Discovery of genetic defects in unexplained colorectal cancer syndromes

Jansen, A.M.L.

Chapter 6

Distinct patterns of somatic mosaicism in the APC gene in neoplasms from patients with unexplained adenomatous polyposis

AML Jansen[1,2], S Crobach[1], WRR Geurts-Giele[3], BEWM van den Akker[1], M Ventayol Garcia[1], D Ruano[1], M Nielsen[4], CMJ Tops[4], JT Wijnen[2,4], FJ Hes[4], T van Wezel[1], WNM Dinjens[3], H Morreau[1]

Gastroenterology, 2017; 152:546-549

Department of [1]Pathology, [2]Human Genetics, [4]Clinical Genetics,
Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Pathology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Abstract

To further unravel mechanisms of *APC* mosaicism in 27 genetically unexplained patients with colonic neoplasms, we used deep sequencing to analyze at least two adenomas or carcinomas per patient. Identical mosaic *APC* variants were identified in adenomas of 9/18 patients with 21-~100 adenomas. Mosaic variants were then variably detected in leukocyte DNA and normal mucosa or were confined to normal mucosa. In one patient comprehensive analysis found no evidence for the mosaic *APC* variant in normal mucosa. One patient was found to carry the mosaic *APC* variant in 10/16 adenomas, underlying the importance of screening ≥ two adenomas.

Keywords: unexplained polyposis; APC mosaicism

Main body

Somatic *APC* mosaicism is estimated to occur in 20% of *de novo* adenomatous polyposis cases.[1-3] Current methods to detect *APC* mosaicism are designed to test for germline variants in leukocyte DNA and the effectiveness of these methods depends on both method-specific sensitivity and the relative contribution of the mosaic variant to the tissue analyzed. Recently, a next-generation sequencing (NGS) study of at least two colonic adenomas in twenty patients found somatic pathogenic *APC* variants in five patients (25%).[4] Analysis of leukocyte DNA confirmed the mosaic *APC* variant in 4/5 cases, but no other tissues were analyzed.

With the aim of identifying *APC* mosaicism we analyzed at least two adenomas or carcinomas with NGS in six patients with 5-20 adenomas, 18 patients with 21-~100 adenomas, one patient with >1000 adenomas, two patients with multiple primary colorectal carcinomas and two positive controls.[5] Nine patients with 21-~100 adenomas (50%) and the two positive controls (APC-08 and APC-09), showed somatic mosaicism, with identical *APC* variants in adenomas tested (Table 1). No difference was observed in location (left/right) or distribution (segmental/whole colon) of the adenomas in patients with or without somatic *APC* mosaicism.

Leukocyte DNA was available for 10/11 of the mosaic patients, including one of the controls. Five of these previously tested negative for a germline mosaic *APC* variant (APC-03, APC-08, APC-09, APC-17 and APC-18, Supplementary Table 1). With NGS deep sequencing a pathogenic variant was identified in three of 10 leukocyte DNAs (APC-03, APC-17 and APC-18) at a 1-4% variant frequency, indicating the detection limit of previously used techniques. Initial testing of adenomas allows detection of very low frequency germline variants and mosaic variants confined to the colon, while initial leukocyte deep sequencing is shown to result in a high percentage false positives.[4] Human gastrulation, the process by which the three germ layers are established, is thought to occur at approximately day 16.[3, 6] The presence of a variant in both leukocyte DNA (mesoderm) and colon mucosa (endoderm) would imply the appearance of this variant before day 16. Primordial germ cells are thought to arise from the primary ectoderm during the second week of development. Therefore, the presence of a somatic variant in both leukocyte DNA and colon mucosa indicates that the variant arose early enough to potentially also be present in germ cells and therefore transmissible to the next generation.

In leukocyte DNA of six mosaic *APC* patients (APC-07, APC-09, APC-15, APC-21, APC-24 and APC-25), the *APC* variant could not be detected by deep sequencing. In four patients (APC-07, APC-09, APC-21 and APC-24), with available normal colonic mucosa, the mosaic *APC* variant was either not detected in normal mucosa (APC-07) or was present at a frequency of 2-29% (APC-09, APC-21 and APC-24, Table 1). In patient APC-24 the variant was present in one of the normal mucosa samples with a 2% frequency, while absent from the other sample, probably due to sampling bias. Variants only present in adenoma- and normal colonic DNA, but absent from leukocyte DNA , indicate a somatic event after the third week of embryogenesis and are unlikely present in the germ cells. For the other two mosaic patients without positivity in leukocyte DNA (APC-15 and APC-25), normal mucosa was not available. DNA from buccal mucosa, fibroblasts and urine from patient APC-09 and APC-21 tested negative for the mosaic *APC* variant.

