Cell, Vol. 66, 589–600, August 9, 1991, Copyright © 1991 by Cell Press

Identification and Characterization of the Familial Adenomatous Polyposis Coli Gene

Joanna Groden,* Andrew Thliveris,*† Wade Samowitz,‡ Mary Carlson,* Lawrence Gelbert,*† Hans Albertsen,* Geoff Joslyn,*† Jeff Stevens,*† Lisa Spirio,* Margaret Robertson,*† Leslie Sargeant,*† Karen Krapcho,*† Erika Wolff,*† Randall Burt,§ J. P. Hughes,‖ Janet Warrington,# John McPherson,# John Wasmuth,# Denis Le Paslier,° Hadi Abderrahim,° Daniel Cohen,° Mark Leppert,*† and Ray White*†

*Department of Human Genetics
†Howard Hughes Medical Institute
‡Department of Pathology
§Gastroenterology Division, Department of Medicine
University of Utah Health Sciences Center
Salt Lake City, Utah 84132
‖St. Mark's Hospital
Salt Lake City, Utah 84117
#Department of Biological Chemistry
College of Medicine
University of California
Irvine, California 92717
°CEPH
27 rue Juliette Dodu
75010 Paris
France

Summary

DNA from 61 unrelated patients with adenomatous polyposis coli (APC) was examined for mutations in three genes (*DP1*, *SRP19*, and *DP2.5*) located within a 100 kb region deleted in two of the patients. The intron–exon boundary sequences were defined for each of these genes, and single-strand conformation polymorphism analysis of exons from *DP2.5* identified four mutations specific to APC patients. Each of two aberrant alleles contained a base substitution changing an amino acid to a stop codon in the predicted peptide; the other mutations were small deletions leading to frameshifts. Analysis of DNA from parents of one of these patients showed that his 2 bp deletion is a new mutation; furthermore, the mutation was transmitted to two of his children. These data have established that *DP2.5* is the *APC* gene.

Introduction

Familial adenomatous polyposis coli (APC) is a dominantly inherited autosomal disorder characterized by an early onset of multiple adenomatous polyps of the colon and a high likelihood of colon carcinoma. As the gene responsible for familial APC also may play a key role in the development of sporadic colon cancer, identification of the *APC* gene would provide a valuable tool for examination of early events in the progression of normal colonic epithelium to adenoma and to carcinoma.

Pulsed-field gel analysis of the *APC* region in DNA from 40 patients revealed deletions in two patients (see Joslyn et al., 1991 [this issue]). The smaller deletion, of approximately 100 kb, is nested within the second, 260 kb deletion. Three candidate genes that map into the smaller deletion have been identified (Figure 1). To determine which of them, if any, is responsible for the inherited disease, mutations specific to APC patients must be demonstrated.

Cytogenetically visible deletions at the locus of the *APC* gene, in addition to the two smaller deletions, have been observed in affected individuals (Herrera et al., 1986; Rivera et al., 1990). If the deletions in these patients are causing the disease by removing all or part of a gene, one would expect that single-base substitutions that inactivate the gene also might be able to cause APC. Therefore, we undertook a search for inactivating point mutations in the three genes within the doubly deleted region, using single-strand conformation polymorphism (SSCP) analysis (Orita et al., 1989a, 1989b). SSCP has proven useful in the analysis of other disease-causing mutations, such as those in neurofibromatosis type 1 and cystic fibrosis, where DNA sequence variation has been characterized successfully at the level of single-base substitutions (Cawthon et al., 1990; Dean et al., 1990). Examination of 12 of the 17 exons of *MCC* by this method has so far failed to reveal any variants unique to APC patients (J. G., unpublished data).

Identification of intron–exon boundaries provided intronic sequence flanking each of the exons of the three new candidate genes, and specific oligonucleotides primed polymerase chain reaction (PCR) amplification of genomic exon sequences from APC patients and controls. SSCP conformers unique to APC patients were identified; sequencing revealed several distinct stop codons and frameshift mutations in the *DP2.5* gene, one of which was a new mutation.

Results

Intron–Exon Boundaries

Figure 1 provides a map of the region deleted in patients 3214 and 3824, indicating the relative locations of the restriction sites and genes ascertained from within the deletions. To obtain DNA sequence adjacent to the exons of the genes *DP1*, *DP2.5*, and *SRP19*, sequencing substrate was obtained by inverse PCR amplification of DNAs from two YACs, 310D8 and 183H12, that span the deletions (Joslyn et al., 1991). Ligation at low concentration cyclized the restriction enzyme–digested YAC DNAs. Oligonucleotides with sequencing tails, designed in inverse orientation at intervals along the cDNAs, primed PCR amplification from the cyclized templates. Comparison of these DNA sequences with the cDNA sequences (see Joslyn et al., 1991) placed exon boundaries at the divergence points. *SRP19* and *DP1* were each shown to have five exons

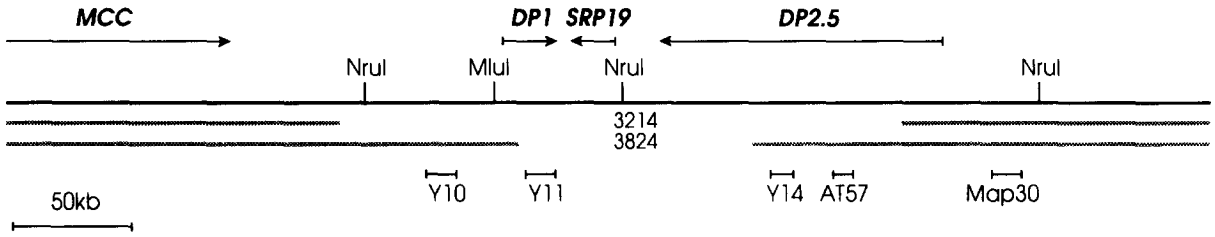


Figure 1. Restriction Map of the Central 500 kb of the 1200 kb NotI Fragment Spanning the *APC* Region

Genomic DNA is shown in black with restriction sites marked vertically. The two deletion chromosomes identified by Joslyn et al. (1991) are shown as stippled lines; clones mapping in the deleted regions are shown below. Positions and orientations of three genes in the doubly deleted region, as well as the 3′ end of *MCC*, are shown above the map.

(G. J. and A. T., unpublished data [*DP1* and *SRP19*, respectively]). *DP2.5* consisted of 15 exons.

The sequences of the oligonucleotides synthesized to provide PCR amplification primers for the exons of each of these genes are listed in Table 1. With the exception of exons 1, 3, 4, 9, and 15 of *DP2.5* (see below), the primer sequences were located in intron sequences flanking the exons. The 5′ primer of exon 1 is complementary to the cDNA sequence, but extends just into the 5′ Kozak context sequence for the initiator methionine, allowing a survey of the translated sequences. The 5′ primer of exon 3 is actually in the 5′ coding sequences of this exon, as three separate intronic primers simply would not amplify. The 5′ primer of exon 4 just overlaps the 5′ end of this exon, and we thus fail to survey the 19 most 5′ bases of this exon. For exon 9, two overlapping primer sets were used, such that each had one end within the exon. For exon 15, the large 3′ exon of *DP2.5*, overlapping primer pairs were placed along the length of the exon; each pair amplified a product of 250–400 bases.

Identification of Nucleotide Alterations in APC Patients

SSCP analysis identifies most single- or multiple-base changes in DNA fragments up to 400 bases in length (Orita et al., 1989a, 1989b). Sequence alterations are detected as shifts in electrophoretic mobility of single-stranded DNA on nondenaturing acrylamide gels; the two complementary strands of a DNA segment usually resolve as two SSCP conformers of distinct mobilities. However, if the sample is from an individual heterozygous for a base-pair variant within the amplified segment, often three or more bands are seen. In some cases, even the sample from a homozygous individual will show multiple bands. Base-pair-change variants are identified by differences in pattern among the DNAs of the sample set.

Exons of the candidate genes were amplified by PCR from the DNAs of 61 unrelated APC patients and a control set of 12 normal individuals. The five exons from *DP1* revealed no unique conformers in the APC patients, although common conformers were observed with exons 2 and 3 in some individuals of both affected and control sets, indicating the presence of DNA sequence polymorphisms. Likewise, none of the five exons of *SRP19* revealed unique conformers in DNA from APC patients in the test panel.

Testing of exons 1 through 14 and primer sets A through N of exon 15 of the *DP2.5* gene, however, revealed variant conformers specific to APC patients in exons 7, 8, 10, 11, and 15. These variants were in the unrelated patients 3746, 3460, 3827, 3712, and 3751, respectively (Figures 2A–2E). The PCR–SSCP procedure was repeated for each of these exons in the five affected individuals and in an expanded set of 48 normal controls. The variant bands were reproducible in the APC patients but were not observed in any of the control DNA samples. Additional variant conformers in exons 11 and 15 of the *DP2.5* gene were seen; however, each of these was found in both the affected and control DNA sets.

