[CANCER RESEARCH 52, 643–645, February 1, 1992]

Disruption of the *APC* Gene by a Retrotransposal Insertion of L1 Sequence in a Colon Cancer[1]

Yoshio Miki, Isamu Nishisho, Akira Horii, Yasuo Miyoshi, Joji Utsunomiya, Kenneth W. Kinzler, Bert Vogelstein, and Yusuke Nakamura[2]

Department of Biochemistry, Cancer Institute, 1-37-1, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan [Y. Mik., I. N., A. H., Y. Miy., Y. N.]; Second Department of Surgery, Hyogo College of Medicine, 1-1, Mukogawa-cho, Nishinomiya, Hyogo 663, Japan [Y. Mik., J. U.]; and Molecular Genetics Laboratory, The Johns Hopkins Oncology Center, Baltimore, Maryland 21231 [K. K., B. V.]

ABSTRACT

The *APC* gene is responsible for familial adenomatous polyposis and is considered to be a tumor suppressor gene associated with development of sporadic colorectal tumors. Here we report the disruption of the *APC* gene caused by somatic insertion of a long interspersed repetitive element (LINE-1 sequence) into the last exon of the *APC* gene in a colon cancer. The inserted sequence was composed of a 3′ portion of the LINE-1 consensus sequence and nearly 180 base pairs of polyadenylate tract. Furthermore, since an 8-base pair target site duplication was observed, retrotranscriptional insertion of an active LINE-1 sequence is suspected as the cause of this insertion event. This is the first report of the disruption of a tumor suppressor gene caused by somatic insertion of a mobile genetic element.

INTRODUCTION

The mechanism of somatic inactivation of putative tumor suppressor genes such as the *RB* (retinoblastoma), *WT* (Wilms' tumor), *DCC* (deleted in colorectal carcinomas), *p53*, and *MCC* (mutated in colorectal cancers) genes have thus far been detected as point mutation, chromosomal loss, intragenic deletion, and/or chromosomal rearrangement (1–6). We and others have recently identified a gene responsible for familial polyposis coli, which was designated as the *APC* gene (7–10). Familial polyposis coli is one of the most common autosomal-dominant diseases which lead to cancer predisposition (11). Affected patients usually develop hundreds to thousands of benign adenomatous polyps in colon, one or some of which will progress to colorectal cancer if not surgically treated. Furthermore, somatic mutations of the *APC* gene were also detected in sporadic colorectal cancers. Hence this gene is considered to be a tumor suppressor gene and probably plays a significant role in the initiation of sporadic colorectal tumors (8).

During our search for somatic mutations in the *APC* gene, we examined over 150 cases of colorectal cancers by Southern blotting using cDNA[3] clones from the *APC* gene as probes. Here we report one case in which the *APC* gene was disrupted by a somatic insertion of a mobile genetic element.

MATERIALS AND METHODS

Samples. Tumors and their corresponding normal mucosal tissues were obtained at surgery from 150 patients with colorectal cancers. Immediately after removal, tissue specimens were rinsed in phosphate-buffered saline, frozen in liquid nitrogen, and stored at −70°C. Routine histopathological diagnoses were performed for all tumors.

Southern Blot Analysis. Frozen samples were ground to a very fine powder in liquid nitrogen, transferred to a 15-ml tube, and suspended in 4 ml of lysis buffer. High-molecular-weight DNAs were extracted in phenol/chloroform/isoamyl alcohol as described by Sato *et al.* (12). DNA from each sample was digested with an appropriate restriction enzyme, electrophoresed in a 0.8% agarose gel, and transferred to a nylon membrane (Pall Biodyne, New York) with 0.1 N NaOH/0.1 M NaCl. The membrane was washed in 2× standard saline-citrate and fixed by UV cross-linking. Prehybridization was performed in 7% PEG 8000/10% sodium dodecyl sulfate at 65°C for 2 h. Hybridization was done at 65°C for 16–24 h in the same solution with a probe labeled with [^{32}P]dCTP by the random priming method (13). After hybridization, the membrane was washed twice in 0.1× standard saline-citrate/0.1% sodium dodecyl sulfate at 65°C for 15 min and then exposed to Kodak XAR film at −70°C.