Patient APC-04 and APC-08 were both found to carry an *APC* c.4666dupA variant in the adenomas initially sequenced with NGS. This variant is located in a homopolymeric nucleotide sequence and could not be reliably detected with the NGS technique alone.[7] However, the variant could be confirmed using the less sensitive HRMA and Sanger sequencing, with a 30-70% variant frequency in both adenomas of patient APC-08 and 10/16 tested adenomas of patient APC-04. In the six adenomas of patient APC-04 not carrying the c.4666dupA variant, other *APC* variants were detected with NGS (Supplementary Table 2, APC-04/P4-P9). The pathology report did not give the exact location of each adenoma and investigating possible clustering of adenomas with and without the *APC* c.4666dupA variant was therefore not possible. Nevertheless, this unusual pattern of adenomas with and without the *APC* mosaic variant illustrates the possible co-occurrence of sporadic adenomas within a mosaic environment, and highlights the importance of screening multiple adenomas. Additionally, leukocyte DNA of the children of patient APC-04, APC-07 and APC-21 was tested for the mosaic *APC* variant detected in the parent (Table 1), but none showed the mosaic variants. These findings are consistent with the *APC* variants occurring after primordial germ cell specification, as would be predicted by the absence of the *APC* variants in leukocyte DNA in APC-07 and APC-21.

In three patients (APC-03, APC-07 and APC-17) the mosaic pattern was subjected to a comprehensive analysis (Figure 1). Deep sequencing of normal mucosa samples identified an *APC* variant at frequencies between 0.4-29% (APC-03) and 0.8-15% (APC-17). In patient APC-07, the *APC* c.3340C>T variant was not found in DNA from the muscularis propria of the ileum, one draining lymph node of the colon, rectal mucosa, small bowel mucosa or seven normal colon mucosa samples. The lymph node sample was analyzed to a 700,000 sequence depth, but still the variant was not detected in this sample. A possible explanation for the anatomically widespread occurrence of the *APC* variant could be field cancerization, a process in which the normal cell population is replaced with tumorigenic clones.[8] This mechanism has been previously described in patients with inflammatory bowel disease, in whom "pre-tumor" clones with identical *TP53* founder mutations, but different driver mutations spread throughout the colon. In agreement with this idea, patient APC-07 showed other additional driver mutations in the different adenomas tested (Supplementary Table 2). A possible mechanism of this clonal expansion of tumorigenic cells is crypt fission, and chronic inflammation of the intestine as seen in inflammatory bowel disease patients is a likely growth stimulus.[8] It is not known whether chronic inflammation was present in patient APC-07. It is also conceivable that previous endoscopies may have aided the spread of tumorigenic cells throughout the intestine, although this does not explain the presence of 20-30 adenomas at first colonoscopy.[8] The size gradient from distal to proximal further accords with the notion that the underlying mechanism of polyposis in this patient was field cancerization developing from one distal tumorigenic cell in the colon and then spreading more proximally.

In detecting APC mosaic variants we underline the importance of testing at least two but preferentially more adenomas. We detected remarkable patterns of mosaicism, with either sporadic adenomas in a mosaic environment or spread of tumorigenic clones throughout the colon suggesting a mechanism of field cancerization.

Table 1: Patients with one APC variant in multiple adenomas

Patient ID	Age	#ad	#T	Mutation	Normal mucosa			Leukocyte		Other tissues tested	Offspring
					n	Freq.	Cov.	Freq.	Cov.		
APC-03	54	~100	6/6	c.3904delC	5	0.4-29%	>2000	1%	2564	LN[2]	0/0
APC-04	54	38	10/16	c.4666dupA	11	NP	NP	NP	NP		0/2
APC-07	48	25	23/23	c.3340C>T	7	0%	NA	0%	NA	I, LN	0/1
APC-08*	35	~100	2/2	c.4666dupA	0	-	-	-	-		0/0
APC-09*	25	62	2/2	c.4057G>T	1	14%	5589	0%	4442	BM, F, U	0/0
APC-15	33	30	3/3	c.2566_2572delCGCGGAA	0	-	-	0%	11446		0/0
APC-17	42	~100	8/8	c.4110_4111delAA	12	0.8 – 15%	>2000	4%	1478		0/0
APC-18	47	80	2/2	c.2493dupA	2	NP	NP	3%[1]	162		0/0
APC-21	67	40	6/6	c.1959-1G>A	2	15-29%	>10000	0%	97256	BM, U	0/1
APC-24	51	58	3/3	c.1974_1975delGA	2	0-2%	>10000	0%	100312		0/0
APC-25	67	30	3/3	c.3211C>T	0	-	-	0%	120062		0/0