Sequencing of Unique Conformers from APC Patients

The five sets of conformers unique to the APC patients were sequenced to determine the nucleotide changes responsible for their altered mobilities. The normal conformers from the host individuals were sequenced also. Bands were cut from the dried acrylamide gels, and the DNA was eluted. PCR amplification of these DNAs provided template for sequencing.

The sequences of the unique conformers from exons 7, 8, 10, and 11 of *DP2.5* revealed dramatic mutations in the *DP2.5* gene. The sequence of the new mutation creating the exon 7 conformer in patient 3746 was shown to contain a deletion of two adjacent nucleotides, at positions 730 and 731 in the cDNA sequence reported by Joslyn et al. (1991). The normal sequence at this splice junction is CAGAGGTCA (intronic sequence underlined), with the intron–exon boundary between the two repetitions of AG. The mutant allele in this patient has the sequence CAGGTCA (Figure 3A). Although this change is at the 5′ splice site, comparison with known consensus sequences of splice junctions would suggest that a functional splice junction is maintained. If this new splice junction were functional, the mutation would introduce a frameshift that creates a stop codon 15 nucleotides downstream. If the new splice junction were not functional, messenger processing would be significantly altered.

To confirm the 2 base deletion, the PCR product from APC patient 3746 and a control DNA were electrophoresed on an acrylamide–urea denaturing gel, along with the products of a sequencing reaction. The sample from

Table 1. Sequences of Primers Used for SSCP Analyses

Exon	Primer 1	Primer 2
	DP1	
1	UP-TCCCCGCCTGCCGCTCTC	RP-GCAGCGGCGGCTCCCGTG
2	UP-GTGAACGGCTCTCATGCTGC	RP-ACGTGCGGGGAGGAATGGA
3	UP-ATGATATCTTACCAAATCATATAC	RP-TTATTCCTACTTCTTCTATACAG
4	UP-TACCCATGCTGGCTCTTTTTC	RP-TGGGGCCATCTTGTTCCTGA
5	UP-ACATTAGGCACAAAGCTTGCAA	RP-ATCAAGCTCCAGTAAGAAGGTA
	SRP19	
1	UP-TGCGGCTCCTGGGTTGTTG	RP-GCCCCTTCCTTTCTGAGGAC
2	UP-TTTTCTCCTGCCTCTTACTGC	RP-ATGACACCCCCCATTCCCTC
3	UP-CCACTTAAAGCACATATATTTAGT	RP-GTATGGAAAATAGTGAAGAACC
4	UP-TTCTTAAGTCCTGTTTTTCTTTTG	RP-TTTAGAACCTTTTTTGTGTTGTG
5	UP-CTCAGATTATACACTAAGCCTAAC	RP-CATGTCTCTTACAGTAGTACCA
	DP2.5	
1	UP-AGGTCCAAGGGTAGCCAAGG*	RP-TAAAAATGGATAAACTACAATTAAAAG
2	UP-AAATACAGAATCATGTCTTGAAGT	RP-ACACCTAAAGATGACAATTTGAG
3	UP-TAACTTAGATAGCAGTAATTTCCC*	RP-ACAATAAACTGGAGTACACAAGG
4	UP-ATAGGTCATTGCTTCTTGCTGAT*	RP-TGAATTTTAATGGATTACCTAGGT
5	UP-CTTTTTTTGCTTTTACTGATTAACG	RP-TGTAATTCATTTTATTCCTAATAGCTC
6	UP-GGTAGCCATAGTATGATTATTTCT	RP-CTACCTATTTTTATACCCACAAAC
7	UP-AAGAAAGCCTACACCATTTTTGC	RP-GATCATTCTTAGAACCATCTTGC
8	UP-ACCTATAGTCTAAATTATACCATC	RP-GTCATGGCATTAGTGACCAG
9	UP-AGTCGTAATTTTGTTTCTAAACTC	RP-TGAAGGACTCGGATTTCACGC*
9a	UP-TCATTCACTCACAGCCTGATGAC*	RP-GCTTTGAAACATGCACTACGAT
10	UP-AAACATCATTGCTCTTCAAATAAC	RP-TACCATGATTTAAAAATCCACCAG
11	UP-GATGATTGTCTTTTTCCTCTTGC	RP-CTGAGCTATCTTAAGAAATACATG
12	UP-TTTTAAATGATCCTCTATTCTGTAT	RP-ACAGAGTCAGACCCTGCCTCAAAG
13	UP-TTTCTATTCTTACTGCTAGCATT	RP-ATACACAGGTAAGAAATTAGGA
14	UP-TAGATGACCCATATTCTGTTTC	RP-CAATTAGGTCTTTTTGAGAGTA
15-A	UP-GTTACTGCATACACATTGTGAC	RP-GCTTTTTGTTTCCTAACATGAAG*
-B	UP-AGTACAAGGATGCCAATATTATG*	RP-ACTTCTATCTTTTTCAGAACGAG*
-C	UP-ATTTGAATACTACAGTGTTACCC*	RP-CTTGTATTCTAATTTGGCATAAGG*
-D	UP-CTGCCCATACACATTCAAACAC*	RP-TGTTTGGGTCTTGCCCATCTT*
-E	UP-AGTCTTAAATATTCAGATGAGCAG*	RP-GTTTCTCTTCATTATATTTTATGCTA*
-F	UP-AAGCCTACCAATTATAGTGAACG*	RP-AGCTGATGACAAAGATGATAATG*
-G	UP-AAGAAACAATACAGACTTATTGTG*	RP-ATGAGTGGGGTCTCCTGAAC*
-H	UP-ATCTCCCTCCAAAAGTGGTGC*	RP-TCCATCTGGAGTACTTTCTGTG*
-I	UP-AGTAAATGCTGCAGTTCAGAGG*	RP-CCGTGGCATATCATCCCCC*
-J	UP-CCCAGACTGCTTCAAAATTACC*	RP-GAGCCTCATCTGTACTTCTGC*
-K	UP-CCCTCCAAATGAGTTAGCTGC*	RP-TTGTGGTATAGGTTTTACTGGTG*
-L	UP-ACCCAACAAAAATCAGTTAGATG*	RP-GTGGCTGGTAACTTTAGCCTC*
-M	UP-ATGATGTTGACCTTTCCAGGG*	RP-ATTGTGTAACTTTTCATCAGTTGC*
-N	UP-AAAGACATACCAGACAGAGGG*	RP-CTTTTTTGGCATTGCGGAGCT*
-O	UP-AAGATGACCTGTTGCAGGAATG*	RP-GAATCAGACGAAGCTTGTCTAGAT*
-P	UP-CCATAGTAAGTAGTTTACATCAAG*	RP-AAACAGGACTTGTACTGTAGGA*
-Q	UP-CAGCCCCTTCAAGCAAACATG*	RP-GAGGACTTATTCCATTTCTACC*
-R	UP-CAGTCTCCTGGCCGAAACTC*	RP-GTTGACTGGCGTACTAATACAG*
-S	UP-TGGTAATGGAGCCAATAAAAAGG*	RP-TGGGAGTTTTCGCCATCCAC*
-T	UP-TGTCTCTATCCACACATTCGTC*	RP-ATGTTTTTCATCCTCACTTTTTGC*
-U	UP-GGAGAAGAACTGGAAGTTCATC*	RP-TTGAATCTTTAATGTTTGGATTTGC*
-V	UP-TCTCCCACAGGTAATACTCCC	RP-GCTAGAACTGAATGGGGTACG
-W	UP-CAGGACAAAATAATCCTGTCCC	RP-ATTTTCTTAGTTTCATTCTTCCTC

All primers are read in the 5′ to 3′ direction. The first primer in each pair lies 5′ of the exon it amplifies; the second primer lies 3′ of the exon it amplifies. Primers that lie within the exon are identified by an asterisk. UP represents the −21M13 universal primer sequence; RP represents the M13 reverse primer sequence.

patient 3746 showed two bands differing in size by 2 nucleotides, with the larger band identical in mobility to the control sample; this result was independent confirmation that patient 3746 is heterozygous for a 2 bp deletion (data not shown).

The unique conformer found in exon 8 of patient 3460 was found to carry a C→T transition, at position 904 in the published cDNA sequence of *DP2.5*, which replaced the normal sequence of CGA with TGA (Figure 3B). This point mutation, when read in frame, results in a stop codon replacing the normal arginine codon. This single-base change had occurred within the context of a CG dimer, a potential hot spot for mutation (Barker et al., 1984).