Isolation of the Fragment Containing the Insertion. A genomic phage library was constructed from colon cancer DNA. The DNA was partially digested with *Sau*3AI. After fractionated by sucrose density gradient centrifugation, DNA was partially filled in, and ligated with Lambda FIX vector, which was digested with *Xho*I and partially filled in (Stratagene, San Diego, CA). After packaging and transfection, the library was plated and screened with cDNA clones from the *APC* gene.

DNA Sequencing. DNA sequencing was done by a modified T7 polymerase method as described by Del Sal *et al.* (14).

RESULTS

Detection of Insertional Mutation in a Colon Cancer. We examined 150 cases of colorectal cancers by Southern blotting using cDNA clones from the *APC* gene as probes. In one case, we found a 4.1-kilobase extra *Bgl*II fragment in addition to the normal 3.4-kilobase *Bgl*II band. This 4.1-kilobase fragment was not seen in 150 normal controls (Fig. 1). DNAs from this tumor and from normal colonic tissue surrounding the tumor were further compared following digestion with other enzymes (*Msp*I, *Hin*dIII, *Eco*RI or *Pst*I) (Fig. 1). All enzymes tested showed extra bands, which were approximately 750 base pairs larger than their normal counterparts when DNA was hybridized with a probe containing *APC* nucleotides 4425 to 5592. These results indicated the possibility of somatic insertion of a 750-base pair fragment into the *APC* gene in this cancer.

Characterization of the Inserted Sequence. To further characterize the inserted sequence, we have screened a genomic phage library constructed from the DNA of this colon cancer. Five positive clones were isolated and two of them contained a 4.1-kilobase *Bgl*II fragment presumptively containing the insertion. A 4.1-kilobase *Bgl*II fragment was subcloned into a pBluescript vector and sequenced (Figs. 2 and 3). As shown in Fig. 3*a*, the *APC* gene was disrupted at codon 1526. The inserted sequence was highly homologous to the L1 consensus sequence (15, 16). The first fifth [nucleotides 1 to 150] of the insertion was almost identical [2-base pair difference] to the inverse complement of the L1 consensus sequence [nucleotides 5697 to 5548 (Fig. 3)]. The central part of the insertion (nucleotides 151 to 531) was highly homologous to the 3′ portion of the L1 consensus sequence (nucleotides 5778 to 6161, more than 98% homology). The last fifth of the insertion (nucleotides 532 to 750) was composed of a polyadenylation signal and subsequent 180-base polyadenylate tract. Furthermore, dupli-

Received 9/4/91; accepted 11/13/91.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported in part by a Grant-in-Aid from the Ministry of Education, Culture and Science, Japan, and Grants for the Vehicle Racing Commemorative Foundation, the Clayton Fund, and National Cancer Institute.

[2] To whom requests for reprints should be addressed.

[3] The abbreviation used is: cDNA, complementary DNA.

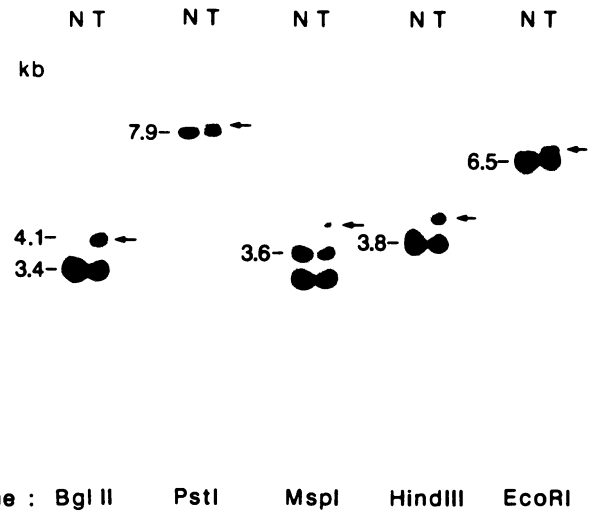

Fig. 1. Southern blot analysis of the *APC* gene from normal tissue (*N*) and colon cancer (*T*) of the same patient. *Abscissa*, restriction endonucleases used; *Arrows*, extra bands in colon cancer which were approximately 750 base pairs larger than those of their normal counterparts.