*Cumulative adenomas at age of last visit. #ad, number of adenomas; #T, samples with variant/ samples tested; n, number of normal mucosa samples; Freq, frequency; cov, coverage; "-", DNA not available; NP, sequencing not possible due to variant; NA, not applicable - tested with mutation-specific PCR; LN, lymph nodes; I, ileum muscularis propria; F, fibroblasts; BM, buccal mucosa; U, urine. Offspring, children positive for the variant/#children tested *positive control [1]percentage estimated from KASPar assay (Supplementary Figure 1) [2]variant present in 2-3%*

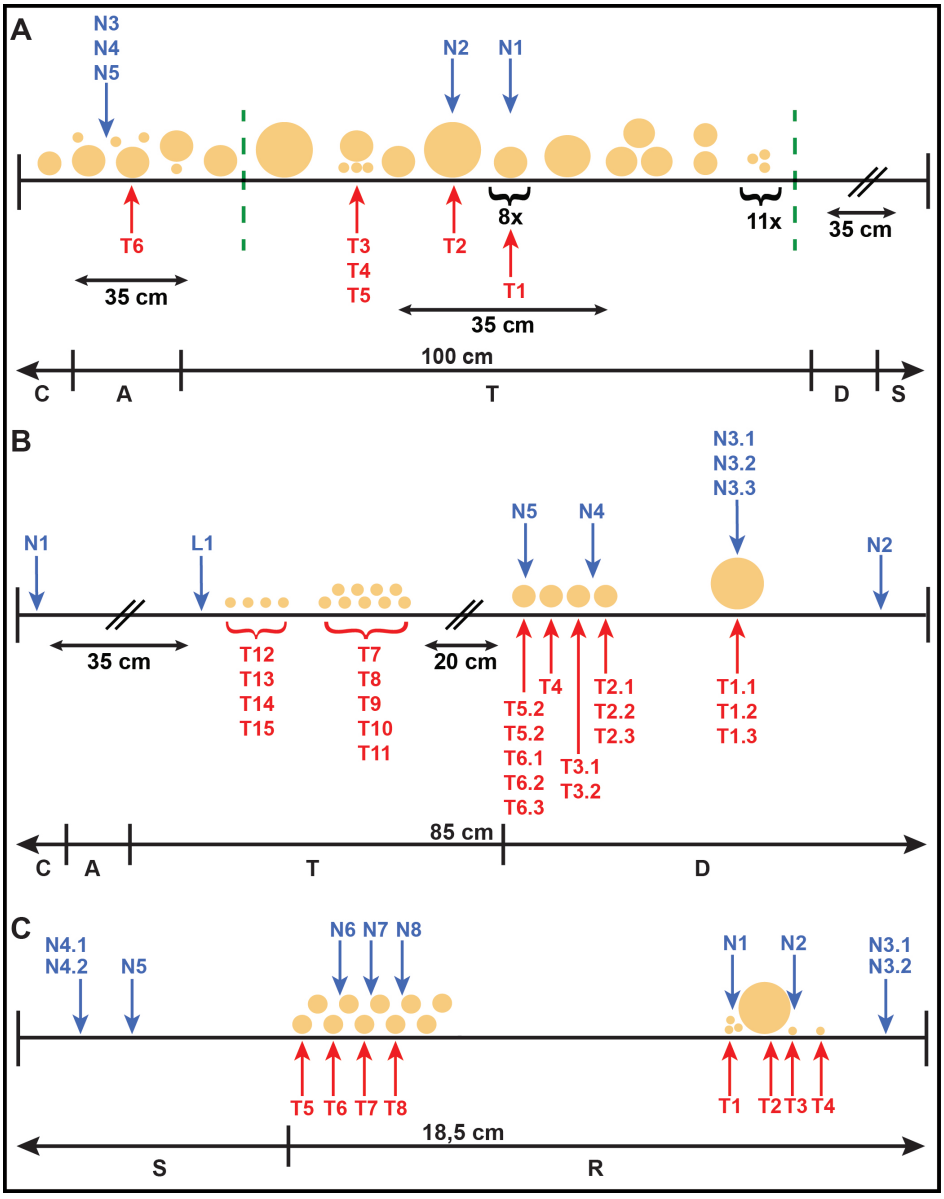


Figure 1: Mosaic patterns of patient APC-03, APC-07 and APC-17

Representation of hemicolectomy patient APC-03(A) and APC-07(B) and rectosigmoid-resection of patient APC-17(C) with location adenoma-derived (T), normal mucosa (N) and lymph node (LN) samples. C, cecum; A, colon ascendens; T, colon transversum; D, colon descendens; S, sigmoid colon; R, rectum. Size of the adenomas in the figure is proportional to real adenoma size, ranging between 4 mm and 3 cm diameter.

References

1. Hes FJ, Nielsen M, Bik EC, et al. Somatic APC mosaicism: an underestimated cause of polyposis coli. Gut 2008;57:71-6.
2. Aretz S, Stienen D, Friedrichs N, et al. Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). Hum Mutat 2007;28:985-92.
3. Tuohy TM, Burt RW. Somatic mosaicism: a cause for unexplained cases of FAP? Gut 2008;57:10-2.
4. Spier I, Drichel D, Kerick M, et al. Low-level APC mutational mosaicism is the underlying cause in a substantial fraction of unexplained colorectal adenomatous polyposis cases. J Med Genet 2016;53:172-9.
5. Out AA, van Minderhout IJ, van der Stoep N, et al. High-resolution melting (HRM) re-analysis of a polyposis patients cohort reveals previously undetected heterozygous and mosaic APC gene mutations. Fam Cancer 2015;14:247-57.
6. Gilbert SF. Early Mammalian Development. Developmental Biology 6th edition ed. Sunderland (MA): Sinauer Associates, 2000.
7. Bragg LM, Stone G, Butler MK, et al. Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. PLoS Comput Biol 2013;9:e1003031.
