Jpn J Human Genet 38, 243–255, 1993

ISOLATION AND CHARACTERIZATION OF A DNA FRAGMENT CONTAINING VARIOUS KINDS OF REPETITIVE SEQUENCES LOCATED ON HUMAN CHROMOSOME 21

Ryoji YAO,[1,*] David PATTERSON,[2] and Kazukiyo ONODERA[1,**]

[1]*Department of Agricultural Chemistry, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*
[2]*Eleanor Roosevelt Institute for Cancer Research, 1899 Gaylord St. Denver, CO 80206, USA*

Summary In order to investigate the repetitive sequences located on human chromosome 21, we have isolated DNA fragments containing Alu sequences. One of the clones, p1, was chosen for further study, because it contained repetitive sequences different from the Alu sequence. Nucleotide sequence analysis of p1 indicates that p1 contains L1 and O-family sequences. Interestingly, when the L1 sequence was used as a probe, a discrete band of 5 kb was seen in HindIII-digested DNA from somatic cell hybrids containing human chromosome 21 as the sole human chromosome. The L1 sequence was rearranged and was interrupted by O-family sequence, which was flanked by 6 bp target site duplications. Since all three repetitive sequences are known to act as retroposons, these results imply that there is an integration hot spot on human chromosome 21. The sequence was mapped within 21q11-21.

Key Words human chromosome 21, repetitive sequence, retroposon, mapping

INTRODUCTION

At least 15% of the eukaryotic genome is composed of interspersed sequences (Jelinek and Schmid, 1982). The interspersed repetitive sequences have been classified as SINES (short interspersed repeated sequence) or LINES (long interspersed sequence) (Singer and Skowronski, 1985). The major SINES in the human genome

Received February 15, 1993; Revised accepted May 17, 1993.
*Present address: Division of Molecular Genetics, B452, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA.
**To whom correspondence should be addressed.

is the Alu sequence family. It has a consensus sequence of about 300 bp and is reiterated 500,000 times in the human genome (Weiner *et al.*, 1986). The major human LINES sequence family is L1 (also described as the KpnI family). The L1 family has a consensus sequence of about 6.4 kb and is reiterated at least 10,000 times in the human genome. Some of them are internally rearranged or deleted. In addition to these SINES and LINES, retrovirus-like repetitive sequences have been identified. The O-family sequence is one such repetitive sequence and is considered to be a solitary LTR (long terminal repeat) (Sun *et al.*, 1984; Paulson *et al.*, 1985).

Many of these repetitive sequences are transposable elements which can replicate and insert into new positions in the genome to increase their copy numbers. Transposable elements can also facilitate DNA rearrangement such as inversions and deletions (Jagadeeswaran *et al.*, 1982; Lehrman *et al.*, 1986). Sometimes, transposable elements may cause mutations by interrupting genes (Wallace *et al.*, 1991). Therefore, transposable elements may generate genetic variations in the human genome for natural selection.

Additionally, since a major sequence of human chromosomes consists of repetitive sequences, repetitive sequences may play important roles in chromosome organization. Many single copy DNA fragments have been mapped because of their usefulness for landmarks of each chromosome. However, mapping repetitive sequences in human chromosomes has been difficult precisely because they are reiterated many times, and the distribution of repetitive sequences remains somewhat unclear.

In many mapping efforts, human chromosome 21 has served as a prototype for the human genome. Physical, genetic, and compositional maps of human chromosome 21 have been reported (Petersen *et al.*, 1991; Gardiner *et al.*, 1990a). The distribution of unique sequences, rare restriction enzyme sites, CpG islands and chromosomal break points has been described (Gardiner *et al.*, 1990b).

In order to investigate further the organization of human chromosome 21, we have isolated DNA fragments containing Alu sequence from mouse-human somatic cell hybrids containing human chromosome 21 as the only human chromosome. Here we report the characterization of one of the clones which contains various kinds of repetitive sequences.

MATERIAL AND METHOD

Somatic cell hybrids. The human and rodents cell lines, as well as human/rodent somatic cell hybrids, WA17, 2Fur, ACEM, R2-10, and 21q$^+$ cells were used and their properties have been described previously (Oates and Patterson, 1977; Kozak *et al.*, 1977; Van Keuren *et al.*, 1986; Patterson *et al.*, 1983; Wong *et al.*, 1989; Drabkin *et al.*, 1895).