The conformer unique to APC patient 3827 in exon 10 was found to contain a deletion of one nucleotide (1367, 1368, or 1369) when compared to the normal sequence found in the other bands on the SSCP gel. This deletion, occurring within a set of three T's, changed the sequence from CTTTCA to CTTCA (Figure 3C); this 1 base frameshift creates a downstream stop within 30 bases. The PCR product amplified from this patient's DNA also was electrophoresed on an acrylamide–urea denaturing gel, along with the PCR product from a control DNA and products from a sequencing reaction. The patient's PCR product showed two bands differing by 1 bp in length, with the larger identical in mobility to the PCR product from the normal DNA; this result confirmed the presence of a 1 bp deletion in patient 3827 (data not shown).

Sequence analysis of the variant conformer of exon 11 from patient 3712 revealed the substitution of a T by a G at position 1500, changing the normal tyrosine codon to a stop codon (Figure 3D).

The pair of conformers observed in exon 15 of the *DP2.5* gene from APC patient 3751 also was sequenced. These conformers were found to carry a nucleotide substitution of C to G at position 5253, the third base of a valine codon (Figure 3E). No amino acid change resulted from this substitution, suggesting that this conformer reflects a genetically silent polymorphism.

The observation of distinct inactivating mutations in the *DP2.5* gene in four unrelated patients strongly suggested that *DP2.5* is the *APC* gene.

Transmission of the New Mutation

Confirmation that the *DP2.5* is, in fact, the *APC* gene was provided by family studies. Patient 3746, described above as carrying an APC allele with a frameshift mutation, is an affected offspring of two normal parents. The pedigree of this family is shown in Figure 4. Colonoscopy revealed no polyps in either parent nor among the patient's three siblings.

DNA samples from both parents, from the patient's wife, and from their three children were examined. As shown in Figure 4, SSCP analysis of DNA from both of the patient's parents displayed the normal pattern of conformers for exon 7, as did DNA from the patient's wife and one of his offspring. The two other children, however, displayed the same new conformers as their affected father. Testing of the patient and his parents with highly polymorphic VNTR

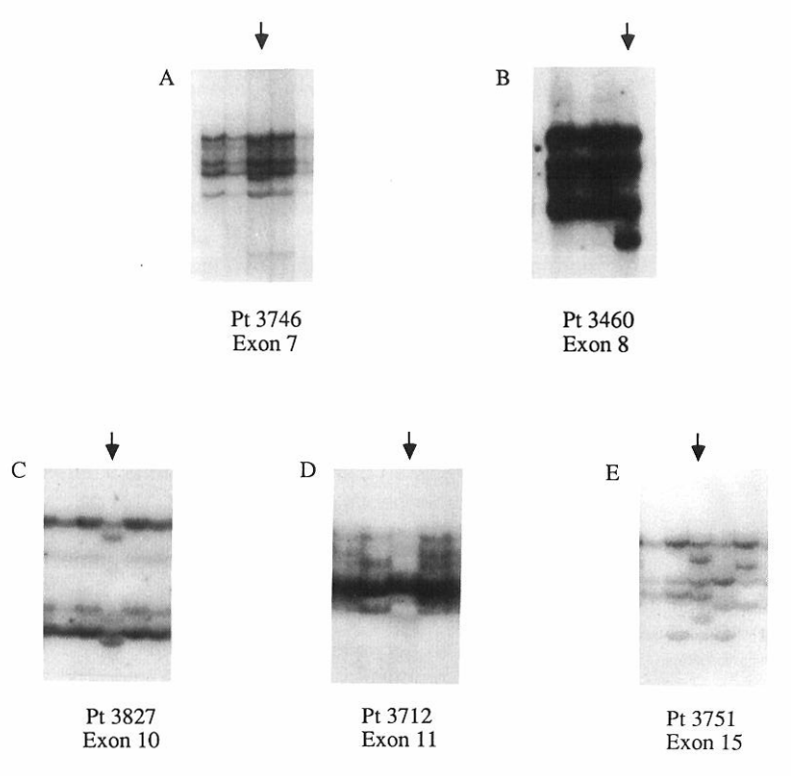


Figure 2. SSCP Conformers Specific to APC Patients in the *DP2.5* Gene

Each panel shows one of the unique SSCP conformers found in five unrelated APC patients. Arrows mark the lanes containing the new conformers. The flanking lanes in each panel show the expected pattern of conformers in DNA from other affected individuals and from normal controls. For exons 11 and 15, the conformer patterns of these control DNAs also show the common polymorphic conformers found in these exons. The PCR products analyzed are from exon 7 (A), exon 8 (B), exon 10 (C), exon 11 (D), and exon 15 (E). The patients with the unique conformers are identified by the numbers 3746, 3460, 3827, 3712, and 3751, respectively.

markers showed a 99.98% likelihood that they are his biological parents.

These observations confirmed that this novel conformer, known to reflect a 2 bp deletion mutation in the *DP2.5* gene, appeared spontaneously with APC in this pedigree and was transmitted to two of the children of the affected individual.

APC Polymorphisms

Sequencing of variant conformers found among controls as well as individuals with APC has revealed the following polymorphisms in the *APC* gene: first, in exon 11, at position 1458, a substitution of T to C creating an RsaI restriction site but no amino acid change; and second, in exon 15, at positions 5037 and 5271, substitutions of A to G and G to T, respectively, neither resulting in amino acid substitutions. These nucleotide polymorphisms in the *APC* gene sequence may be useful for diagnostic purposes.

Genomic Structure of the *APC* Gene

The positions of the exons of the *APC* gene, and the sequences of the intron–exon boundaries, are shown in Figure 5. The continuity of the very large (6.5 kb), most 3′ exon in *DP2.5* was shown in two ways. First, inverse PCR with primers spanning the entire length of this exon revealed no divergence of the cDNA sequence from the genomic sequence. Second, PCR amplification with converging primers placed at intervals along the exon generated products of the same size whether amplified from the originally isolated cDNA, cDNA from various tissues, or genomic template. Two forms of exon 9 were found in *DP2.5*: one is the complete exon; and the other, labeled exon 9A, is the result of a splice into the interior of the exon that deletes bases 934 to 1236 in the mRNA and removes 101 amino acids from the predicted protein (see Figure 9 in Joslyn et al., 1991).

Mapping of *APC* Exons in Somatic Cell Hybrids from Deletion Patients

Somatic cell hybrids carrying the segregated chromosomes 5 from the 100 kb (HHW1291) and 260 kb (HHW1155) deletion patients (Joslyn et al., 1991) were used to determine the distribution of the *APC* gene exons across the deletions. DNAs from these cell lines were used as template, along with genomic DNA from a normal control, for PCR-based amplification of the *APC* exons.

PCR analysis of the hybrids from the 260 kb deletion of patient 3214 showed that all but one (exon 1) of the *APC* exons are removed by this deletion (Figure 6A). PCR analysis of the somatic cell hybrid HHW1291, carrying the chromosome 5 homolog with the 100 kb deletion from patient 3824, revealed that exons 1 through 9 are present but exons 10 through 15 are missing (Figure 6B). This result placed the deletion breakpoint either between exons 9 and 10 or within exon 10.

Expression Properties

Tissues that express the *APC* gene were identified by PCR amplification of cDNA made to mRNA with primers located within adjacent *APC* exons. In addition, PCR primers that flank the alternatively spliced exon 9 were chosen so that the expression pattern of both splice forms could be assessed. All tissue types tested (brain, lung, aorta, spleen, heart, kidney, liver, stomach, placenta, and colonic mu-

