npg

ORIGINAL ARTICLE

Inactivation of promoter 1B of APC causes partial gene silencing: evidence for a significant role of the promoter in regulation and causative of familial adenomatous polyposis

A Rohlin[1], Y Engwall[1], K Fritzell[2], K Göransson[2], A Bergsten[1], Z Einbeigi[3], M Nilbert[4,5], P Karlsson[3], J Björk[2] and M Nordling[1]

[1]Department of Clinical Genetics, Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg, Sahlgrenska University Hospital, Gothenburg, Sweden; [2]The Swedish Polyposis Registry, Department of Gastroenterology and Hepatology, Karolinska Institute, Karolinska University Hospital, Stockholm, Sweden; [3]Department of Oncology, Sahlgrenska Academy at University of Gothenburg, Sahlgrenska University Hospital, Gothenburg, Sweden; [4]Department of Oncology, Clinical Sciences, Lund University, Lund, Sweden and [5]Clinical Research Centre, Hvidovre Hospital, Copenhagen University, Hvidovre, Denmark

Familial adenomatous polyposis (FAP) is caused by germline mutations in the adenomatous polyposis coli (*APC*) gene. Two promoters, 1A and 1B, have been recognized in *APC*, and 1B is thought to have a minor role in the regulation of the gene. We have identified a novel deletion encompassing half of this promoter in the largest family (Family 1) of the Swedish Polyposis Registry. The mutation leads to an imbalance in allele-specific expression of *APC*, and transcription from promoter 1B was highly impaired in both normal colorectal mucosa and blood from mutation carriers. To establish the significance of promoter 1B in normal colorectal mucosa (from controls), expression levels of specific transcripts from each of the promoters, 1A and 1B, were examined, and the expression from 1B was significantly higher compared with 1A. Significant amounts of transcripts generated from promoter 1B were also determined in a panel of 20 various normal tissues examined. In FAP-related tumors, the *APC* germline mutation is proposed to dictate the second hit. Mutations leaving two or three out of seven 20-amino-acid repeats in the central domain of APC intact seem to be required for tumorigenesis. We examined adenomas from mutation carriers in Family 1 for second hits in the entire gene without any findings, however, loss of the residual expression of the deleterious allele was observed. Three major conclusions of significant importance in relation to the function of *APC* can be drawn from this study; (i) germline inactivation of promoter 1B is disease causing in FAP; (ii) expression of transcripts from promoter 1B is generated at considerable higher levels compared with 1A, demonstrating a hitherto unknown importance of 1B; (iii) adenoma formation in FAP, caused by impaired function of promoter 1B, does not require homozygous inactivation of *APC* allowing for alternative genetic models as basis for adenoma formation.

Introduction

Familial adenomatous polyposis (FAP) is caused by dominant inheritance of germline mutations in the adenomatous polyposis coli (*APC*) tumor suppressor gene. In classical FAP, mutation carriers develop hundreds to thousands of adenomas in the large bowel and untreated carriers inevitably develop colorectal cancer at the age of 35–40 years (Bisgaard *et al.*, 1994). Patients with FAP also have increased risk of developing extra-colonic manifestations, such as duodenal adenomas, fundic gland polyps, desmoid tumors, osteomas and lipomas. Desmoid tumors have a prevalence of 10–20% in FAP patients, and are a major cause of morbidity and mortality. A milder form of FAP also exists referred to as attenuated polyposis (AFAP).

The *APC* gene (5q21–q22) encodes several tissue-specific transcripts in which the major transcript has an open reading frame of 8543 bp (Horii *et al.*, 1993), two promoter regions have been identified, 1A and 1B (Lambertz and Ballhausen 1993; De Rosa *et al.*, 2007; Hosoya *et al.*, 2009). Promoter 1B has not been extensively studied, and has been proposed to have a minor role in the regulation of the *APC* gene compared with promoter 1A (Charames *et al.*, 2008, Segditsas *et al.*, 2008, Tsuchiya *et al.*, 2000). More than 1400 disease-causing mutations causing premature truncation or absence of the APC protein are included in the Human Gene Mutation Database (Stenson *et al.*, 2009), and the mutation detection frequency generally reported in families with a classical FAP is around 80–90%. A number of screening methods to detect mutations in exons of *APC* exists, but as sequence-based methods gradually are becoming more available and easy to use,

this is the screening method that today dominate in routine diagnostics. Together with multiplex ligation-dependent probe amplification (MLPA; Schouten et al., 2002), widely used to detect copy-number variations (CNVs) in defined regions, the majority of disease-causing mutations can be identified. However, detection of point mutations and CNVs located in introns and in the regulatory regions of the gene still constitute a challenge. Imbalance in allele-specific expression (ASE) of the two APC alleles as a cause of FAP has been recognized in several families (Castellsague et al., 2010a). In majority of the cases, no deleterious mutation explaining the lowered expression levels has been found. Higher proportions of patients without extra-colonic manifestations have been observed in these families and in APC-mutation-negative FAP families in general (Moisio et al., 2002; Renkonen et al., 2005).

In this study, we have screened 54 FAP/AFAP families for mutations in APC and in the largest and first enrolled family in the Swedish Polyposis Registry, Family 1, we identified two adjacent large novel deletions in the regulatory region of the gene. One of the deletions included more than half of the promoter 1B. We have previously identified imbalance in ASE in this family (Kanter-Smoler et al., 2008). The family includes 150 individuals (seven generations), of whom 59 are affected.