cation of an AT-rich 8 nucleotides (GAATAATG) of the *APC* sequence was observed at the target site (*underline* in Fig. 3*a*) without deletion of any *APC* sequence (Fig. 3*b*).

DISCUSSION

L1 sequences are up to 7 kilobases long, flanked by target site duplications of various length. Unlike retroviral transposons, they have no terminal repeats, but have A-rich regions at their 3′ ends, usually preceded by a polyadenylation signal. Many L1 sequences have two long open reading frames (ORF-1, ORF-2) (17), one of which (ORF-2) contains a region which is highly homologous to retroviral reverse transcriptase and RNase H (18, 19). These features indicate that L1 elements may be "nonviral retrotransposons" (16, 19), which transpose through an RNA intermediate with reverse transcriptase activity, similar to a postulated mechanism generating pseudogenes (20). However, because most L1 elements are heterogeneously truncated from their 5′ end [up to 95% in mammals (19)], it is extremely difficult to identify a structure or sequence which may still function as a transposon. In the case described here, the insertional sequence has a long stretch of polyadenylate, and it was therefore suspected that the functional L1 sequence was transcribed and integrated into the *APC* gene by a retrotranscriptional mechanism. In addition, the evidence of an 8-base pair target site duplication, since this is commonly found in class II retrotransposons (16), supports this speculation.

The biological function of L1 elements is unclear, but several cases indicating its mobility have been reported. For example *de novo* germ line insertions of L1 sequences occur in the factor VIII gene sometimes resulting in hemophilia A (21, 22). In rats, L1 insertion into the *Igh* (immunoglobulin heavy chain) and *Mlvi-2* (Moloney leukemia virus integration 2) loci were also reported (23). Furthermore, the somatic insertion of an L1 sequence into an intron of the *myc* locus in a human breast carcinoma (24) or the 5′ flanking region of the c-*myc* gene in canine transmissible venereal tumor (25) was reported. It is interesting that somatic insertions described in the present report and Ref. 24 involved an inversion of part of the inserted L1 sequences, while the germ line insertions (21, 22) had the expected colinear organization. Although the effect of the L1 insertion into or upstream of the myc gene was not clear, the *APC* insertion described above can be expected to dramatically alter the gene product, inasmuch as it occurred in an exon in middle of the gene. It will be of interest to determine the frequency at which L1-mediated inactivation of tumor suppressor genes occurs in somatic cells and whether inherited or environmental factors influence this frequency.