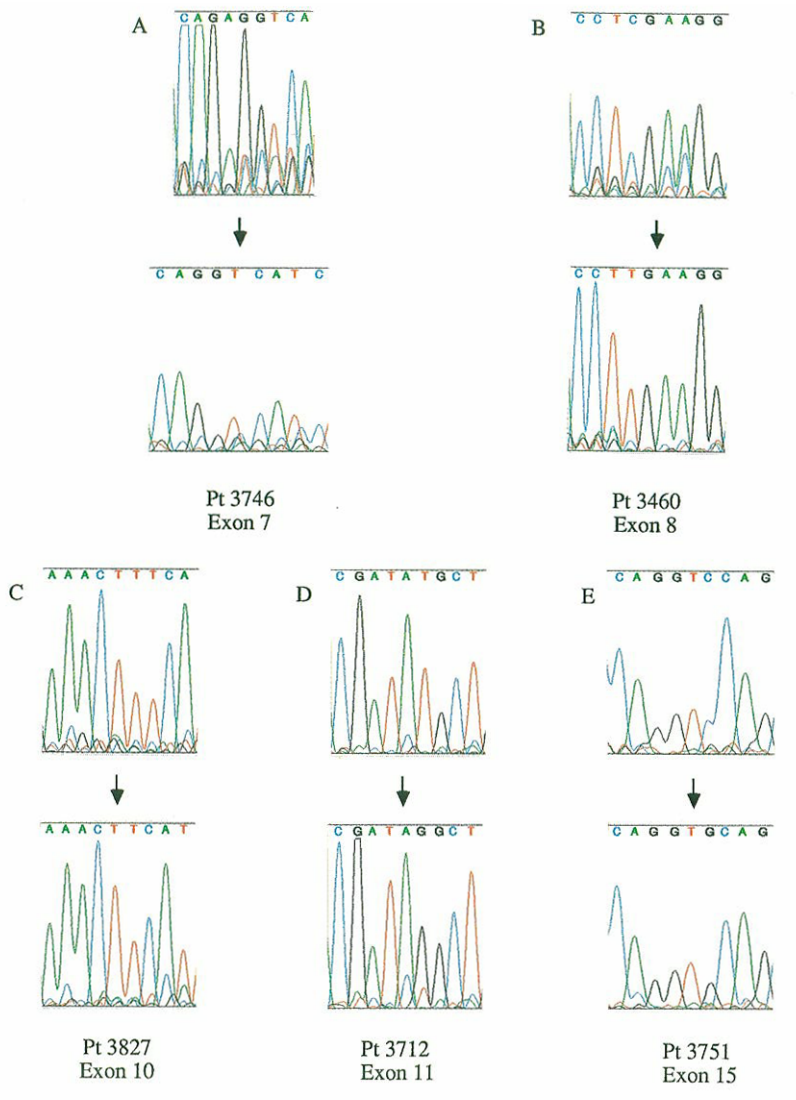


Figure 3. Sequences of the Unique SSCP Conformers Found in the *DP2.5* Genes of Five APC Patients

(A) Sequences of the SSCP conformers from patient 3746. The normal sequence at the 5′ intron–exon boundary of exon 7 (3 intronic nucleotides followed by cDNA nucleotides 730–735) is shown at the top; the lower panel shows the deletion of an AG at positions 730 and 731 in the unique SSCP conformer from this patient. This change introduces a frameshift in the coding sequence of the gene.

(B) Sequences of the SSCP conformers from patient 3460. The normal sequence at positions 901–909 is shown at the top; the lower panel shows a C→T transition at position 904 in the unique SSCP conformer from this patient. This change introduces a stop codon in the coding sequence.

(C) Sequences of the SSCP conformers from patient 3827. The normal sequence at positions 1363–1371 is shown at the top; the lower panel shows the deletion of one of the three T's at positions 1367–1369 in the unique SSCP conformer from this patient. This change introduces a frameshift in the coding sequence.

(D) Sequences of the SSCP conformers from patient 3712. The normal sequence at positions 1495–1503 is shown at the top; the lower panel shows a T→G transversion at position 1500 in the unique SSCP conformer from this patient. This change introduces a stop codon in the coding sequence.

(E) Sequences of the SSCP conformers from patient 3751. The normal sequence at positions 5248–5256 is shown at the top; the lower panel shows a C→G transversion at position 5253 in the unique SSCP conformer from this patient. This change does not alter the coding sequence.

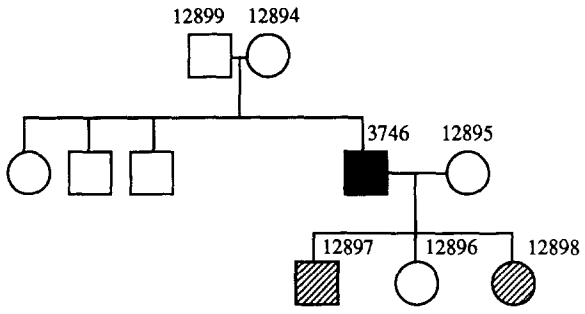
cosa) and cultured cell lines (lymphoblasts, HL60, and choriocarcinoma) expressed both splice forms of the *APC* gene (Figure 7). We note, however, that expression by lymphocytes normally residing in some tissues, including colon, prevents unequivocal assessment of expression. The larger mRNA, containing the complete exon 9 rather than only exon 9A, appears to be the more abundant message.

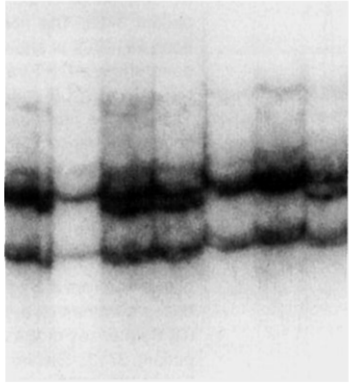
Northern analysis of poly(A)-selected RNA from lymphoblasts revealed a single band of approximately 10 kb, consistent with the size of the sequenced cDNA (data not shown).

Structure of the Predicted Protein Product of the *APC* Gene

The cDNA consensus sequence of *APC* predicts that the longer, more abundant form of the message codes for a 2844 amino acid peptide with a mass of 311.8 kd (see Figure 9 in Joslyn et al., 1991). This predicted *APC* peptide was compared with the current data bases of protein and DNA sequences using both Intelligenetics and GCG software packages. No genes with a high degree of amino acid sequence similarity were found. Although many short (approximately 20 amino acid) regions of sequence similarity were uncovered, none was sufficiently strong to reveal which, if any, might represent functional homology. Interestingly, multiple similarities to myosins and keratins did appear. The *APC* gene also was scanned for sequence motifs of known function; although multiple glycosylation, phosphorylation, and myristoylation sites were seen, their significance is uncertain.

Analysis of the *APC* peptide sequence did identify features important in considering potential protein structure. Hydropathy plots (Kyte and Doolittle, 1982) indicate that the *APC* protein is notably hydrophilic. No hydrophobic domains suggesting a signal peptide or a membrane-spanning domain were found. Analysis of the first 1000 residues indicates that α-helical rods may form (Cohen and Parry, 1986); there is a scarcity of proline residues and, as shown in Figure 8, there are a number of regions





containing heptad repeats (apolar-X-X-apolar-X-X-X). Interestingly, in exon 9A, the deleted form of exon 9, two heptad repeat regions are reconnected in the proper heptad repeat frame, deleting the intervening peptide region. After the first 1000 residues, the high proline content of the remainder of the peptide suggests a compact rather than a rod-like structure.

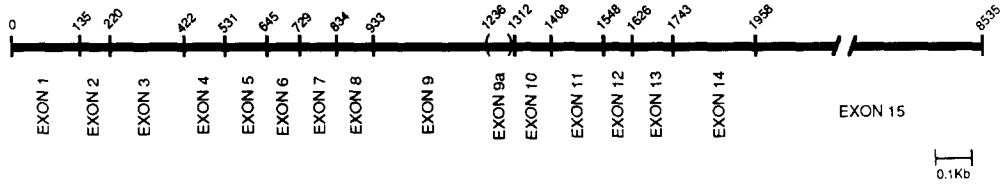
The most prominent feature of the second 1000 residues is a 20 amino acid repeat that is iterated seven times with semiregular spacing (Table 2). The intervening sequences between the seven repeat regions contained 114, 116,

Figure 4. Inheritance of a New Mutation in the *APC* Gene

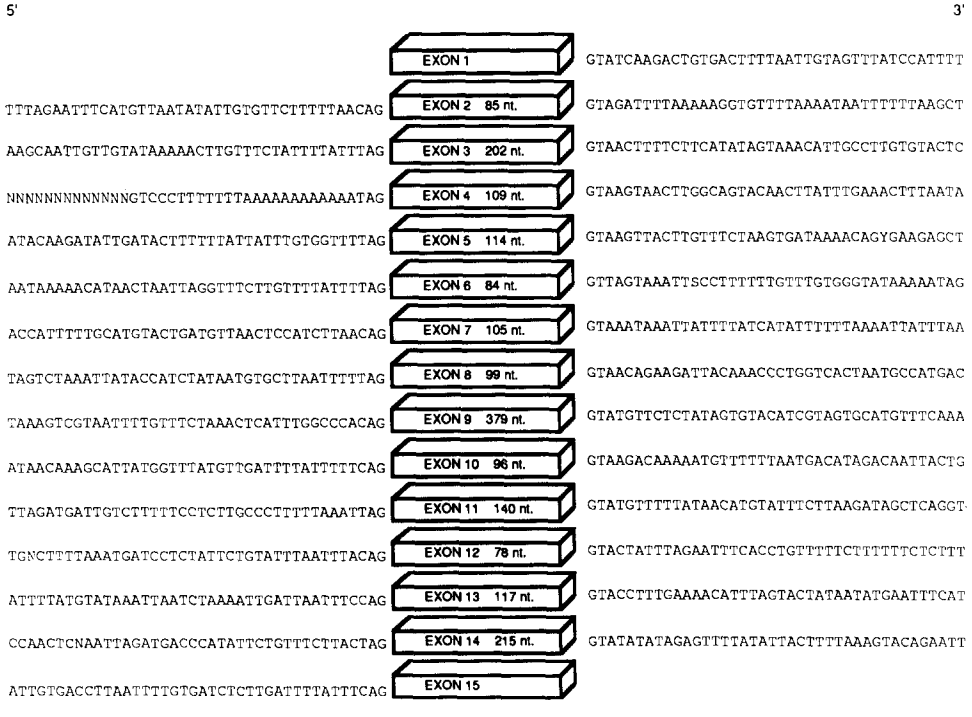
(Top) The pedigree of patient 3746 is shown, with numbers denoting the individuals from whom DNA was obtained. This kindred was ascertained originally by patient 3746, who also is identified by the blackened square in this pedigree.

