Detection of *APC* gene deletion by double competitive polymerase chain reaction in patients with familial adenomatous polyposis

MAHO TAKAHASHI1,3, MARIKO KIKUCHI1,3, NAGANARI OHKURA1, HIROKO YAGUCHI1, YUKO NAGAMURA1, SUMIKO OHNAMI2, MINEKO USHIAMA2, TERUHIKO YOSHIDA2, KOKICHI SUGANO4, TAKEO IWAMA5, SHINJI KOSUGI6 and TOSHIHIKO TSUKADA1

1Tumor Endocrinology Project and 2Genetics Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045; 3Department of Ophthalmology, Teikyo University School of Medicine, 2-11-1 Kaga, Itabashi-ku, Tokyo 173-8605; 4Oncogene and Cancer Prevention Unit, Tochigi Cancer Center Research Institute, 4-9-13 Younan, Utsunomiya, Tochigi 320-0834; 5Kyoudo Hospital, Sasaki Institute, 2-2 Kandasurugadai, Chiyoda-ku, Tokyo 101-0062; 6Departments of Clinical Genetics and Medical Ethics, Kyoto University Graduate School of Medicine, 54 Shogoinkawahara-cho, Sakyo-ku, Kyoto 606-8397, Japan

Received February 3, 2006; Accepted April 19, 2006

Abstract. Familial adenomatous polyposis (FAP) is an autosomal dominant familial cancer syndrome caused by germline mutations of the tumor suppressor adenomatous polyposis coli (*APC*) gene. Heterozygous *APC* mutations have been identified in the majority of classical FAP patients who develop more than 100 colorectal adenomas. However, classical FAP patients often fail to display germline *APC* mutations detectable by routine mutation analysis. These apparently mutation-negative cases may be caused by heterozygous large genomic deletions. In the present study, FAP patients who showed no *APC* germline mutation detectable by the protein truncation assay and direct sequencing of protein coding exons were screened for *APC* gene deletion by a gene dosage assay based on double competitive polymerase chain reaction. Gene dosage measurements within exon 15 of the *APC* gene identified two patients with gene deletion and one with possible gene duplication among 41 apparently mutation-negative patients. The deleted sequences in the two patients were determined by fine gene dosage mapping around the *APC* gene and nucleotide sequencing of the deletion breakpoints. They were ~435-kilobase pair (kb) and 737-kb regions including the whole *APC* gene and flanked by a 4-base pair repeat and LINE-1 repetitive sequences, respectively. The chimeric LINE-1 element created at the breakpoint in the latter case also contained a short sequence derived from another LINE-1 element, suggesting a complex unequal homologous recombination event. These findings indicate that this gene dose assay is a useful technique to detect large gene deletions of the *APC* gene and to determine their genomic breakpoints.

Introduction

Familial adenomatous polyposis (FAP) (MIM# 175100) is an autosomal dominant familial cancer syndrome characterized by an early onset of multiple adenomatous polyps of the colon that can progress to malignant forms (1). This disorder is caused by germline mutations of the adenomatous polyposis coli (*APC*) gene, the 15 protein-coding exons of which are distributed along a 100-kilobase pair (kb) region on chromosome 5q21 and encode ~10 kb mRNA that is translated into a tumor suppressor protein of 2843 amino acids (2,3). Various heterozygous germline mutations of the *APC* gene have been identified in the majority of classical FAP patients who develop more than 100 colorectal adenomas and early-onset colorectal carcinoma as well as other features including retinal pigment epithelial hypertrophy. However, 20-50% of classical FAP patients displayed no detectable germline mutations when examined by polymerase chain reaction (PCR)-based analytical methods such as single strand conformational analysis, protein truncation assay (PTT) and direct DNA sequencing of PCR-amplified exons (4-9). Although these mutation-negative cases may represent disease entities distinct from that caused by *APC* gene mutation, it is also likely that routine analysis may have missed mutations in the untested part of the *APC* gene or large genomic deletions (6,10).