8. Galandiuk S, Rodriguez-Justo M, Jeffery R, et al. Field cancerization in the intestinal epithelium of patients with Crohn's ileocolitis. Gastroenterology 2012;142:855-864 e8.
9. Hes FJ, Ruano D, Nieuwenhuis M, et al. Colorectal cancer risk variants on 11q23 and 15q13 are associated with unexplained adenomatous polyposis. J Med Genet 2014;51:55-60.
10. Elsayed FA, Kets CM, Ruano D, et al. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. Eur J Hum Genet 2014 Aug; 23(8):1080-4
11. Lurkin I, Stoehr R, Hurst CD, et al. Two multiplex assays that simultaneously identify 22 possible mutation sites in the KRAS, BRAF, NRAS and PIK3CA genes. PLoS One 2010;5:e8802.
12. Wang K, Mingyao L, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research 2010;38:e164.
13. Tavtigian SV, Deffenbaugh AM, Yin L, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 2006;43:295-305.

Supplementary Methods

Study subjects
In total, 27 patients and two positive controls were included. Patients presented with 5-20 adenomas (n=6), 21-100 (n=18), 1000 adenomas (n=1) or multiple colorectal carcinomas (n=2), while the two positive controls both had 21-100 adenomas. All patients were negative for germline *APC/MUTYH* variants by Sanger sequencing and Multiplex Ligation-dependent Probe Amplification (MLPA). In addition, the cohort included patients previously screened for germline (mosaic) *APC* variants by denaturing gradient gel electrophoresis (DGGE)[1], the protein truncation test (PTT)[1] and high-resolution melting analysis (HRMA)[5]. Patients were also screened for CRC risk-associated SNPs[9] and *POLE* and *POLD1* hotspot mutations[10] (see Supplementary Table 1). No pathogenic germline (including mosaic) *APC*, *MUTYH, POLE* or *POLD1* variants were detected in these patients.

Demographic and clinical data and informed consent were obtained during the consultation. The study was approved by the LUMC medical ethical committee (P01-019E). Patients presented with polyposis affecting the entire colon (n= 12) or segmental polyposis, either right-sided (n=2) or left-sided (n= 9) (see Table 1). The location of the adenomas was unknown for four patients, and two patients had multiple colorectal cancers but no adenomas.