(Bottom) SSCP analysis of DNA from individuals in the kindred described above, amplified with primers flanking exon 1 of *DP2.5*. Each lane is centered below the individual from whom the corresponding DNA was obtained. The unique conformer originally observed in DNA from patient 3746 also can be seen in the DNA of two of his offspring (identified by hatching in the pedigree) but not in DNA from his two normal parents.

A)



B)



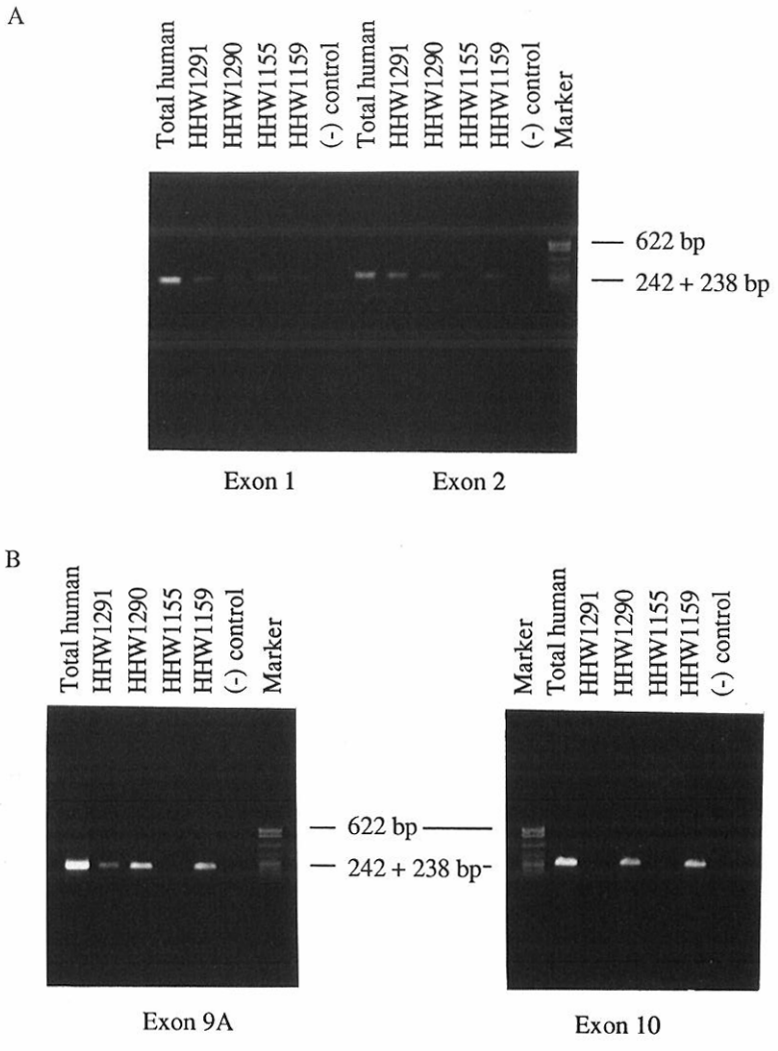


Figure 6. PCR Analysis of the *APC* Gene in the Deleted Chromosomes of Patients 3214 and 3824

(A) Primers flanking *APC* exons 1 and 2 were used to amplify DNA from the four somatic cell hybrids derived from deletion patients 3214 and 3824. HHW1155 contains the deleted chromosome 5 from 3214; HHW1291 contains the deleted chromosome 5 from 3824; HHW1159 and HHW1290 carry the respective normal homologs. Total human DNA from a normal individual also is included as a control. Exon 1 is amplified from all four hybrids; exon 2 is not amplified from hybrid HHW1155, indicating a breakpoint between these two exons or within exon 2 in this chromosome.

(B) Primers flanking *APC* exons 9A and 10 were used to amplify DNA from the same four somatic cell hybrids. Exon 9A is not amplified from HHW1155 but is amplified from HHW1291. Exon 10 is not amplified from either deletion-carrying hybrid. Both exons are amplified from hybrids with the normal homologs. This result indicates that the breakpoint in the HHW1291 hybrid is either between exons 9A and 10 or within exon 10.

151, 205, 107, and 58 amino acids, respectively. Finally, residues 2200–2400 contain a 200 amino acid basic domain.

Discussion

Four inactivating mutations found in APC patients specify *DP2.5* as the *APC* gene. Two of the mutations create stop codons; the other two are frameshift mutations (1 bp and 2 bp deletions) that alter the reading frame and create stop codons immediately downstream. None of these mutations is seen among a panel of unaffected, control samples. Most importantly, however, the 2 bp deletion of patient 3746 is a new mutation; it is found in neither of the parents, both of whom are clinically normal. We conclude that mutations in *DP2.5* are capable of causing APC and that *DP2.5* is the *APC* gene.

Relatively little suggestion as to the function of the *APC* gene product has been revealed by searching data bases for amino acid sequence similarities with known proteins. The strongest similarities detected are to intermediate-filament structural proteins such as the myosins and keratins. These similarities are due largely to the presence in the predicted peptide of a series of heptad repeat motifs (Cohen and Parry, 1986) that also are found, interestingly, in the *MCC* gene product (Bourne, 1991). The heptad repeats are believed to be capable of facilitating protein–protein interactions through formation of extended α helices that present a series of hydrophobic residues extending along one side. These hydrophobic regions can stabilize

Figure 5. Exons in *DP2.5*

(A) Positions of exon boundaries. Exons 9 corresponds to nucleotides 933–1312; exon 9A corresponds to nucleotides 1236–1312. The stop codon in the cDNA is at nucleotide 8535.

(B) Partial intronic sequence surrounding each exon.

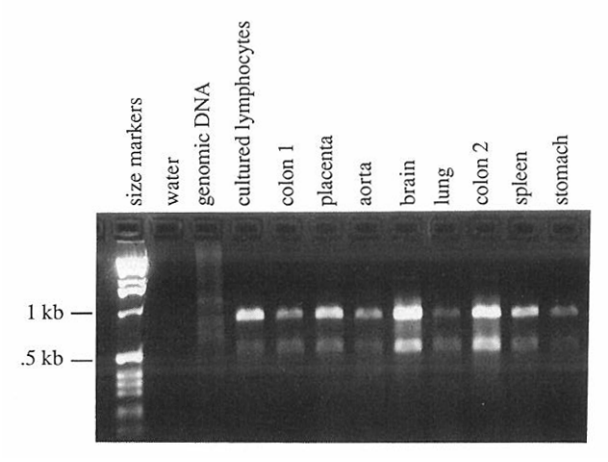


Figure 7. Expression of *APC* Product in Various Human Tissues, from Exon-Connection Experiments

PCR with primers flanking the alternatively spliced exon shows expression of two products in all tissues examined and a size difference consistent with that predicted by cDNA sequence analysis (303 bp; see Joslyn et al., 1991). The larger form appears to be more abundant in most tissues.

the interaction of two such α helices to form either homo- or heterodimers through the formation of a coiled coil (Cohen and Parry, 1986). However, the short length of the heptad repeat regions in the predicted *APC* gene product does not suggest an extended filamentous structure and may reflect intrapeptide bundling of α helices.

On the other hand, in the case of the *APC* gene products, one of which contains a fusion of two heptad repeat regions, such interactions might result in the formation of multimers containing both kinds of subunits. Moreover, the presence of heptad repeats in the putative product of the closely located *MCC* gene, which shows mutations in tumors, might permit an interaction between the products of the *APC* and *MCC* genes.

In addition to the heptad repeats, multiple potential serine phosphorylation site motifs, glycosylation motifs, and myristoylation sites are found. The functional significance of these potential sites remains to be established. No hydrophobic regions suggesting a transmembrane domain have been found, and the overall hydrophilic nature of the predicted protein suggests that the *APC* protein may be located in the interior of the cell.