In the routine PCR-based mutation screening, genomic deletions are often masked by the presence of the normal allele in the heterozygous state. Several different analytical techniques have been developed to detect heterozygous gene deletions and applied to DNA diagnosis of familial cancer syndromes (11). We previously demonstrated that a gene
A dose assay based on double competitive PCR was useful to determine a large germline deletion of the \textit{MEN1} gene in a family with multiple endocrine neoplasia type 1 (12,13). We partly modified this technique and applied it to detect germline \textit{APC} gene deletion in apparently mutation-negative FAP patients. We here demonstrate that this method is useful to identify large genomic deletion and genomic breakpoints in the \textit{APC} gene region.

\textbf{Materials and methods}

\textit{Subjects.} Patients were diagnosed as having colorectal polyposis by endoscopic examinations at individual hospitals and registered in a multi-institutional research program for hereditary colorectal cancer. Written informed consent was obtained for the \textit{APC} gene mutation analysis, which was conducted at the National Cancer Center Research Institute.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Gene dose assay. A, construction of internal standards by the original method. In the first PCR, two half portions of the internal standard were produced with primer pairs a/b or c/d and with genomic DNA (G) as a template. Primers b and c had 2-bp deletion (shown by a triangle) in otherwise normal sequences and were complementary to each other within their overlapping sequence. In the second PCR, a full-length internal standard (IS) was generated by extension and amplification of the two half portions with primers a and d. Horizontal arrows and solid lines represent PCR primers and DNA sequences, respectively. B, construction of internal standards by the modified method. Shorter internal standards were produced by extension of two synthetic primers a' and b' that were designed to have short complementary sequences in the 3' ends, where 2-bp deletion was introduced. C, in the gene dose assay, genomic DNA (G') and internal standard (IS') sequences for a test region were amplified with primers e' and f' (left), and genomic DNA (G') and internal standard (IS') sequences for the reference gene were amplified with primers e'' and f'' (right). The PCR products of internal standards i's and i's' were 2 bp smaller than those of genomic sequences g' and g', respectively. D, representative electropherogram of PCR products from test genomic (g') and internal standard (i's') sequences (left) and the reference genomic (g'') and internal standard (i's'') sequences (right). Each PCR product generated two fluorescent peaks representing 3'-end deoxyadenylated and non-deoxyadenylated forms. Nucleotide numbers below were derived from the GeneScan-500 TAMRA size marker. Gene dose quotient was calculated by the following formula: Gene dose quotient = (g'/i's)/(g''/i's'').}
\end{figure}
Forty-one apparently unrelated colorectal polyposis patients who had no germline APC gene mutation detectable by PTT (14) and DNA sequencing (15) of the whole protein-coding exons were studied. This study was approved by the Ethics Review Committees of the National Cancer Center, Tochigi Cancer Center, Kyoudo Hospital and Kyoto University Graduate School of Medicine.

Construction of internal standard DNA. The gene dose assay was based on quantitative PCR and employed internal standard DNA for competitive PCR, which had a 2-base pair (bp) deletion within otherwise normal sequences. The internal standards for the target sequence within exon 15 of the APC gene (Ex15) and some other target sequences around the gene and that for the hMLH1 gene, which was used as a reference sequence, were prepared as described previously (13). Briefly, the internal standards were produced by two-step amplification (Fig. 1A and Table I), and introduced into a plasmid vector. The plasmid clones verified to have a desired sequence were selected and cut with restriction enzymes that cleaved only the vector DNA. This linearization of internal standard DNA was essential for high reproducibility of this gene dose assay. This original method provided relatively long internal standard DNA (>300 bp).