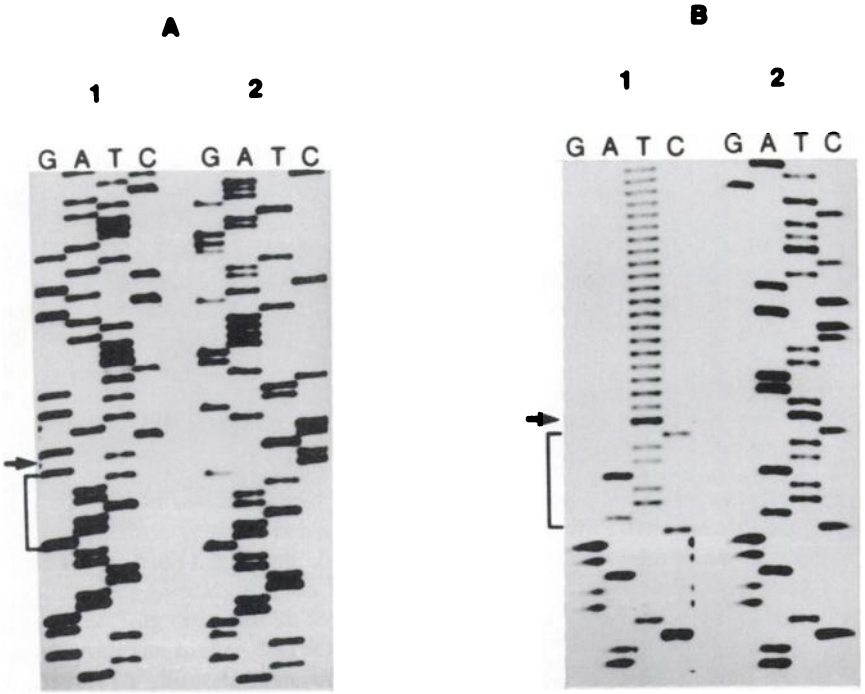

Fig. 2. Sequence analysis of the target site. The 4.1-kilobase *Bgl*II fragment containing the insertion was subcloned and pertinent clones were sequenced (17). *A*, 5′ end of the insertion (*A1*) (*arrow*) and the normal counterpart (*A2*). *B*, 3′ end of the insertion (*arrow*) (*B1*), preceded by polyadenylate and the normal counterpart (B2). Eight-base pair sequences in *brackets* (*A1*, *B1*) indicate the target site duplication.