Table 2. Seven Different Versions of the 20 Amino Acid Repeat

Consensus:	F * V E * T P * C F S R * S S L S S L S
1262:	Y C V E D T P I C F S R C S S L S S L S
1376:	H Y V Q E T P L M F S R C T S V S S L D
1492:	F A T E S T P D G F S C S S S L S A L S
1643:	Y C V E G T P I N F S T A T S L S D L T
1848:	T P I E G T P Y C F S R N D S L S S L D
1953:	F A I E N T P V C P S H N S S L S S L S
2013:	F H V E D T P V C F S R N S S L S S L S

Numbers denote the first amino acid of each repeat. The consensus sequence at the top reflects a majority amino acid at a given position.

All six of the mutations characterized here in APC patients—two stop codons, two frameshift deletions, and two larger deletions—would be expected to inactivate the *APC* gene. These inactivating mutations in the *APC* gene support the hypothesis that the *APC* gene is a tumor suppressor, analogous in its formal genetics to the retinoblastoma gene (Knudson et al., 1975). Loss of the chromosome carrying the normal homolog is diagnostic for the presence of a tumor suppressor gene (Cavenee et al., 1985). The significant frequency of loss of chromosome 5 homologs observed in both sporadic and APC colon carcinomas (Solomon et al., 1987; Vogelstein et al., 1988; Okamoto et al., 1988) suggests that, at the cellular level, the mutant alleles of the *APC* gene are recessive with respect to carcinoma formation.

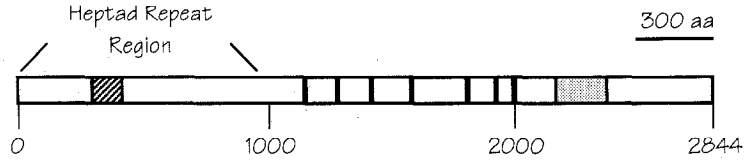
A distinction between retinoblastoma and colon carcinoma, however, is that in the case of colon carcinoma a well-defined intermediate in the progression pathway, the adenomatous polyp, can be examined. Although the tumor suppressor model might suggest that the adenomatous polyp would likewise show loss of a chromosome 5 homolog, almost none of the polyps of individuals with APC show a chromosome 5 loss (Solomon et al., 1987; Vogelstein et al., 1988). This finding suggests that the inherited *APC* alleles are dominant with respect to polyp formation, even at the level of the individual cellular precursor. It should be noted that a second event, perhaps a mutation in some other, as yet unidentified, gene, is required for polyp formation, as all of the colonic epithelial cells of an APC patient are heterozygous for the inherited APC mutation, but only a few become precursors to polyps.

Bourne (1991) has recently proposed that dominance with respect to polyp formation of inherited *APC* mutations could be due to incorporation of altered gene products into a multisubunit structure. However, the finding reported here that four mutant alleles of the *APC* gene are, in fact, inactivating stop codons and frameshifts seems inconsistent with that model.

Bodmer et al. (1987) have suggested an alternative view of polyp formation, accounting for the dominance of *APC* alleles with respect to polyp formation and proposing a stochastic basis for the second event. In this view *APC* gene mutations reduce the activity of the gene, and polyp formation occurs when the level of activity of the *APC* gene falls below some critical, or threshold, value. This hypothesis would be consistent with the gene-inactivating mutations described here, as loss of one allele might be expected to reduce the activity of the *APC* gene in the cell. If this quantitative view were correct, it might also suggest an explanation for regression of adenomas in APC patients following treatment with the nonsteroidal anti-inflammatory drug sulindac (Waddell et al., 1989). Such treatment might alter the cellular environment in such a way that the threshold requirement for activity of the polyposis gene product is reduced, and lower levels of *APC* gene activity might then be sufficient to prevent development or maintenance of the adenoma.

Although missense amino acid substitutions resulting in reduced activity would be expected to arise at a frequency comparable to that of gene-inactivating mutations, such

A



B

```
6
Tyr Asp Gln Leu Leu Lys Gln
Val Glu Ala Leu Lys Met Glu
Asn Ser Asn Leu Arg Gln Glu
Leu Glu Asp Asn Ser Asn His
Leu Thr Lys Leu Glu Thr Glu
Ala Ser Asn Met Lys Glu Val
Leu Lys Gln Leu Gln Gly Ser
Ile Glu Asp
        68

           123
           Phe Val Asn Gly
Ser Arg Glu Ser Thr Gly Tyr
Leu Glu Glu Leu Glu Lys Glu
Arg Ser Leu Leu Leu Ala Asp
Leu Asp Lys
       150

221
Ile Gln Gln Ile Glu Lys Asp
Ile Leu Arg Ile Arg Gln Leu
Leu Gln Ser Gln Ala Thr Glu
Ala Glu Arg Ser
            245
```

Repeats specific to the longer gene product.

```
305
Thr Ser His Leu Gly Thr Lys
Val Glu Met Val Tyr Ser Leu
Leu Ser Met Leu Gly Thr His
                        325
406
Val Leu His Leu Leu Glu Gln
Ile Arg Ala Tyr Cys Glu Thr
Cys Trp Glu Trp Gln Glu Ala
                        426
```

Repeat specific to the shorter gene product.

```
Thr Ser His Leu Gly Thr Lys
Ile Arg Ala Tyr Cys Glu Thr
Cys Trp Glu Trp Gln Glu Ala
```

```
469
Leu Gly Gly Leu Gln Ala Ile
Ala Glu Leu Leu Gln Val Asp
Cys Glu Met Tyr Gly Leu Thr
Asn Asp His Tyr Ser Ile Thr
Leu Arg Arg Tyr Ala Gly Met
Ala Leu Thr Asn Leu Thr Phe
Gly Asp Val Ala Asn Lys Ala
                        517

545
Ala Ser Val Leu Arg Asn Leu
Ser Trp Arg Ala Asp Val Asn
Ser Lys Lys Thr Leu Arg Glu
Val Gly Ser Val Lys Ala Leu
Met Glu Cys Ala Leu Glu Val
                        579

585
Leu Lys Ser Val Leu Ser Ala
Leu Trp Asn Leu Ser Ala His
Cys Thr Glu Asn Lys Ala Asp
Ile Cys Ala Val Asp Gly Ala
Leu Ala Phe Leu Val Gly Thr
                        619

631
Ile Ile Glu Ser Gly Gly Gly
Ile Leu Arg Asn Val Ser Ser
Leu Ile Ala Thr Asn Glu Asp
His Arg Gln Ile Leu Arg Glu
Asn Asn Cys Leu Gln Thr Leu
Leu Gln His Leu Lys Ser His
Ser Leu Thr Ile Val Ser Asn
Ala Cys Gly Thr Leu Trp Asn
                        686

880
Ile Ser Thr Thr Ala Ala Gln
Ile Ala Lys Val Met Glu Glu
Val Ser Ala Ile
            897
```

Figure 8. Structure of the Predicted *APC* Gene Product and Amino Acid Sequences of Heptad Repeat Regions

(A) Schematic of the *APC* protein. The first 1000 amino acids contain heptad repeat motifs. The hatched region is coded by the alternative exon (exon 9) and is absent from the smaller mRNA. Vertical bars are the seven repeats of 20 amino acids (Table 2); stippling denotes a basic domain.

(B) Heptad repeats (apolar-X-X-apolar-X-X-X) present in the first 1000 amino acids of the *APC* protein. The numbers, based on the longer gene product, denote the first and last residue of each repeat block; the apolar residues that define the repeats appear in boldface. The three repeat units that are boxed are specific to the different gene products. Underlined amino acids in the repeats beginning at residues 305 and 406 are derived from the alternative exon (exon 9) and are present only in the longer gene product. Amino acids that are not underlined in the above repeats become contiguous in product derived from the alternative splice of exon 9 and form a heptad repeat block that is specific to the shorter gene product.

mutations have not yet been seen. One reason might be that, in general, amino acid substitutions are not sufficiently damaging to give the characteristic APC phenotype. Such mutations may be present, therefore, but create a milder phenotype, currently unrecognized as APC.

Families with a high incidence of colon cancer but relatively small numbers of polyps (10–20 per patient on average) have been identified recently, and the gene causing the phenotype has been mapped to the *APC* locus (Leppert et al., 1990; L. S., unpublished data). Indeed, allelic heterogeneity at the *APC* locus was proposed to account for the phenotypic heterogeneity of the APC pedigrees. Evidence for genetic transmission of small numbers of polyps in some families, and the proposition that this transmission is the basis for much of the population risk of colon carcinoma (Cannon-Albright et al., 1988; Burt et al., 1985), suggests that even milder *APC* alleles might exist, predisposing a carrier to the development of only one or a few polyps (Leppert et al., 1990).

