.F. and Watkins,P.C. (1985) in Smith,G.F. (ed.), *Molecular structure of chromosome 21 and Downs Syndrome, Ann. N.Y. Acad Sci.*, **450**, 33-42.

Kunkel,L.M., Smith,K.D., Boyer,S.H., Borgaonkar,D.S., Wachtel,S.F., Miller,O.J., Berg,W.R., Jones,H.W.,Jr. and Rary,J.M.H. (1978) *Proc. Natl. Acad. Sci. USA*, **74**, 1245-1249.

Lebo,R.V., Chakravarti,A., Buetow,K.H., Cheung,M.-C., Cann,H., Cordell,B.R. and Goodman,H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4808-4812.

Levanon,D., Lieman-Harwitz,J., Dafni,N., Winderson,M., Sherman,L., Bernstein,Y., Laver-Rudich,Z., Danciger,E., Stein,O. and Groner,Y. (1985) *EMBO J.*, **4**, 77-84.

Maniatis,T., Kee,G.F., Efstratiadis,A. and Kafatos,F. (1976) *Cell*, **8**, 163-182.

Maynard-Smith,S., Penrose,L.S. and Smith,C.A.B. (eds.) (1961) *Mathematical tables for research workers in human genetics*, Little, Brown, Boston, pp. 44.

Moore,E.E., Jones,C., Kao,F.T. and Oates,D.C. (1977) *Am. J. Hum. Genet.*, **29**, 389-396.

Morton,N.E. (1955) *Am. J. Hum. Genet.*, **7**, 277-318.

Morton,N.E. (1982) *Outline of Genetic Epidemiology*, Basel, Karger.

Ott,J. (1974) *Am. J. Hum. Genet.*, **26**, 588-597.

Patterson,D.B., Craw,S. and Jones,C. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 405-409.

Patterson,D.B. and Schandle,V.B. (1983) in Caskey,C.T. and White,R. (eds.), *Recombinant DNA applications to human disease*, Banbury Report 14, Cold Spring Harbor Laboratory Press, pp. 215-223.

Schmickel,R.D. and Knoller,M. (1977) *Ped. Res.*, **11**, 929-935.

Scott,A.F., Phillips,J.A.,III and Migeon,B.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4563-4565.

Skovby,F., Krassikoff,N. and Francke,U. (1984) *Hum. Genet.*, **65**, 291-294.

Sinet,P.M., Couturier,J., Dutrillaux,B., Poissonier,M., Raoul,O., Rethore,M.O., Allard,D., Lejeune,J. and Jerome,H. (1976) *Exp. Cell. Res.*, **97**, 47-55.

Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.

Tan,Y.H., Tischfield,J., Ruddle,F.H. (1973) *J. Exp. Med.*, **137**, 317-330.

Vora,S. and Francke,U. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3738-3742.

Watkins,P., Tanzi,R., Landes,G., Tricoli,J.V., Shows,T. and Gusella,J.F. (1984) *Birth Defects Original Article Series 20*: 602.

Received on 28 February 1985; revised on 7 June 1985