Of the 29 patients included, 11 patients represented a retrospective cohort collected between 2000 and 2013. Nine of these patients presented with multiple colorectal adenomas (5-100), while one patient showed 1000 adenomas distributed throughout the colon. The final patient in this group (APC-12) presented with two primary colorectal carcinomas before the age of 50. The average age at diagnosis was 46.8 years (range 26 – 63 years). In addition, sixteen patients were prospectively recruited.

All but one presented with 20-100 adenomas, with an average age at diagnosis of 60.1 years (range 33-75 years). One patient (APC-11) was diagnosed with four primary colorectal carcinomas at the age of 53. Adenomas from two positive controls were also included. These patients (APC-08 and APC-09) were previously tested for mosaic *APC* variants in a pilot study using high resolution melting analysis (HRMA).[5] Patient APC-08 carried an *APC* frameshift variant (c.4666dupA) in both adenomas tested, while patient APC-09 displayed an *APC* nonsense variant (c.4057G>T) in both adenomas tested (see Supplementary Table 1).

Twenty-three tumor or adenoma-derived DNA samples from patient APC-07 were analyzed with Sanger sequencing for six *APC* hotspot variants (c.3340C>T, c.3927_3931delAAAG, c.4348C>T, c.4348C>G, c.4391_4394delAGAG and c.4666delA) at the Erasmus MC, Rotterdam. DNA isolated from the muscularis propria of the ileum, one draining lymph node from the colon and several normal mucosa samples throughout the colon and near the adenomas was tested with SNAPshot[11] and mutation-specific PCR (*APC* mutation-specific primers available on request). Additionally, DNA from rectal mucosa, small bowel mucosa and normal colon mucosa was analyzed by deep sequencing (see Figure 1).

Tissue micro-dissection and DNA extraction
Formalin-fixed paraffin embedded (FFPE) tissue blocks were collected for all patients. Hematoxylin and eosin-stained tumor tissue slides were examined for tissue sections with tumor percentages >20%. After examination, 5 to 10 μm sections were prepared and stained with hematoxylin (eosin staining was omitted to preserve the integrity of the DNA). After staining, slides were visualized with an inverted microscope and manually microdissected with a sharp, pointed knife. When frozen tissue was used, tumor enrichment was achieved by removing non-tumorous tissue as much as possible after frozen section analysis. DNA was isolated with the Nucleospin® Tissue kit (Bioké, Leiden, the Netherlands). If possible, DNA from normal colon mucosa was also isolated from FFPE tissue blocks.

Target enrichment, DNA sequencing and data analysis
A custom *APC* Panel was designed with the Ion AmpliSeq™ Designer tool. The complete sequencing panel consisted of 115 amplicons (11216 bp), covering 99.3% of the coding regions of *APC*. Next-generation sequencing of *APC* was performed with the Ion PGM™ System (Life Technologies, Carlsbad, CA, USA). Raw data analysis, alignments, and variant calling was done using the default parameters in Torrent Suite v4.0. The Variant Caller Parameter Setting was set on 'Somatic – PGM – Low Stringency'. Variants were functionally annotated using ANNOVAR.[12] Variants were annotated to the Genbank reference sequence NM_000038.5. Coding variants were analyzed for their effect on function with *in silico* protein prediction software Align GVGD[13], SIFT (*http://sift.jcvi.org/*), PolyPhen-2 (*http://genetics.bwh.harvard.edu/pph2/*) and MutationTaster (*http://www.mutationtaster.org/*). The Leiden Open Variation Database (LOVD, http://www.lovd.nl/APC) and the catalogue of somatic mutations in cancer (COSMIC, *http://cancer.sanger.ac.uk/cosmic*) were consulted to find variants previously described and classified. All variants predicted to affect function were visually inspected with the Integrative Genomics Viewer (IGV, *https://www.broadinstitute.org/igv/home*) and confirmed by Sanger sequencing. Loss of heterozygosity (LOH) was called if all *APC* SNPs showed a variant frequency between 10 – 30% or between 60-90% (Supplementary Table 2).