Southern blot analysis. Chromosomal DNA was isolated according to a method

described by Zsebo *et al.* (1990). DNA (10 μg) was digested with HindIII, developed on an agarose gel, and blotted onto GeneScreenPlus (New England Nuclear) membranes. These blots were prehybridized in a solution containing 50% formamide, 5 × Denhardt's solution (1 × Denhardt's solution = 0.2 g/*l* Ficoll, 0.2 g/*l* polyvinylpyrolidone, 0.2 g/*l* BSA), 0.5% SDS, 100 μg/ml sheared, denatured salmon sperm DNA, 6 × SSC (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate) at 42°C for 2 h, hybridized with ^{32}P-labeled probes in a solution containing 50% formamide, 0.5% SDS, 100 μg/ml sheared, denatured salmon sperm DNA, 6 × SSC at 42°C for 40 h, and washed at a final stringency of 0.1 × SSC, 0.2% SDS, at 65°C.

DNA sequencing. Each fragment was subcloned into pUC118 or 119 and nested deletions were generated by progressive digestion with exonuclease III (Henikoff, 1984). Automated analysis of DNA sequence was performed by the chain-termination method, adopted for ssDNA using Taq dye-primer sequencing kit (Applied Biosystems, Foster City, CA) by means of an ABI 370A DNA sequencer (Applied Biosystems). Sequences were analyzed using software provided by the Genetics Computer Group.

Polymerase chain reaction (PCR). PCR was performed essentially as described by Saiki *et al.* (1988). A set of primers, F2, 5′-CTTAGACATAGATGTCTGCC and R2, 5′-GGTACTTCCATTAAACTACC were synthesized according to the p1 nucleotide sequence at the position betewen 547 and 566 and at the position between 1501 and 1520, respectively. A set of primers which was designed to amplify exon 15 of amyloid protein precursor (APP) was used as internal controls (Levy *et al.*, 1990). The sequences of APP primers are 5′-CCTCATCCAAATGTCCCCGTCATT and 5′-GCCTAATTCTCTCATAGTCTTAATTCCCAC. PCR was carried out in 100 μl of reaction mixtures containing PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin), 20 nmol each dNTP, 100 pmol of each primer, 0.1 μg DNA, and 2.5 units of *Taq* polymerase and was covered with 100 μl of mineral oil. The reaction was carried out for 30 cycles. One cycle was denaturation (94°C, 1 min), annealing (58°C, 1 min), and extension (72°C, 2 min). Initial denaturation was done for 5 min and the final extension was done for 10 min. The PCR products were analyzed on 3% NuSieve GTG agarose gels (FMC Bioproducts, Rockland, ME).

RESULTS

Southern blot analysis of clone p1

In order to isolate genomic sequences encoded by human chromosome 21, the genomic DNA from human-mouse somatic cell hybrid WA17, which contains human chromosome 21 as its only human chromosome, were isolated and cloned into lambda ZAPII vector to construct a genomic library. The genomic library was screened with an Alu sequence, BLUR-8 (Rubin *et al.*, 1980), as a probe by means of plaque hybridization. Positive clones were converted into pBluescript-

SK(−) and were used for further studies. To characterize further, each clones was subcloned and the subcloned fragments were subjected to the Southern blotting hybridization analysis. On this way to identify the location of Alu sequence, we happened to find that one of the clones, p1, contains repetitive sequences in addition to the Alu sequence and that it has a distinctive structure. The restriction map of clone p1 is shown in Fig. 1A. Clone p1 has a size of approximately 4 kb and Alu sequence was found in the indicated XbaI-BglII fragment (Fig. 1A). When the KpnI-XhoI fragment (Fig. 1A, probe C), which does not carry Alu sequence,