In the present study, we additionally adopted a modified method for internal standard preparation, which provided relatively short internal standard DNA (~100 bp). In this new method, synthetic PCR primers (Fig. 1B, Table I) were designed to have a short complementary sequence in their 3'-ends, where 2-bp deletion was introduced into an otherwise normal sequence, and were simply extended by PCR under the following conditions: 1 min at 95°C for 1 cycle; 1 min at 95°C for denaturing, 1 min at 51, 55, 59 or 63°C for annealing, 1 min at 72°C for extension for 12 cycles, 5 min at 72°C for 1 cycle in the reaction mixture containing Taq DNA polymerase (rTaq, Takara, Japan), primers (50 μM each) and other components recommended by the enzyme supplier. The PCR products, purified with QIAquick PCR purification kit (Qiagen Inc., Valencia, CA), were diluted to 1:10-10 and used as internal standard DNA. The genomic sequence data were obtained from internet database (http://www.ncbi.nlm.nih.gov/). The accession numbers of a chromosome 5q clone CTC-554D6 containing the APC gene and 3q clones RP11-129K12 and RP11-491D6 containing the hMLH1 gene were AC008575.7, AC011816.17, respectively.

Gene dose assay. Genomic DNA was prepared either directly from whole blood with a QIAamp blood kit (Qiagen Inc.) or from isolated peripheral blood cells by the ordinary phenol-chloroform method or a SepaGene Kit (Sanko Junyaku, Tokyo, Japan). Double competitive PCR was described previously (Fig. 1C) (13). Some PCR primers for a test or the reference sequence are shown in Table I. One of the primers of each primer pair was labeled with fluorescein. Each sample was amplified in triplicate. The relative amounts of PCR products derived from the genomic and internal standard DNA in each reaction tube were evaluated by measuring fluorescence intensity with an automated capillary sequencer as described previously (Fig. 1D) (13).

Gene dose quotient was calculated from the amounts of four PCR products: Gene dose quotient = (PCR product of test genomic sequence ÷ PCR product of internal standard for test sequence) + (PCR product of reference genomic sequence ÷ PCR product of internal standard for reference sequence). The mean value obtained from triplicate PCR was used to calculate each gene dose quotient. Gene dose was normalized to the normal control DNA by using the following formula: Gene dose = (gene dose quotient of test sample ÷ gene dose quotient of normal control sample) x 2.

Microsatellite analysis. Genomic DNA was amplified by PCR with a pair of primers, one of which was labeled with fluorescein. The size of the PCR product was estimated by electrophoresis with an automated capillary sequencer as described previously (12). The primer sequences and PCR conditions were obtained from the Genome Database (http://gdbwww.gdb.org/).

Sequencing analysis of deletion breakpoints. In one patient (KT25) having a gene deletion, PCR primers were prepared at an ~1-kb interval within the region where the deletion breakpoints were located by the gene dose mapping (Fig. 4A). Genomic DNA of the patient or a normal volunteer was amplified by PCR with every possible combination of upstream and downstream primers under the following conditions: 1 min at 95°C for 1 cycle; 30 sec at 95°C, 3 min at 68°C for 35 cycles; 3 min at 68°C for 1 cycle. PCR products were separated by electrophoresis on 1% agarose gel, and visualized by ethidium bromide staining. PCR products specific to the patient were analyzed by direct nucleotide sequencing with primers used for PCR.

In another patient (KT64), the deletion breakpoints were considered to be within long repetitive sequences. Therefore, long-range PCR was conducted with primers used for gene dose assay at sites where both alleles have proven to be retained. Long-range PCR was performed in a 50 μl reaction mixture containing 0.5 unit Takara LA Taq (Takara, Japan), 25 ng genomic DNA, 50 μM each primer and components recommended by the enzyme supplier under the following conditions: 1 min at 94°C for 1 cycle, 10 sec at 98°C and 15 min at 68°C for 30 cycles, and extension of 72°C for 10 min. PCR products were separated by electrophoresis on 0.8% agarose gel, and visualized by ethidium bromide staining. PCR products specific to the patient were subjected to restriction enzyme analysis. The restriction fragments considered to have the boundary sequences were introduced into a plasmid vector pBluescriptSK(+) (Stratagene, La Jolla, CA) and analyzed by nucleotide sequencing.

Results

Gene dose analysis of apparently mutation-negative FAP patients. Gene dose at the test sequence Ex15, which was located near the 5'-end of exon 15 of the APC gene, was measured in 6 normal subjects and 41 apparently mutation-negative polyposis patients. All normal subjects and all but three polyposis patients displayed gene dose values distributed around 2, suggesting that this group represents subjects having two copies of the Ex15 sequence and that deviation of each
Table I. PCR primers for preparing internal standards and for gene dose assay.