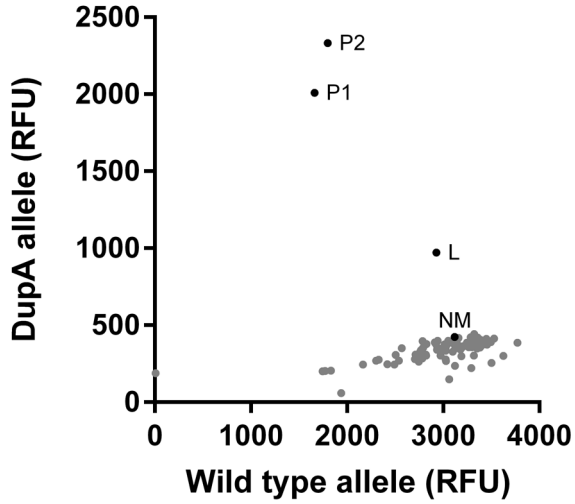
Leukocyte and/or DNA from normal colon mucosa of all patients with identical *APC* variants in multiple adenomas was screened for a specific mosaic *APC* variant. Primers amplifying specific amplicons containing the variant were used for deep sequencing (coverage minimal 1000x). All other available tumor or adenoma-derived DNA was also tested for the mosaic *APC* variant in order to study the pattern of mosaicism (Supplementary Table 1).

Confirmation of variants using KASPar or HRMA
Confirmation of the *APC* c.2493dupA variant (patient APC-18) was done using the competitive allele-specific PCR (KASPar) assay, following the manufacturer's protocol (LGC Genomic, Berlin, Germany). The primers were designed using Primerpicker (KBioscience, Hoddesdon, UK) and the following primers were used: APC_c.2493dupA_C1; **5'** GGC AAC ATG ACT GTC CTT TCA CCA T-3', APC_c.2493dupA_A1; 5'- GAA GGT GAC CAA GTT CAT GCT CCT CTT GAT GAA GAG GAG CTG GGT A-3' and APC_c.2493dupA_A2; 5' GAA GGT CGG AGT CAA CGG ATT CTC TTG ATG AAG AGG AGC TGG GTT- 3'. Variants were identified using CFX manager software v3.0 (Bio-rad, Veenendaal, the Netherlands). Confirmation of the variant is shown in Supplementary Figure 1.

The *APC* c.4666dupA variant (patients APC-04 and APC-08) is a duplication in a homopolymeric nucleotide stretch and could in many cases be a false positive as the NGS workflow used (PGM) is known to produce errors in homopolymeric sequences.[7] This variant was confirmed by high resolution melting analysis (HRMA). Sixteen adenoma and 11 matched normal mucosa DNA samples (normal colon mucosa isolated from the same FFPE block as matched tumor) were available for patient APC-04 and were tested for the variant. HRMA and HRMA data analysis were performed according to the LUMC clinical diagnostics protocol, as previously described.[5]

Testing offspring

The offspring above the age of 18 years of all patients within the cohort was offered genetic screening if a mosaic variant was found in the index patient, irrespective whether the mutation was present in adenomas, normal mucosa and/or leukocyte DNA. (Supplementary Table 1)



Supplementary Figure 1: KASPar assay APC c.2493dupA

KASPar assay of the APC c.2493dupA variant. In black two positive controls (P1 and P2), the leukocyte DNA sample (L) with 7% mutant allele and the normal mucosa (NM) sample. Cluster analysis of the endpoint fluorescence of the KASPar assay shows three distinct clusters. The two adenoma samples cluster together and display a heterozygous phenotype, containing the wildtype and mutant allele. All negative controles (in grey) cluster together and are shown to be homozygous wildtype. The normal mucosa sample falls within this cluster. The leukocyte DNA sample shows a low, but certainly visible percentage of mutant allele.