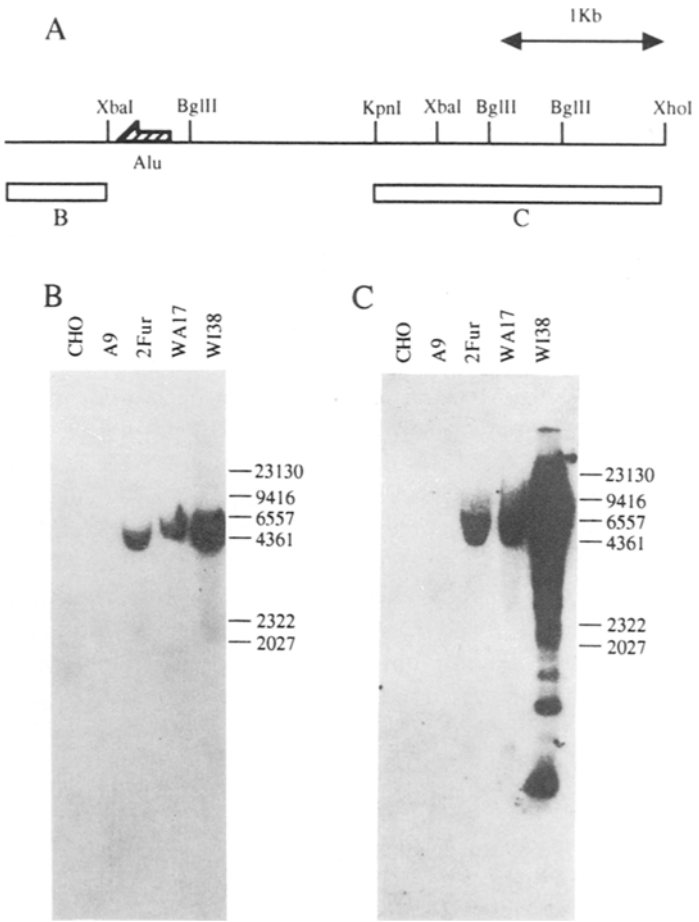


Fig. 1. A: Restriction map of p1. Probe for Southern hybridization are shown as open boxes. An arrow represents the location and orientation of Alu sequence. B: Southern hybridization analysis of clone p1. DNAs from cell lines indicated at the top were digested with HindIII, applied on 0.7% agarose gel, and transferred to a nylon filter. The filter was hybridized with probe B (panel A). C: A same series of Southern hybridization analysis was carried out using probe C.

was used as a probe, Southern blot hybridization analysis of HindIII-digested human DNA revealed a smear suggesting that the fragment contains repetitive sequences different from Alu sequence (Fig. 1C).

Interestingly, a discrete band of 5 kb was seen in HindIII-digested DNA from WA17 or 2Fur cells, which are somatic cell hybrids containing human chromosome 21 or the long arm of human chromosome 21 as their only human chromosome material respectively. The intensity of the band was much stronger than that given by single copy genes (data not shown). These results suggest that the repetitive sequences contained in the KpnI-XhoI fragment are dispersed in human chromosomes in HindIII fragments of different sizes. However, on human chromosome 21, the repetitive sequence appears to exist in HindIII fragments of a single size. These results suggest the possibility of the specific and/or dominant organization of the repetitive sequence on human chromosome 21.

Sequence of clone p1

The complete sequence and the organization of the clone p1 is shown in Figs. 2 and 3. The Alu sequence was found to be located at the position between 608 and 922 and was flanked by 10 bp target site duplications (TSDs). The Alu sequence in clone p1 was 80% homologous to that of consensus sequence (Jurka and Smith, 1988).

The region homologous to L1 repetitive sequence was found in between nucleotide position of 1021 and 4142. This region has an unique structure. The first 2640 bp (nucleotide position 1021 to 3660) is homologous to L1 sequence (nucleotide position 2269 to 4584 of the L1 sequence) in inverted orientation (Skowronski *et al.*, 1988). On the contrary, the second part (nucleotide position 3668 to 4142) was also homologous to L1 sequence (nucleotide position 1024 to 1502) in the forward orientation. The sequence of 7 nucleotides between the two parts was not homologous to the L1 sequence in either orientation. No open reading frame (ORF) was found in this L1 homologous region.

In addition to Alu and L1 sequences, O-family repetitive sequences (Sun *et al.*, 1984) were found at the position between 1076 and 1418. The O-family repeat is flanked by 6 bp TSDs. The O-family sequence is believed to be a solitary LTR-

Fig. 2. Nucleotide sequence of clone p1 and its comparison with repetitive sequences. The p1 sequence is shown on the top and Alu, L1, and O-family sequences are in the second or third lines. Alu indicates the Alu sequence described by Jurka and Smith (1988); O-F, O-family sequence described by Sun *et al.* (1984); L1-R, L1 sequence described by Skowronski *et al.* (1988) in reverse orientation; L1, L1 sequence described by Skowronski *et al.* in forward orientation. Numbers indicate nucleotide positions corresponding the original reports. Dots represent the identical bases and dashes represent deletion. The box indicates the sequence between two regions homologous to the L1 sequence in different orientation. Heavy underlines show target site duplications (TSDs) for the Alu sequence and light underlines show TDSs for the O-family sequence.