The role of *MCC*, a gene that is mutated in some colonic tumors (Kinzler et al., 1991), in the inherited disorder APC is unresolved. The *MCC* mutations in carcinomas suggest that the function of this gene is relevant to tumorigenesis. However, the finding of inherited alterations in a different gene in some APC patients, one of them a new mutation that entered the pedigree concurrently with the disease, is evidence that, in these families, an inherited defect in *MCC* is not the cause of their polyposis phenotype.

Another possibility—suggested by the potential involvement of both the *MCC* and *APC* genes in colonic tumorigenesis, their physical proximity, and the presence in both of heptad repeat motifs—is that the *MCC* and *APC* gene products interact, perhaps through heterodimer formation, to form an active structure. If this were true, mutations

in either gene during tumorigenesis might affect the activity of the heterodimer. The failure to find inherited mutations in *MCC* might then be explained either by the lower mutation frequency expected from a smaller gene, or by lethality of constitutional mutations in *MCC*.

A more immediate use of the cloned *APC* gene will be in the development of DNA-based diagnostic tests for the presence of mutant *APC* alleles. The most effective tests directly detect the presence of a mutant allele in a patient; for now, this level of test will be available only for those relatively few examples where the mutation has been characterized. For the remaining individuals known to be at risk because of family history, linkage tests involving other family members are required to determine whether they have inherited a mutant *APC* allele. The several *APC* gene polymorphisms reported here will be helpful in providing diagnostic information. Furthermore, a multi-allelic CA repeat polymorphism present in the very closely colocated *DP1* gene (L. S., unpublished data) should also provide accurate diagnostic information.

Experimental Procedures

Identification of APC Patients

APC patients were identified through referrals from private physicians and hospital clinics associated with the diagnosis and treatment of APC. Patients were selected on the basis of their availability and their willingness to donate blood samples. Participating institutions included the following: the Department of Surgery and the Division of Gastroenterology, Department of Medicine, at the University of Utah Medical Center; the Johns Hopkins School of Medicine, Baltimore, MD; St. Mark's Hospital, Salt Lake City, UT; LDS Hospital, Salt Lake City, UT; Cleveland Clinic, Cleveland, OH; and the University of Illinois College of Medicine at Peoria Medical Center, Peoria, IL. Medical records were obtained and reviewed to verify the correct diagnosis of APC in each patient. All human tissue samples used in this project were obtained with the approval of the Institutional Review Board at the University of Utah Health Sciences Center.

Tissue Culture

Epstein–Barr virus was used to transform lymphocytes from normal and APC individuals. These lymphoblastoid cell lines then were cultured at 37°C, 5% CO_2 in RPMI 1640 medium (Cellgro/Mediatech) with 1% Nutridoma (Boehringer Mannheim), 5% fetal bovine serum (Hyclone), and 50 U/ml gentamycin sulfate. Hybrid cell lines were derived as described in Joslyn et al. (1991) and were cultured at 39°C, 5% CO_2 in DMEM (Cellgro/Mediatech) with 10% fetal bovine serum (Hyclone) and 2% penicillin–streptomycin (GIBCO). Patient DNA was obtained directly from peripheral blood or from lymphoblastoid cell lines initiated in our laboratory from patients' blood.

Identification of Intron–Exon Boundaries

Yeast cells carrying a YAC were grown in 100 ml of YDP medium at 32°C, and DNA was harvested according to published protocols (Ausubel et al., 1989). Two micrograms of total yeast plus YAC DNA (310D8) was digested with several enzymes (Sau3A, HaeIII, TaqI, AluI, HhaI, and MspI). Following digestion, DNAs were phenol extracted, precipitated with 2.5 vols of ethanol, and collected by centrifugation. DNAs were washed with 70% ethanol, dried under vacuum, and resuspended in 50 μl of TE-4 (10 mM Tris [pH 7.5], 0.1 mM EDTA). DNA molecules were ligated to maximize circularization (Ochman et al., 1990), with final DNA concentrations between 0.2 and 2.0 μg/ml. Ligated DNAs were precipitated, washed, and resuspended in 40 μl of TE-4. Five microliters of each ligation was used for PCR experiments.

To obtain genomic sequence flanking putative exons of *DP2.5*, diverging primers designed within cDNA sequence (Table 3) were used with the above ligation mixtures in PCR reactions (30 cycles; each cycle 1 min at 94°C, 1 min at 52°C, 2 min at 72°C). Amplified products were resolved on 2% agarose gels (Maniatis et al., 1982) and isolated from the gel with Geneclean (Bio101, La Jolla, CA). These products were sequenced using the dideoxy termination method (Sanger et al., 1977) with fluorescently tagged M13 primers on an Applied Biosystems model 373A DNA sequencer (Applied Biosystems, Foster City, CA).

Table 3. Primer Sequences for Inverse PCR

Exon	Primer 1	Primer 2
1	RP-GTAAGATGATTGGAATTATCTTCTA	UP-GGAAACTGAGGCATCTAATATGAAG
2	RP-GTCCAGAAGAAGCCATAG	UP-TGATTTATTAGAGCGTCTTA
3	RP-AGAGAAAGTACTGGATATTTAGAA	UP-CGGCTTCCATAAGAACGGAG
	RP-CCAGGGAAATTACTGCTATCTA	UP-TCAAAAATGTCCCTCCGTTCT
4	RP-TCAAGATCAGCAAGAAGCAATG	UP-AAGGAAAAAGACTGGTATTACG
5	RP-CTGGCAAATCTGTTTGTAAGGA	UP-ATATGAGCAAGGCAAATCAGAG
6	RP-TATGTCCTTTTCGATTTGCTGA	UP-GTATACGACAGCTTTTACAGTC
7	RP-TGACCTTCATTCTGCCGCTC	UP-ATCAACATGGCAACTTCTGGTA
8	RP-CAGAGTGTGTGCTACTAGAAC	UP-GGCTGACAAGTCATCTGGGA
9	RP-AGCATTGACAACAATGAATACAC	UP-TCATGATAAGGATGATATGTCGC
9A	RP-CTTCCTGCCACTCCCAACAG	UP-CATGGACCAGGACAAAAATCCA
10	RP-CATTAGAACACACACAGCAGG	UP-GAAGAGCATAGACATGCAATGAA
11	RP-AAGTTTGTCAAAGCCATTCCAG	UP-TTGGAGATGTAGCCAACAAGG
12	RP-CAAGTGCTCTCATGCAGCCT	UP-CCAACTAAAATCTGAAAGTGAAG
13	RP-CGCAACGTCTTTTTACTATTTACA	UP-GAAGTGTGAAAGCATTGATGGAA
14	RP-TCAATACGCTTTTGAGGGTTGAT	UP-ATTGCACTGAGAATAAAGCTGAT
15	RP-AAAGTTTGTAGACAGTTGTTCTC	UP-ACACTTAAAATCTCATAGTTTGAC
	RP-TCATGTCGATTGGTGTCAAAAAC	UP-CAGACAATTTTAATACTGGCAAC
	RP-CAGAATAGGATTCAATCGAGG	UP-ATGAAAGTAAGTTTTGCAGTTATG
	RP-ACTATGAAGATGATAAGCCTACC	UP-AATTCCATGATTAGAACCCACTC
	RP-CAGCTGACCTAGTTCCAATC	UP-CTGTGAGCGAAGTTCCAGCA

UP = TGTAAAACGACGGCCAGT; RP = CAGGAAACAGCTATGACC.

PCR Amplification

DNA samples were amplified for SSCP analysis using the polymerase chain reaction (5 min at 95°C, 1 time; 1 min at 95°C, 1 min at annealing temperature, 1 min at 72°C, 35 times; 10 min at 72°C, 1 time) with the primer pairs listed in Table 1. The annealing temperature for each reaction was 2°C–4°C below the T_m for the primer. The reaction mixture was made up of the following: 200 ng of DNA, 0.5 μM each primer, 70 μM each deoxynucleoside triphosphate, 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 0.25 U of Taq polymerase, 0.25 mM spermidine, and 0.1 μl of [α-^{32}P]dCTP (3000 Ci/ml) in a volume of 10 μl.

Somatic cell hybrid DNAs were amplified similarly, but in a volume of 25 μl, without isotope, and with 1.25 U of Taq polymerase and 2.5 mM dNTPs.

SSCP Gel Analysis

PCR products were diluted in 50 μl of 0.1% SDS, 10 mM EDTA and 50 μl of 95% formamide, 20 mM EDTA, 0.05% bromphenol blue, 0.05% xylene cyanol. Products were denatured at 95°C for 3 min and kept on ice until loaded onto 5% polyacrylamide gels. Two conditions were run for each set of reactions: a 90 mM Tris–borate (pH 7.5), 2 mM EDTA gel at 4°C, and a 90 mM Tris–borate (pH 7.5), 2 mM EDTA, 10% glycerol gel at room temperature. Electrophoresis was carried out for both conditions at 40 W, constant power.