<table>
<thead>
<tr>
<th>Ex15</th>
<th>hMLH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>a: (127178) CGAAGTACAAGGATGCAAT</td>
<td>a: (94389) ACTTGGCTACCAGGACCTTGGC</td>
</tr>
<tr>
<td>b: (127401) ATTATCATCATGTATTGGTGTC</td>
<td>b: (94228) CAGAGGCTGAGAACATGCT</td>
</tr>
<tr>
<td>c: (127380) CACCAATACATGATGATAATAG</td>
<td>c: (94247) ATGCAATTCTCAGGCTCTGAG</td>
</tr>
<tr>
<td>d: (127624) CAGTGTTGGAGATCTGCAA</td>
<td>d: (94200) TGGGGTTGCTGGAAGTAGGTC</td>
</tr>
<tr>
<td>e: (127319) AGGCATCTCATCGTAGTAAGCA</td>
<td>e: (94359) GGGGAGATGGTTAAATCCAA</td>
</tr>
<tr>
<td>f: (127486) TTAGTGAAGAGGAGCTGGTAA</td>
<td>f: (94179) TCTTATCCCTGTGACAAAGC</td>
</tr>
</tbody>
</table>

Primer positions are indicated in the parenthesis by the 5' end position numbered in the clone CTC-554D6 (AC008575.7) for APC, in the clone RP11-491D6 (AC006583.31) for hMLH1, in the clone RP11-159K7 (AC137549.1) for 4 and 5, in the clone CTC-487M23 (AC008536.7) for 13 and 14, in the clone RP11-56P2 (AC136500.1) for 18 and 19, and in the clone CTD-2201G3 (AC010431.7) for 28 and 29.

The value from 2 in this group should represent an experimental error (Fig. 2). Two patients (KT25 and KT64) showed gene dose values around 1, suggesting that they had only one copy of the Ex15 sequence. Another patient reproducibly displayed a gene dose of ~3, suggesting the presence of three copies of the Ex15 sequence. Gene dose values of these three patients deviated from 2 by more than 2x standard deviation calculated from all samples (standard deviation =0.384, n=47).
Gene dose mapping in patients with only one copy of the \textit{Ex15} sequence. The sequence copy number at multiple sites around the \textit{APC} gene was determined by gene dose assay and microsatellite polymorphism analysis in patients KT25 and KT64. The gene dose at several target sites around the gene was measured, and then the next set of target sites was selected within the regions where the genomic breakpoints were found to be located by the first gene dose assay. These procedures were repeated until the breakpoints were located within at most 10-kb regions.

In KT25, the copy number was examined at sites 1-15 and was a half of the normal value at sites 5-13 (Fig. 3), suggesting that the breakpoints were within a 6.2-kb region between sites 4 and 5 and within a 6.1-kb region between sites 13 and 14. In KT64, the copy number was examined at sites 16-31 and was a half of the normal value at sites 19-28 (Fig. 3), suggesting that the breakpoints were within a 6.9-kb region between sites 18 and 19 and within a 6.6-kb region between sites 28 and 29. Microsatellite analysis revealed only one major peak at D5S346 and two peaks at D5S1965 in both KT25 and KT64 patients (data not shown), being compatible with the gene dose assay results (Fig. 3).