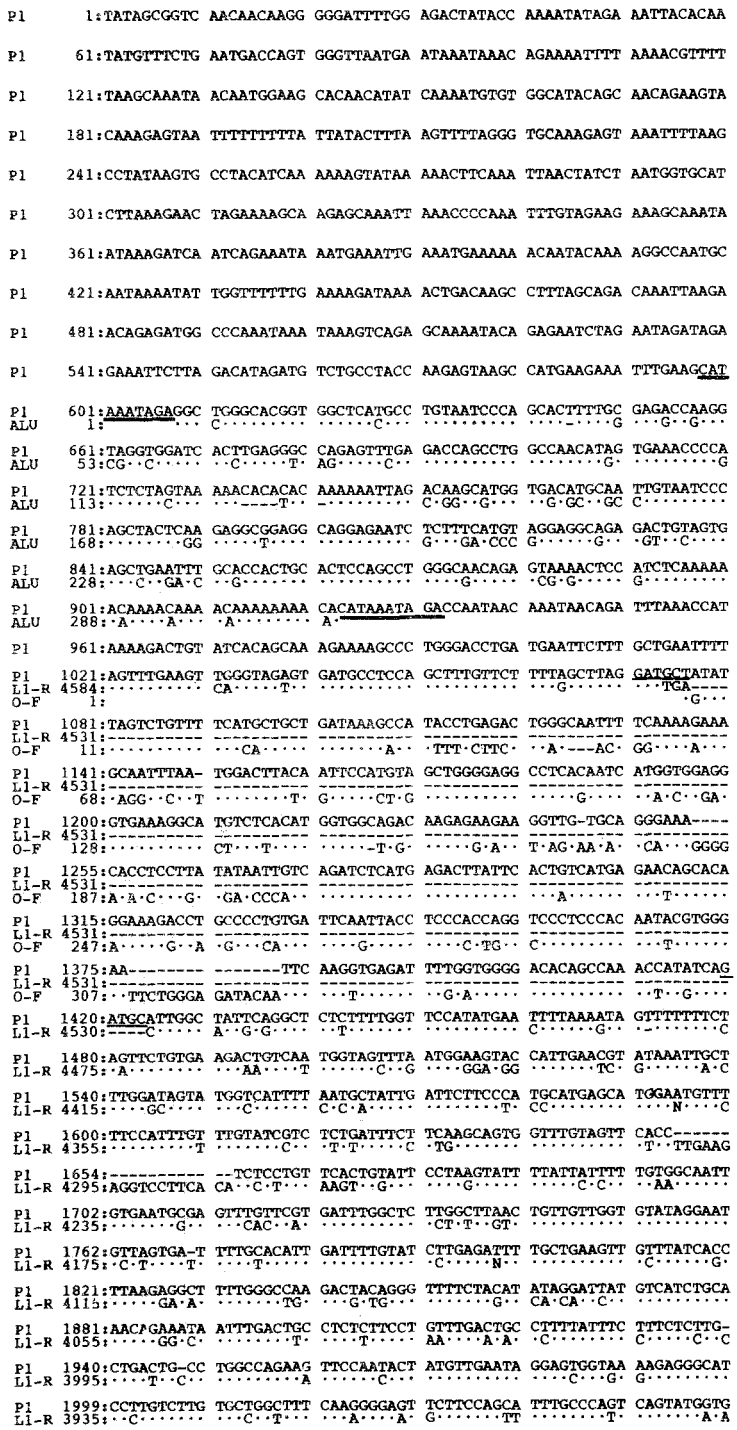
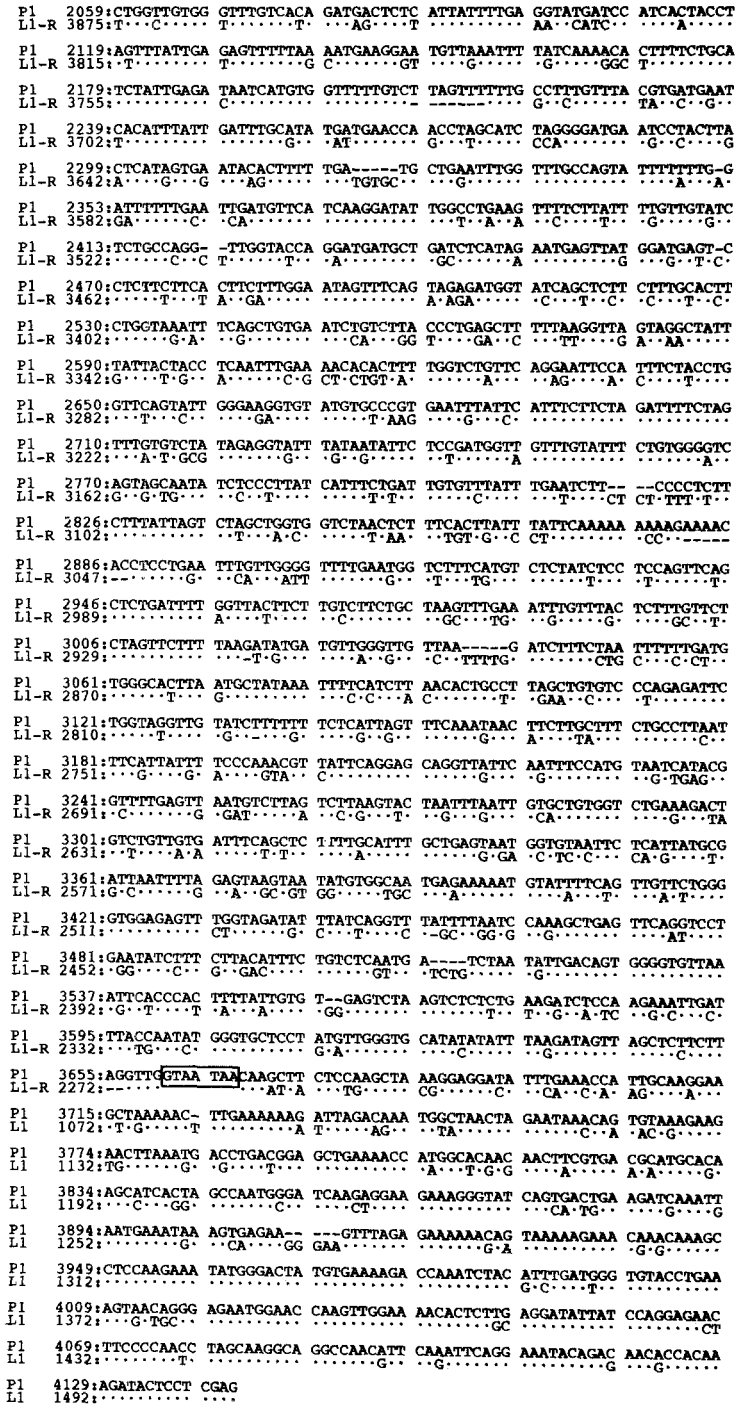