After electrophoresis, gels were transferred to Whatman 3MM paper and dried on vacuum slab dryers. Autoradiography with Kodak X-Omat AR film and without intensifying screens, overnight, was usually sufficient to detect bands on film.

Sequencing of SSCP Conformers

Individual SSCP bands were cut directly from the dried gels, placed in 100 μl of distilled water, shaken for 30 min at room temperature, and

centrifuged briefly. A 10 μl aliquot of this supernatant was used for PCR amplification with the same primer pair that was used in the original amplification. Reaction mixtures were as follows: 0.5 μM each primer, 10 μl of 10× Perkin-Elmer Cetus PCR buffer, and 2.0 μl of a 1.25 mM dNTP mix. Samples were vortexed, spun down, placed on ice, and overlaid with mineral oil. Two units of Taq polymerase was added to each reaction. Reactions were placed in a preheated DNA thermal cycler at 95°C and run with the following parameters: 95°C, 1 min; 60°C, 1 min; 72°C, 1 min (25 times). Samples were purified by a centrifugation wash with a Centricon 100 column (Amicon). Samples were stored frozen until ready for sequencing on the Applied Biosystems model 373A DNA sequencer (Applied BioSystems, Foster City, CA). Samples were sequenced with the dideoxy termination method with Taq polymerase and fluorescently tagged M13 universal or reverse sequencing primers.

RNA Preparation and Northern Analysis

RNA was extracted from tissue by the guanidium–acid–phenol method described by Chomczynski and Sacchi (1987). Approximately 3 μg of poly(A)-selected RNA was electrophoresed through 1% agarose gels containing 0.66 M formaldehyde and blotted overnight onto uncharged nylon membranes. Probe preparation and hybridization were performed by the methods described above.

Exon Connection

RNA was extracted from tissues as described by Chomczynski and Sacchi (1987). cDNA was made through reverse transcription of 1 μg of total RNA, with synthesis primed by random hexamers (Perkin-Elmer Cetus Corp., Norwalk, CT). PCR was then performed on the cDNA (Fearon et al., 1990) using one oligonucleotide primer from exon 8 and a second primer from exon 13. PCR with these same primers on genomic DNA served as a control for contamination of RNA samples with DNA.

Acknowledgments

The laboratory experiments reported here reflect a communal effort on the part of authors H. Albertsen, M. C., J. G., L. G., G. J., M. R., W. S., J. S., L. S., and A. T. The authors are grateful to Steve Gherken for help in scanning the data bases, to Ed Meenen for timely preparation of oligonucleotides, and to Ruth Foltz for editing and compiling the manuscript. We are also indebted to Bob Weiss for his helpful comments and discussions. In addition to the sources acknowledged in the companion paper, we were provided patient material by Drs. David Jagelman and Philip Paul at the Cleveland Clinic and by the DNA Diagnostic Laboratory at the University of Utah. We also wish to thank Dr. Allen Sinclair for his examination of the relatives of patient 3746. H. Albertsen was supported by a grant from the Danish Medical Research Council, J. G. by National Cancer Institute grant 5T32 CA09602, M. C. and L. G. by National Institutes of Health grant R01-GM40886-03, and J. W. by NIH grant HG-00320. R. W. is an Investigator of the Howard Hughes Medical Institute.

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 USC Section 1734 solely to indicate this fact.

Received June 11, 1991; revised June 26, 1991.

References

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K. (1989). Current Protocols in Molecular Biology, Vol. 1 (New York: John Wiley and Sons), pp. 13.03–13.12.3.

Barker, D., Schafer, M., and White, R. (1984). Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell *36*, 131–138.

Bodmer, W., Bailey, C., Bodmer, J., Bussey, H., Ellis, A., Gorman, P., Lucibello, F., Murday, V., Rider, S., Scambler, P., Sheer, D., Solomon, E., and Spurr, N. (1987). Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature *328*, 614–616.

Bourne, H. (1991). Consider the coiled coil. Nature *351*, 188–190.

Burt, R., Bishop, D. T., Cannon, L., Dowdle, M., Lee, R., and Skolnick, M. (1985). Dominant inheritance of adenomatous colonic polyps and colorectal cancer. N. Engl. J. Med. *312*, 1540–1544.

Cannon-Albright, L., Skolnick, M., Bishop, D. T., Lee, R., and Burt, R. (1988). Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. N. Engl. J. Med. *319*, 533–537.

Cavenee, W., Hansen, M., Nordenskjold, M., Kock, E., Maumenee, I., Squire, J., Phillips, R., and Gallie, B. (1985). Genetic origin of mutations predisposing to retinoblastoma. Science *228*, 501–503.

Cawthon, R. M., Weiss, R., Xu, G., Viskochil, D., Culver, M., Stevens, J., Robertson, M., Dunn, D., Gesteland, R., O'Connell, P., and White, R. (1990). A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations. Cell *62*, 193–201.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyante–phenol–chloroform extraction. Anal. Biochem. *162*, 156–159.

Cohen, C., and Parry, D. (1986). α-Helical coiled coils—a widespread motif in proteins. Trends Biochem. Sci. *11*, 245–248.

Dean, M., White, M. B., Amos, J., Gerrard, B., Stewart, C., Khaw, K.-T., and Leppert, M. (1990). Multiple mutations in highly conserved residues are found in mildly affected cystic fibrosis patients. Cell *61*, 863–870.

Fearon, E., Cho, K., Nigro, J., Kern, S., Simons, J., Ruppert, J., Hamilton, S., Preisinger, A., Thomas, G., Kinsler, K., and Vogelstein, B. (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. Science *247*, 49–56.

Herrera, L., Kakati, S., Gibas, L, Pietrzak, E., and Sandberg, A. (1986). Brief clinical report: Gardner syndrome in a man with an interstitial deletion of 5q. Am. J. Med. Genet. *25*, 473–476.

Joslyn, G., Carlson, M., Thliveris, A., Albertsen, H., Gelbert, L., 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., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. (1991). Identification of deletion mutations and three new genes at the familial polyposis locus. Cell *66*, this issue.

Kinzler, K., Nilbert, M., Vogelstein, B., Bryan, T., Levy, D., Smith, K., Preisinger, A., Hamilton, S., Hedge, P., Markham, A., Carlson, M., Joslyn, G., Groden, J., White, R., Miki, Y., Miyoshi, Y., Nishisho, I., and Nakamura, Y. (1991). Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science *251*, 1366–1370.

Knudson, A., Hethcote, H., and Brown, B. (1975). Mutation and childhood cancer: a probabilistic model for the incidence of retinoblastoma. Proc. Natl. Acad. Sci. USA *72*, 5116–5120.

Kyte, J., and Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. *157*, 105–132.

Leppert, M., Burt, R., Hughes, J., Samowitz, W., Nakamura, Y., Woodward, S., Gardner, E., Lalouel, J.-M., and White, R. (1990). Genetic analysis of an inherited predisposition to colon cancer in a family with a variable number of adenomatous polyps. N. Engl. J. Med. *322*, 904–908.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Ochman, H., Medhora, M., Garza, D., and Hartl, D. (1990). Amplification of flanking sequences by inverse PCR. In PCR Protocols: A Guide to Methods and Applications, M. Innis, D. Gelfand, J. Sninsky, and T. White, eds. (San Diego: Academic Press), pp. 219–227.

Okamoto, M., Sasaki, M., Sugio, K., Sato, C., Iwama, T., Ikeuchi, T., Tonomura, A., Sasazuki, T., and Miyaki, M. (1988). Loss of constitutional heterozygosity in colon carcinoma from patients with familial polyposis coli. Nature *331*, 273–277.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989a). Detection of polymorphisms of human DNA by gel electrophoresis as single strand conformation polymorphisms. Proc. Natl. Acad. Sci. USA *86*, 2766–2770.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989b). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics *5*, 874–879.

Rivera, H., Simi, P., Rossi, S., Pardelli, L., and Di Paolo, M. (1990). A constitutional 5q23 deletion. J. Med. Genet. *27*, 267–268.

Sanger, F., Nicklen, S., and Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Solomon, E., Voss, R., Hall, V., Bodmer, W., Jass, J., Jeffreys, A., Lucibello, F., Patel, I., and Rider, S. (1987). Chromosome 5 allele loss in human colorectal carcinomas. Nature *328*, 616–619.

Vogelstein, B., Fearon, E., Hamilton, S., Kern, S., Preisinger, A., Leppert, M., Nakamura, Y., White, R., Smits, A., and Bos, J. (1988). Genetic alterations during colorectal tumor development. N. Engl. J. Med. *319*, 525–532.

Waddell, W., Ganser, G., Cerise, E., and Loughry, R. (1989). Sulindac for polyposis of the colon. Am. J. Surg. *157*, 175–179.