Determination of deleted sequences in KT25 and KT64 patients. The deletion boundary sequences were amplified by PCR with primers designed to hybridize to the region where the deletion breakpoints were located by the gene dose mapping (Fig. 4). When the DNA of KT25 was used as a template,
Figure 4. Molecular cloning and determination of the deletion boundaries in KT25 (A) and KT64 (B). In the upper diagrams, arrows, the open box and vertical bars indicate PCR primers, the APC gene region and target sites of gene dose assay (numbered as in Fig. 3), respectively. Ethidium bromide-stained agarose gels displaying PCR products are shown in the middle. M, DNA ladder for size marker. Nucleotide sequences of the deletion boundaries are shown below. Nucleotide sequences of the breakpoint-containing PCR products (upper case letters) and registered sequences in the database (lower case letters) are aligned where colons and hyphens indicate identical nucleotides and nucleotide gaps, respectively. A, in KT25, PCR was conducted with every possible combination between upstream and downstream primers indicated. The primers UP1-4 and DP1-4 were designed at an ~1-kb interval within regions between sites 4 and 5 and between 13 and 14. Only the primer pair of UP2 and 13-e (lane 1) and the primer pair UP3 and 13-e (lane 2) generated PCR products with KT25 DNA whereas no such products were generated with healthy control DNA (lanes 3 and 4). a, the deletion boundary sequence of the PCR product. The 4-bp sequence flanking the deletion is underlined. b, the database sequence from position 114956 (left) to 114921 (right) of AC137549.1. c, the database sequence from position 106397 (left) to 106432 (right) of AC008536.7. B, in KT64, long-range PCR was conducted with primers 18-f and 29-e. Lane 1, KT64; lane 2, healthy control DNA. a, the deletion boundary sequence of the PCR product. b, the LINE-1 sequence located between 18 and 19, from position 113107 (left upper) to 112898 (right lower) of AC136500.1. c, the LINE-1 sequence from position 43820 (left upper) to 43603 (right lower) of AC010431.7, which is located 50 kb further downstream from the LINE-1 element present between 28 and 29. d, the LINE-1 sequence located between 28 and 29, from position 99225 (left upper) to 99008 (right lower) of AC010431.7. The double and single horizontal lines indicate intervals in which the recombination was suspected to occur between the LINE-1 sequences b and c and between c and d, respectively.
PCR with the primer pair UP2 and 13-e and that with UP3 and 13-e generated PCR products of 2.3 and 1.3 kb, respectively (Fig. 4A), which were not generated from the control DNA. Sequencing analysis of these PCR products confirmed that the 1.3-kb sequence was a part of the 2.3-kb sequence. Comparison of these sequences with those registered in the database revealed that deletion breakpoints in KT25 were located between an ~224-kb upstream site (nucleotide position 114937 numbered in AC137549.1) and an ~120-kb downstream site (position 106417 in AC008536.7). The deletion was ~435 kb in size and contained the whole APC gene (Fig. 5). There was no highly homologous sequence in the deletion boundaries except a 4-bp identical sequence at the deletion end-points.

In KT64, the regions where the breakpoints were located by gene dose mapping consisted mostly of long interspersed repetitive sequences called LINE-1 (16). Because of the high sequence homology among LINE-1 members, PCR primers for a unique sequence could not be designed within these regions. Therefore, genomic DNA was amplified by long-range PCR with the primers 18-f and 29-e (Fig. 4B). This PCR generated a 7-kb fragment from KT64 DNA but not from the control DNA (Fig. 4B). Analysis of restriction enzyme recognition sites of this fragment and comparison with database sequences revealed that the contiguous 1.2-kb and 1.6-kb fragments generated by KpnI digestion of the 7-kb fragment contained the deletion boundary sequences (data not shown). These KpnI fragments were introduced into a plasmid vector and analyzed by nucleotide sequencing. Comparison of these sequences with the database indicated that the boundary sequences consisted of three LINE-1 elements that were present in the same orientation around the APC gene (Fig. 4B). The upstream sequence of the boundary was approximately half (2.5 kb) of the LINE-1 element located ~80 kb upstream (registered in AC136500.1) from the APC gene. The downstream sequence of the boundary was approximately half (2.7 kb) of the LINE-1 element located ~570 kb downstream (registered in AC010431.7) from the APC gene. Between these two LINE-1 sequences, a short sequence of 46-139 bp, which appears to have been derived from another LINE-1 sequence, was inserted between the two half LINE-1 elements shown in this diagram.