Fig. 2.



like sequence, since it starts with TG and terminates with CA, which is a hallmark of the LTRs (Varmus, 1982). The O-family sequence located in p1 starts with TA and terminates with CA. Deviation from the "TG-CA" rule applicative to the infectious retroviruses has also been seen in the LTRs of a previously described class of human retrovirus-like sequence (Steele *et al.*, 1984; Mager and Henthorn, 1984). The LTRs described by Mager *et al.* also start with TA. It is also interesting to note that the O-family sequence in clone p1 was found within the L1 sequence, because both of the retrovirus-like sequences described by Steele *et al.* and Mager *et al.* were also found within the L1 sequence. These observations may indicate the association of some particular family of retrovirus-like sequence with L1 sequence.

A region between nucleotides 1 and 608 which did not contain any repetitive sequence was AT-rich and the GC content of the region was 30%. It has been suggested that retroposons such as Alu sequence are preferentially inserted into short AT-rich regions (Daniels and Deininger, 1985). This AT-rich region of p1 could be preferred for the insertion of such various kinds of repetitive sequences.

Regional assignment of clone p1

To determine the regional location of clone p1, the location of p1 sequence in somatic cell hybrids containing various regions of human chromosome 21 was analyzed by PCR. A set of primers designed according to the p1 sequence was used. Primers for APP were added to the reaction mixture as internal controls. The amplified fragments are expected to contain both Alu and O-family sequences (Fig. 3). When the DNAs from 2Fur, WA17, R2-10, and $21q^+$ cells, somatic cell hybrids containing 21q11-21, were used as template, fragments with the expected sizes were amplified (Fig. 4A). The fragment was not amplified either from DNAs of hamster or mouse cells or from somatic cell hybrids, $8q^-$ and ACEM. When p1-specific primers were used, only the upper bands were seen (data not shown).