Supplementary Table 1: Cohort information

ID	Age	#ad	Age*	#ad*	Location	Type	Tumor	#T	#NM	Previously tested
Patients with mosaic *APC* variant										
APC-03	54	100	-	-	**CATDS**R	Whole colon	-	6	5	DGGE, HRM, POLE/POLD1, PTT, Risk-SNPs
APC-04	54	38	-	-	CATD**SR**	Segmental (Left)	CRC54	16	11	-
APC-07	48	25	-	-	**CATD**SR	Whole colon	-	23	7	-
APC-08*	27	80	35	100	Unknown	Unknown	-	2	-	HRM, POLE/POLD1
APC-09*	17	35	25	62	CATD**SR**	Segmental (Left)	-	6	1	DGGE, HRM, POLE/POLD1, PTT
APC-15	33	30	-	-	**CADSR**	Whole colon	-	3	-	POLE/POLD1, Risk-SNPs
APC-17	37	10	42	100	CATD**SR**	Segmental (Left)	Sig37, Rect37	8	12	DGGE, POLE/POLD1, PTT, Risk-SNPs
APC-18	35	20	47	80	**S**	Segmental (Left)	-	4	2	DGGE, HRM, POLE/POLD1, PTT, Risk-SNPs
APC-21	67	40	-	-	**CATDSR**	Whole colon	-	6	-	-
APC-24	51	58	-	-	**T**	Whole colon	CRC51 3x	3	-	POLE/POLD1
APC-25	67	30	-	-	**SR**	Segmental (Left)	-	3	-	-
Patients without mosaic APC variant										
APC-01	45	5	-	-	**CATS**	Whole colon	CRC41, CRC45, Des52	2	-	-
APC-02	54	5	-	-	CATD**S**R	Segmental (Left)	-	4	-	-
APC-05	33	1000	-	-	**ATSR**	Whole colon	-	13	-	POLE/POLD1
APC-06	63	10	-	-	CATD**SR**	Segmental (Left)	CRC63	4	-	-
APC-10	65	>20	-	-	**ATDR**	Whole colon	Rect64, 3x Ad duo	2	-	-
APC-11	53	0	-	-	C**ATD**SR	No polyposis	CRC53 4x	3	-	-
APC-12	40	0	-	-	**T**	No polyposis	CRC40, CRC49	2	-	DGGE, POLE/POLD1, PTT
APC-13	62	5	-	-	Unknown	Unknown	Sig62, Abd.desm.	2	-	-
APC-14	26	40	30	65	Unknown	Unknown	-	12	-	POLE/POLD1, Risk-SNPs

Supplementary Table 1: continued

ID	Age	#ad	Age*	#ad*	Location	Type	Tumor	#T	#NM	Previously tested
APC-16	76	50	-	-	**AS**	Whole colon	CRC74, Sig74	2	-	-
APC-19	61	20	-	-	**CT**	Segmental (right)	-	2	-	-
APC-20	66	20	-	-	**CATDSR**	Whole colon	Pr65	4	-	POLE/POLD1
APC-22	51	34	-	-	Unknown	Unknown	-	2	-	-
APC-23	63	>27	-	-	**TD**	Segmental (left)	CRC63, BHD	4	-	POLE/POLD1
APC-26	67	>50	71	100	**CATDS**	Whole colon	CRC67	3	-	-
APC-27	75	21	-	-	**CAT**	Segmental (right)	-	3	-	-
APC-28	64	>45	-	-	**S**	Segmental (left)	-	3	-	-
APC-29	67	17	68	23	**CATD**	Whole colon	CRC67	4	-	POLE/POLD1

Age and #ad indicates the age and number of adenomas at first diagnosis. Age and #ad* shows cumulative age and cumulative number of adenomas ('-' if no additional data was available). The column 'Location' shows the location of the adenomas and/or carcinomas in bold/underlined if adenomas were present, not in bold if no adenomas were present. If a letter is absent, the location is not described in pathology report, and no information is available. c = cecum, a = colon ascendens, t = colon transversum, d= colon descendens, s= sigmoid, r = rectum. Tumor shows tumortype followed by age of diagnosis. CRC = colorectal carcinoma, Des = desmoid carcinoma, Sig = sigmoid carcinoma, Ad duo = adenoma in duodenum, Rect = rectum carcinoma, Abd.desm = abdominal desmoids, Pr = prostate cancer, BHD = Birt-Hogg-Dubé Syndrome. Some patients were previously screened for mosaic APC variants with denaturing gel gradient electrophoresis (DGGE), high-resolution melting (HRM), the protein truncation test (PTT) or were tested for variants in POLE/POLD1 (POLE/POLD1) or for the presence of CRC-risk associated SNPs in APC (Risk-SNPs). *positive control*