**Discussion**

Genomic deletion has increasingly been recognized as a fairly common mutation of a tumor suppressor gene that should be searched for when routine PCR-based analysis fails to identify a causative mutation. The deletions vary widely in size, ranging...
from a single exon to an extremely large chromosomal region containing many genes. To detect these deletions, various techniques have been employed such as Southern blot analysis (5), fluorescence in situ hybridization (18), long-range PCR (19) and several newly developed methods for gene dosage measurement. The latter methods, which estimate the copy number of specific sequences, include competitive PCR (12,20,21), multiplex PCR (22,23), real-time PCR (10), multiplex amplifiable probe hybridization (24), multiplex ligation-dependent probe amplification (25) and comparative genome hybridization (26). Although various methods are thus available, no single method is adequate for every occasion and combination of these techniques is necessary in some cases. In a DNA diagnosis setting, demonstration of gene dosage alteration may not be sufficient for definite conclusion and determination of genomic breakpoints may be required.

We previously demonstrated that gene dosage assay based on competitive PCR was useful for identification of MEN1 gene deletion and determination of deletion end-points (13). In the present study, we partly modified the procedure of preparing internal standard DNA. This modification made the construction of internal standard easier, and facilitated gene dose assay at newly adopted target sites. This modified assay was applied to screening for APC gene deletion and then to determination of exact deletion breakpoints in FAP patients. Two patients were identified as a deletion carrier by the initial screening of the exon 15 copy number variation and were ultimately shown to have a 435-kb and 737-kb genomic deletion including the whole APC gene. Another patient was suspected to have APC gene duplication because there appeared to be three copies of exon 15. Further analysis is necessary to characterize the APC gene mutation in this patient. We found three patients having APC gene deletion or possible rearrangement among 41 patients. Similar detection rate, 7 patients with large deletion among 60 classical FAP patients, was previously reported in the screening for copy number variation of the APC exon 14 by real-time quantitative multiplex PCR assay (10).

We determined the deleted genomic sequences, which were flanked by only a 4-bp repeat in one patient and LINE-1 repetitive sequences of ~6 kb in another patient. The involvement of short direct repeats of a few nucleotides at the deletion breakpoint has been well documented where DNA polymerase slippage has been suggested as the mechanism of generating genomic deletions (13,27,28). Interspersed repetitive sequences such as LINE-1 and Alu retrotransposon elements have also repeatedly been implicated in gene rearrangements (27-29). Generation of large genomic deletions can be mediated by these elements through retrotransposition (30) or non-allelic homologous recombination (27-29). The latter mechanism is likely to be responsible for the creation of the chimeric LINE-1 element identified at the deletion breakpoint in KT64. The lack of structural hallmarks of retrotransposition such as 5' truncation and a 3' poly(A) tail is also consistent with the homologous recombination mechanism. However, the presence of a third LINE-1 sequence of 46-139 bp in the chimeric element at the breakpoint suggests that the recombination was not a simple crossover event between two LINE-1 elements. Similar three-piece chimeras have been reported for Alu elements associated with large genomic deletions, where ~130-bp Alu repeat sequences were inserted (27,28). Although a model has been proposed for its formation process assuming the contribution of an Okazaki fragment from a remote repetitive element (27,28), the precise mechanism for the creation of a three-piece chimera associated with a large genomic deletion remains to be elucidated.

The delineation of genomic breakpoints in affected families will provide a much stronger proof of gene deletion than reduced gene dosage or reduced fluorescence in situ hybridization signals alone, and consequently offer important information for genetic counseling of the patients’ families. Our gene dosage mapping successfully delimited the deletion boundaries and led to the determination of breakpoints only with a small amount of clinical DNA samples that remained after the initial mutation screening. Thus, the gene dose mapping has proved to be a useful technique for a definitive DNA test of large gene deletions in a clinical setting.

Acknowledgements

This study was supported in part by the program for promotion of Fundamental Studies in Health Sciences of the Pharmaceautical and Medical Devices Agency, by Grant-in-aid for Scientific Research from Japan Society for the Promotion of Science and by SRF for Biomedical Research. M.T. is an Awardee of Research Resident Fellowship from the Foundation for Promotion of Cancer Research (Japan) for the 3rd Term Comprehensive 10-Year-Strategy for Cancer Control.

References