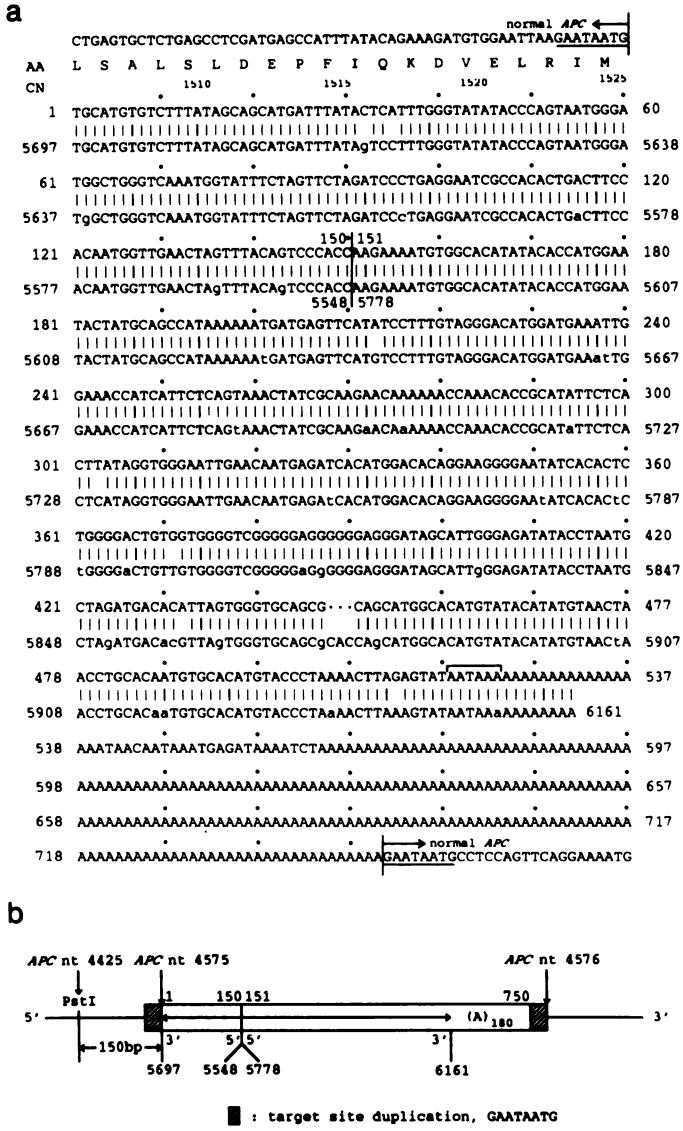

Fig. 3. Comparison of the inserted sequences and L1 consensus sequences. *a*, nucleotide sequence comparison. The inserted sequence is shown above, and the L1 consensus sequences below, numbered as presented in Ref. 5. *Lowercase letters* in the L1 consensus sequences indicate degenerate bases. *Connecting lines* between sequences indicate nucleotide homology. There are only 7 single base pair differences and a 3-base pair deletion relative to the L1 sequence (nucleotides 6057 to 6077). *Two large vertical lines* mark the borders of normal *APC* sequences and the inserted sequences. Amino acids (*AA*) and codon numbers (*CN*) of the *APC* gene are presented below the nucleotide sequences. *Underlines*, target site duplication; *bracket*, polyadenylation signal. *b*, schematic diagram of the L1 insertion into the last exon of the *APC* gene. *Open box*, insertion. The insertion is flanked by an 8-base pair target site duplication shown in *hatched boxes*. Nucleotides 1 to 150 in the insertion are similar to an inverse complement of the L1 sequence (nucleotides 5697 to 5548). The next sequences are homologous to a 3′ portion of the L1 consensus sequence (nucleotides 5778 to 6161) and are followed by a polyadenylation signal and subsequent 180-nucleotide tract of A residues.

ACKNOWLEDGMENTS

We thank Dr. Yoshiyuki Sakaki for helpful advice and Kumiko Koyama and Kiyoshi Noguchi for technical assistance.