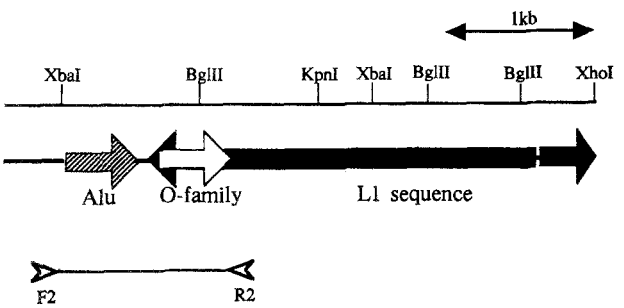


Fig. 3. Schematic diagram of p1. Restriction sites of p1 are shown on the top. Shaded arrow indicates Alu sequence; open arrow, O-family sequence; closed arrow, L1 sequences. The p1 specific primers, F2 and R2, are shown by small open arrows, which are located at both ends of the amplified fragment indicated by thin line.

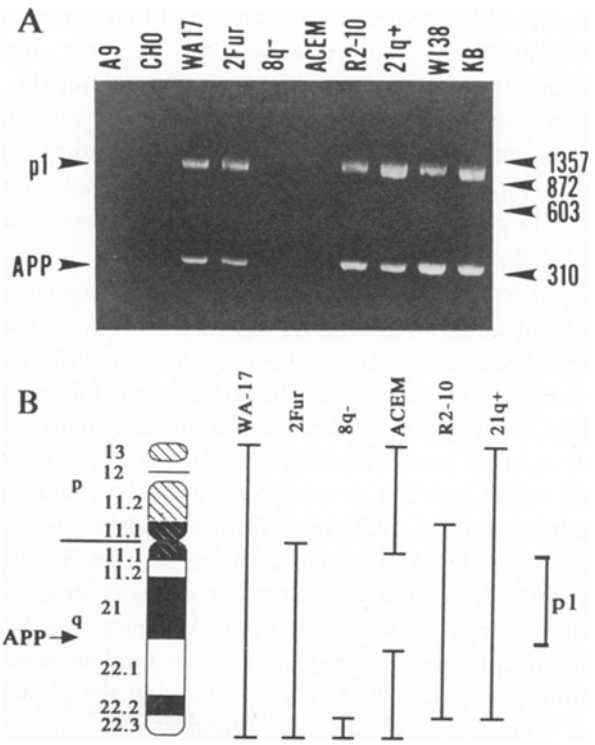


Fig. 4. PCR analysis of human-rodent hybrid cell panel. A: DNAs from cell lines shown on the top of the figure were subjected to PCR analysis with primers specific to p1 and amyloid protein precursor (APP) sequences. Amplified DNAs were visualized by ethdium bromide staining. p1-specific DNA fragments are indicated on the left. The numbers indicate positions for the size markers. B: Chromosomal segments contained in somatic cell hybrids. Idiogramatic representation of chromosome 21 is shown in the right part. The location of APP is shown by arrow. The region in which clone p1 was mapped is shown in the right bracket.

These results indicate that p1 is located within 21q11-21 (Fig. 4B). Since fragments with the expected size were amplified, the structure of the p1 sequence may be conserved among each human genome of these cells.

DISCUSSION

The human genome is composed of two structurally distinctive regions, G-bands (Giemsa positive, or Giemsa dark bands) and R-bands (reverse bands; these are equivalent to Giemsa negative, or Giemsa light bands). It has been suggested that some classes of repetitive sequences are clustered in specific bands, *i.e.*, Alu sequence predominates in the R-bands and L1 in the G-bands (Korenberg and

Rykowski, 1988). In addition, it has been suggested that some interspersed repetitive sequences such as Alu sequence integrate into preferred sites in the chromosomes (Moyzis *et al.*, 1989). Although little is known about the molecular mechanism of the integration of retroposons, the fact that p1 contains three kinds of putative retroposons may indicate that there is an integration hot spot. If this is the case, it is interesting to note that it is likely that p1 is located in the G-band which is believed to be relatively gene poor and which contains relatively few known genes (Gardiner, 1990b).

When the KpnI-XhoI fragment, which contains L1 sequence, was used as a probe, a discrete band was seen in HindIII-digested DNA from somatic cell hybrids containing human chromosome 21 as the only human chromosome. However, Southern blot hybridization analysis of HindIII-digested human DNA resulted in a smear detected by the fragment. These results indicate that one kind of structure would dominantly exist on human chromosome 21. If this is the case, it is difficult to explain how L1 would function as retroposon which can replicate and insert into new positions in the genome. Although further analysis such as Southern blot hybridization analysis of DNAs digested with another restriction enzymes will be necessary to establish firmly the copy number and the exact organization, one possibility is that the region containing the p1 sequence may be amplified at the genome level and the p1 sequence may be exist as tandem repeat. Therefore, it would be interesting to isolate a YAC clone containing the p1 sequence and determine its structure. These experiments are in progress.