Supplementary Table 2: Overview of *APC* variants found in tested adenomas

Patient-ID		P1-Variant	%	Cov.	P2-Variant	%	Cov.	P3-Variant	%	Cov.	P4-Variant	%	Cov.
APC-01	**2**	-	-	-	c.3337_3338del	17%	1257						
APC-02	**2**	c.4069G>T LOH?	24% -	1635	c.3964G>T LOH?	38% -	40						
APC-03	**2**	c.3904delC c.694C>T	31% 34%	2445 7954	c.3904delC LOH	83% -	2533	c.3904delC	74%	27			
APC-04	**4**	c.4666dupA LOH	58% -	1818	c.4666dupA LOH	45% -	967	c.4666dupA LOH	41% -	443	c.4655_4656del	51%	542
APC-05	**4**	c.3880C>T c.4221delT	34% 24%	1490 2541	-	-	-	c.1279_1289del	32%	1762	c.4580_4589del	37%	1205
APC-06	**4**	847C>T	31%	2241	c.2496delC c.4099C>T	39% 43%	2731 2784	c.4611_4612delAG c.4501delT	37% 39%	894 512			
APC-07	**23**												
APC-08	**2**	c.4666dupA	42%	487	c.4666dupA	40%	2061						
APC-09	**2**	c.4057G>T	27%	699	c.4057G>T	28%	953						
APC-10	**2**	c.3030delT	64%	4251	c.2991T>A LOH	68% -	2789 -						
APC-11	**3**	c.4468delC c.1660C>T	23% 25%	1658 1667	c.637C>T c.4228_4229insT	37% 37%	1052 2648	c.3639_3650delinsT	42%	2202			
APC-12	**2**	-	-	-	-	-	-						
APC-13	**2**	c.4314delA LOH	66% -	1296 -	-	-	-						
APC-14	**2**	-	-	-	-	-	-						
APC-15	**3**	c.2566_2572del c.4429C>T	28% 26%	1555 >2000	c.2566_2572del c.4737delT	28% 27%	2000 1979	c.2566_2572del c.4660G>T	44% 40%	>2000 1934			
APC-16	**2**	-	-	-	-	-	-						
APC-17	**2**	c.4110_4111del LOH	66% -	2165	c.4110_4111del LOH	59% -	865						

Supplementary Table 2 : continued

Patient-ID		P1-Variant	%	Cov.	P2-Variant	%	Cov.	P3-Variant	%	Cov.	P4-Variant	%	Cov.
APC-18	**2**	c.2493dupA	38%	846	c.2493dupA	33%	803						
APC-19	**2**	c.4348C>T c.3467_3470del	17% 26%	451 1984	-	-	-						
APC-20	**4**	c.4233delT	30%	2222	-			-			c.1349delT	31%	3322
APC-21	**6**	c.1959-1G>A	30%	>2000	c.1959-1G>A	14%	>2000	c.1959-1G>A c.3982C>T	30% 30%	>2000 2000	c.1959-1G>A c.4062_4063del c.1972G>T	30% 40% 30%	>2000 1982 1998
APC-22	**2**	c.2626C>T c.2805C>A	23% 22%	1999 1990	-	-	-	c.591_592del c.4157delG	17% 18%	1985 22			
APC-23		c.3991A>T c.847C>T	16% 16%	>2000 >2000	c.2626C>T c.4240delG	47% 31%	2000 62	c.2413C>T c.4233delT	19% 52%	>2000 58	c.1495C>T	41	>2000
APC-24	**3**	c.1974_1975del LOH	70% -	>2000 -	c.1974_1975del	27%	>2000	c.1974_1975del c.994C>T	40% 39%	>2000 >2000			
APC-25	**3**	c.3211C>T	26%	>2000	c.3211C>T c.4057G>T	37% 24%	>2000 >2000	c.3211C>T	34%	1987			
APC-26	**3**	-	-	-	-	-	-	-	-	-			
APC-27	**4**	c.5771_5778del c.3927_3931del	32% 26%	2000 >2000	-	-	-	-	-	-	c.694C>T c.1213C>T	12 11	319 >2000
APC-28	**3**	-	-	-	c.694C>T	28%	94	c.2626C>T c.4271delC	42% 45%	1998 1998			
APC-29	**4**	c.2626C>T c.4348C>T	33% 34%	1998 2000	c.3682C>T c.4733_4734del	37% 36%	2000 1979	c.2932C>T c.4358delC	12% 12%	1998 1952	c.1269G>A c.4459dupA	46 41	1996 1993

n indicates the number of adenomas initially tested. In patients with identical mutations in multiple adenomas, all available tumor material was tested. % shows the variant frequency, cov. indicates the coverage in reads per amplcion. '–' adenoma is sequenced, but no APC variant was detected. LOH indicates loss of heterozygosity. Patient APC-07 was analyzed for APC hotspot mutations and the c.3340C>T was found in all 23 adenoma-derived samples (not shown in this table