REFERENCES

1. Horowitz, J. M., Yandell, D. W., Park, S-H., Canning, S., White, P., Buchkovich, K., Harlow, E., Weinberg, R. A., and Dryja, T. P. Point mutational inactivation of the retinoblastoma antioncogene. Science (Washington DC), *243:* 937–940, 1989.
2. Hashimoto, T., Takahashi, R., Yandell, D. W., Xu, H-J., Hu, S-X., Gunnell, S., and Benedict, W. F. Characterization of intragenic deletions in two sporadic germinal mutation cases of retinoblastoma resulting in abnormal gene expression. Oncogene, *6:* 463–469, 1991.
3. Haber, D. A., Buckler, A. J., Glaser, T., Call, K. M., Pelletier, J., Sohn, R. L., Douglass, E. C., and Housman, D. E. An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. Cell, *61:* 1257–1269, 1990.
4. Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. *p53* mutations in human cancers. Science (Washington DC), *253:* 49–53, 1991.
5. Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W., and Vogelstein, B. Identification of a chromosome 18q gene that is altered in colorectal cancers. Science (Washington DC), *247:* 49–56, 1990.
6. Kinzler, K. W., Nilbert, M. C., Vogelstein, B., Bryan, T. M., Levy, D. B., Smith, K. J., Preisinger, A. C., Hamilton, S. R., Hedge, P., Markham, A., Carlson, M., Joslyn, G., Groden, J., White, R., Miki, Y., Miyoshi, Y., Nishisho, I., and Nakmura, Y. Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science (Washington DC), *251:* 1366–1370, 1991.
7. Kinzler, K. W., Nilbert, M. C., Su, L-k., Vogelstein, B., Bryan, T. M., Levy, D. B., Smith, K. J., Preisinger, A. C., Hedge, P., McKechnie, D., Finniear, R., Markham, A., Groffen, J., Boguski, M. S., Altschul, S. F., Horii, A., Ando, H., Miyoshi, Y., Miki, Y., Nishisho, I., and Nakamura, T. Identification of *FAP* locus genes from chromosome 5q21. Science (Washington DC), *253:* 661–665, 1991.
8. Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyoma, K., Utsunomiya, J., Baba, S., Hedge, P., Markham, A., Krush, A. J., Petersen, G., Hamilton, S. R., Nilbert, M. C., Levy, D. B., Bryan, T. M., Preisinger, A. C., Smith, K. J., Su, L-K., Kinzler, K. W., and Vogelstein, B. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science (Washington DC), *253:* 665–669, 1991.
9. Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J. P., Warrington, J., McPherson, J., Wasmuth, J., Paslier, D. L., Abderrahim, H., Cohen, D., Leppert, M., and White, R. Identification and characterization of the familial adenomatous polyposis coli gene. Cell, *66:* 589–600, 1991.
10. Joslyn, G., Carlson, M., Thliveris, A., Albertsen, H., Gelbert, R., Samowitz, W., Groden, J., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J. P., Warrington, J., McPherson, J., Wasmuth, J., Paslier, D. L., Abderrahim, H., Cohen, D., Leppert, M., and White, R. Identification of deletion mutations and three new genes at the familial polyposis locus. Cell, *66:* 601–613, 1991.
11. Utsunomiya, J., and Lynch, H. T. (eds.). Hereditary colorectal cancer. Tokyo: Springer-Verlag, 1990.
12. Sato, T., Tanigami, A., Yamakawa, K., Akiyama, F., Kasumi, F., Sakamoto, G., and Nakamura, Y. Allelotype of breast cancer: cumulative allele losses promote tumor progression in primary breast cancer. Cancer Res., *50:* 7184–7189, 1990.
13. Feinberg, A. P., and Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragment to high specific activity. Anal. Biochem., *132:* 6–13, 1983.
14. Del Sal, G., Manfioletti, G., and Schneider, C. The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. Biotechniques, *7:* 514–521, 1989.
15. Scott, A. S., Schmeckpeper, B. J., Abdelrazik, M., Comey, C. T., O'Hara, B., Rossiter, J.p., Cooley, T., Heath, P., Smith, K. D., and Margolet, L. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics, *1:* 113–125, 1987.
16. Fanning, T. G., and Singer, M. F. LINE-1: a mammalian transposable element. Biochim. Biophys. Acta, *910:* 203–212, 1987.
17. Hattori, M., Kuhara, S., Takenaka, O., and Sakaki, Y. L1 family of repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transcriptase-related protein. Nature (Lond.), *321:* 625–628, 1986.
18. Xiong, Y., and Eickbush, T. H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J., *9:* 3353–3362, 1990.
19. Schwarz-Sommer, Z., Leclercq, L., Gobel, E., and Saedler, H. Cin4, an insert altering the structure of the *A1* gene in *Zea mays*, exhibits properties of nonviral retrotransposons. EMBO J., *6:* 3873–3880, 1987.
20. Sharp, P. A. Conversion of RNA to DNA in mammals: Alu-like elements and pseudogenes. Nature (Lond.), *301:* 471–472, 1983.
21. Kazazian Jr, H. H., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G., and Antonarakis, S. E. Haemophilia A resulting from *de novo* insertion of L1 sequences represents a novel mechanism for mutation in man. Nature (Lond.), *332:* 164–166, 1988.
22. Woods-Samuels, P., Wong, C., Mathias, S. L., Scott, A. F., Kazazian, Jr,H. H., and Antonarakis, S. E. Characterization of a nondeleterious L1 insertion in an intron of the human factor VIII gene and further evidence of open reading frames in functional L1 elements. Genomics, *4:* 290–296, 1989.
23. Economou-Pachnis, A., Lohse, M. A., Furano, A. V., and Tsichlis, P. N. Insertion of long interspersed repeated elements at *Igh* (immunoglobulin heavy chain) and *Mlvi-2* (Moloney leukemia virus integration 2) loci of rats. Proc. Natl. Acad. Sci. USA, *82:* 2857–2861, 1985
24. Morse, B., Rotherg, P. G., South, V. J., Spandorfer, J. M., and Astrin, S. M. Insertional mutagenesis of the *myc* locus by a LINE-1 sequence in a human breast carcinoma. Nature (Lond.), *333:* 87–90, 1988.
25. Katzir, N., Rechavi, G., Cohen, J. B., Unger, T., Simoni, F., Segal, S., Cohen, D., and Givol, D. "Retroposon" insertion into the cellular oncogene c-*myc* in canine transmissible venereal tumor. Proc. Natl. Acad. Sci. USA, *82:* 1054–1058, 1985