The clone p1 contains three kinds of retroposons, Alu, L1, and O-family sequence (Weiner *et al.*, 1986). The function of these retroposons remains unknown. However, there is suggestive evidence that some retroposons may be involved in genome rearrangement (Jagadeeswaran *et al.*, 1982; Lehrman *et al.*, 1986; Henthorn *et al.*, 1986). In particular, the retrovirus-like sequence at the globin locus is supposed to facilitate recombinational events to produce thalassemias (Mager *et al.*, 1985). Alternately, the presence of such repetitive sequences may indicate that the region is particularly susceptible to breakage for the integration of foreign DNA. Precise mapping of p1 and further analysis of this locus may provide further clues to the mechanism involved in chromosomal rearrangements in mammalian cells.

Furthermore, it has been suggested that some classes of repetitive sequences may contribute to gene regulation. In chicken, the CR1 sequence, which share features with retrovirus LTRs (Stumph *et al.*, 1984), is supposed to change the chromosome structure and thus change gene expression (Scott *et al.*, 1987). In addition, it has been shown that CR1 can function as a silencer element leading to the reduction of transcription (Baniahmad *et al.*, 1987). In vertebrate, repetitive elements of the L1 and Alu sequences have also been shown to function as silencers (Laimans *et al.*, 1986; Saffer and Thurston, 1989). p1 sequence has three repetitive sequences, all of which may contribute to gene regulation. Although

further study is needed to substantiate this possibility, this clone may be useful to study not only chromosome organization but also the involvement of repetitive sequences in gene regulation.

Acknowledgment This work was supported by the Grant-in-Aid for Science Research on Priority Areas No. 03224104 from the Ministry of Education, Science and Culture of Japan and NIH grant HD 17449 to Dr. D. Patterson. We appreciate the support of Shin-Etu-Chemicals for these studies. The nucleotide sequence data reported will appear in GenBank under the accession number L05173.

REFERENCES

Baniahmad A, Muller M, Steiner C, Renkawitz R (1987): Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. EMBO J **6**: 2297–2303

Daniels GR, Deininger PL (1985): Integration site preferences of the Alu family and similar repetitive DNA sequences. Nucleic Acids Res **13**: 8939–8954

Drabkin HA, Diaz M, Bradley CM, LeBeau MM, Rowley JD, Patterson D (1985): Isolation and analysis of the 21q+ chromosome in the acute myelogenous leukemia 8;21 translocation: Evidence that c-mos is not translocated. Proc Natl Acad Sci USA **82**: 464–468

Gardiner K, Aissani B, Bernardi G (1990a): A compositional map of human chromosome 21. EMBO J **9**: 1853–1858

Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D (1990b): Analysis of human chromosome 21: Correlation of physical and cytogenetic maps; gene and CpG island distribution. EMBO J **9**: 25–34

Henikoff S (1984): Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–359

Henthorn PS, Mager DL, Huisman TH, Smithies O (1986): A gene deletion ending within a complex array of repeated sequences 3′ to the human β-globin gene cluster. Proc Natl Acad Sci USA **83**: 5194–5198

Jagadeeswaran P, Tuan D, Forget BG, Weissman SM (1982): A gene deletion ending at the midpoint of a repetitive DNA sequence in one form of hereditary persistence of fetal haemoglobin. Nature **296**: 469–470

Jelinek WR, Schmid CW (1982): Repetitive sequences in eukaryotic DNA and their expression. Annu Rev Biochem **51**: 813–844

Jurka J, Smith T (1988): A fundamental division in the Alu family of repeated sequences. Proc Natl Acad Sci USA **85**: 4775–4778

Korenberg JR, Rykowski MC (1988): Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell **53**: 391–400

Kozak CA, Lawrence JB, Ruddle FH (1977): A sequential staining technique for the chromosomal analysis of the interspecific mouse/hamster and mouse/human somatic cell hybrids. Exp Cell Res **105**: 109–117

Laimins L, Holmgren-Konig M, Khoury G (1986): Transcriptional "silencer" element in rat repetitive sequences associated with the rat insulin 1 gene locus. Proc Natl Acad Sci USA **83**: 3151–3155

Lehrman MA, Russell DW, Goldstein JL, Brown MS (1986): Exon-Alu recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. Proc Natl Acad Sci USA **83**: 3679–3683

Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, Van Duinen SG, Bots GT, Luyendijk W, Frangione B (1990): Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Duch type. Science **248**: 1124–1126

Mager DL, Henthorn PS (1984): Identification of a retrovirus-like repetitive element in human DNA. Proc Natl Acad Sci USA **81**: 7510–7514

Mager DL, Henthorn PS, Smithies O (1985): A Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$ thalassemia deletion: comparison to other deletions in the human β-globin gene cluster and sequence analysis of the breakpoints. Nucleic Acids Res **13**: 6559–6575

Moyzis RK, Torney DC, Meyne J, Buckingham JM, Wu JR, Burks C, Sirotkin KM, Goad WB (1989): The distribution of interspersed repetitive DNA sequences in the human genome. Genomics **4**: 273–289

Oates DC, Patterson D (1977): Biochemical genetics of Chinese hamster cell mutants with deviant purine metabolism: characterization of Chinese hamster cell mutants defective in phosphoribosylpyrophosphate amidotransferase and phosphoribosylglycinamide synthetase and an examination of alternatives to the first step of purine biosynthesis. Somatic Cell Genet **3**: 561–577

Patterson D, Jones C, Morse H, Rumsby P, Miller Y, Davis R (1983): Structural gene coding for multifunctional protein carrying orotate phosphoribosyltransferase and OMP decarboxylase activity is located on long arm of human chromosome 3. Somatic Cell Genet **9**: 359–374

Paulson KE, Deka N, Schmid CW, Misra R, Schindler CW, Ruch MG, Kadyk L, Leinwand L (1985): A transposon-like element in human DNA. Nature **316**: 359–361

Petersen MB, Slaugenhaupt SA, Lewis JG, Warren AC, Chakravarti A, Antonarakis SE (1991): A genetic linkage map of 27 markers on human chromosome 21. Genomics **9**: 407–419

Rubin CM, Houck CM, Deininger PL, Friedmann T, Schmid CW (1980): Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature **284**: 372–374

Saffer JD, Thurston SJ (1989): A negative regulatory element with properties similar to those of enhancers is contained within an Alu sequence. Mol Cell Biol **9**: 355–364

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**: 487–491

Scott MJ, Tsai MJ, O'Malley BW (1987): Deoxyribonuclease I sensitivity of the ovomucoid-ovoinhibitor gene complex in oviduct nuclei and relative location of CR1 repetitive sequences. Biochemistry **26**: 6831–6840

Singer MF, Skowronski J (1985): Making sense out of LINES: Long interspersed repeated sequences in mammalian genomes. TIBS **10**: 119–122

Skowronski J, Fanning TG, Singer MF (1988): Unit-length line-1 transcripts in human teratocarcinoma cells. Mol Cell Biol **8**: 1385–1397

Steele PE, Rabson AB, Bryan T, Martin MA (1984): Distinctive termini characterize two families of human endogenous retroviral sequences. Science **225**: 943–947

Stumph WE, Hodgson CP, Tsai MJ, O'Malley BW (1984): Genomic structure and possible retroviral origin of the chicken CR1 repetitive DNA sequence family. Proc Natl Acad Sci USA **81**: 6667–6671

Sun L, Paulson KE, Schmid CW, Kadyk L, Leinwand L (1984): Non-Alu family interspersed repeats in human DNA and their transcriptional activity. Nucleic Acids Res **12**: 2669–2690

Van Keuren ML, Watkins PC, Drabkin HA, Jabs EW, Gusella JF, Patterson D (1986): Regional localization of DNA sequences on chromosome 21 using somatic cell hybrids. Am J Hum Genet **38**: 793–804

Varmus HE (1982): Form and function of retroviral proviruses. Science **216**: 812–820

Wallace MR, Andersen LB, Saulino AM, Gregory PE, Glover TW, Collins FS (1991): A *de novo* Alu insertion in neurofibromatosis type 1. Nature **353**: 864–866

Weiner AM, Deininger PL, Efstratiadis A (1986): Nonviral retroposones: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem **55**: 631–661

Wong C, Kazazian HH Jr, Stetten G, Earnshaw WC, Van Keuren ML, Antonarakis SE (1989) Molecular mechanism in the formation of a human ring chromosome 21. Proc Natl Acad Sci USA **86**: 1914–1918

Zsebo KM, Williams DA, Geissler EN, Broudy VC, Martin FH, Atkins HL, Hsu RY, Birkett NC, Okino KH, Murdock DC, Jaconsen FW, Langley KE, Smith KA, Takeishi T, Cattanach BM, Gall SJ, Suggs SV (1990): Stem cell factor is encoded at the SI locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. Cell **63**: 213–224