Results

Mutation screening

In this study, 54 families with FAP or AFAP were screened for mutations in APC. A deleterious mutation was found in 97% of families with FAP and 26% of families with AFAP, two of these families carried bi-allelic MUTYH mutations (Table 1). We have previously reported imbalance in ASE in Family 1 from the Swedish Polyposis Registry (Kanter-Smoler et al., 2008), but the mutation remained undetected, and that is why we undertook this study in an attempt to reveal the mutation.

Analysis of CNVs in the APC region of Family 1 and characterization of the deletions

Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) was used to analyze for CNVs (using copy-number (CN) probes and single-nucleotide polymorphism (SNP) probes) in the chromosomal region of APC. Using this approach, a large deletion of ~61 kb was found in the regulatory region of the APC gene (Figure 1a). Primer walking and sequencing of PCR products revealed two adjacent deletions (Del 1 and Del 2), encompassing 18.7 kb and 42.1 kb, respectively. The two deletions were separated by a retained fragment of 1540 bp (Figure 1b). The deletions are located 46 kb upstream of the coding region of APC and the precise location of Del 1 was chr5:112,008,797_112,027,564; and Del 2; chr5:112,029,076_112,071,227 according to the University of California, Santa Cruz reference sequence March 2006 (National Center for Biotechnology Information/hg18). Six affected and two non-affected individuals were analyzed on the array platform and all affected, but not the unaffected, individuals carried the deletions. The deletions contained 320 bp of promoter 1B and further upstream highly conserved regions and regulatory elements including six transcription factor-binding sites (interferon-regulatory factor-1, Nkx25, TATA, POU6F1 and two OCT1-binding sites). The most highly conserved region was the binding site for the most upstream located OCT1 site.

Five additional affected and two unaffected family members from Family 1 and 50 control individuals were analyzed by PCR amplification to confirm the segregation of the deletion with the disease. The deletions were detected in all of the affected individuals, but not in the unaffected or in the normal controls. Amplification of the deleted region by two different primer pairs in one affected individual is shown in Figure 1c. The same region in unaffected family members is not possible to amplify because of its size.

Determination of expression levels from promoter 1A and 1B in blood from Family 1 using quantitative–PCR (Q–PCR) and expression arrays

The consequence the deletions had on the transcription of the APC gene was investigated using Q–PCR (TaqMan; Applied Biosystems, Foster City, CA, USA) and exon-specific expression arrays (Affymetrix Gene Chip Human Exon 1.0 ST Array). A reduced expression of APC was detected in affected individuals compared with control individuals with both methods. A schematic illustration over the locations of the different TaqMan probes is presented in Figure 2. APC1 binds to transcripts generated from the promoter 1B; APC2 and APC3 bind to different transcripts from promoter 1A and APC4; and APC5 binds to all transcripts generated from both promoters 1A and 1B. Approximately a threefold reduction in relative expression values was observed in affected individuals compared with a healthy family member and controls for APC1 (1B) and APC4 (1A + 1B) and at least a twofold reduction with APC5 (1A + 1B). The probes APC2 (1A) and APC3 (1A) demonstrated unexpectedly at least a threefold higher expression in the affected patients compared with normal controls (Figure 3). For a more comprehensive illustration of the Q–PCR data, we have used normalization where the sample with the lowest expression for each assay is set to zero. Statistical significance was determined with the t-test, and was calculated for all of the assays. All assays had P-values <0.001.

Analyses using the expression array resulted in a twofold reduction in the overall expression of the APC gene in affected family members compared with five controls (data not shown). Twofold reduction in expression represents a theoretically 50% downregulation in gene expression. The expression array analysis of affected individuals in Family 1 also revealed a splice variant lacking exon 2 (besides the lower APC expression, Figure 4). This splice variant was not

Table 1 Mutations in FAP and AFAP patients

Patient	FAP/AFAP	Location	Mutation	Reference
C152, P4 (Family 1)	FAP	Promoter 1B (320 bp deletion)	chr5:g.[112,008,797_112,027,564del;112,029,076_112,071,227del]	This study
C1929	FAP	Gene	chr5:g.(112,112,483_112,112,538)_(112,226,776_112,232,783)del	This study
C1040	FAP	Promoter 1A and gene	chr5:g.(112,091,076_112,092,095)_(112,208,168_112,209,500)del	This study
C1466	FAP	Promoter 1A, 1B and gene	chr5:g.(106,255,243_106,255,351)_(112,284,218_112,288,290)del	This study
C2348	FAP	Promoter 1A, 1B and gene	chr5:g.(110,965,499_110,966,476)_(112,997,172_112,997,631)del	This study
C2474	AFAP	Exon 2–3 (duplication)	NM_000038.3:c.136_422dup, p.Ser142_Val2843delins23X	Mutoh et al. (2006)
C1563	AFAP	Exon 3	NM_000038.3:c.422-3_456del, p.Arg141GlyfsX14	This study
C2420	FAP	Exon 5	NM_000038.3:c.637C > T, p.Arg213X	Miyoshi et al. (1992)
C2027	FAP	Exon 6	NM_000038.3:c.694C > T, p.Arg232X	Miyoshi et al. (1992)
C2462	FAP	Exon 7	NM_000038.3:c.790C > T, p.Gln264X	Cowie et al. (2004)
C1946	AFAP	Exon 9	NM_000038.3:c.1263_1272delins19, p.Trp421X	This study
C1079	FAP	Exon 11, intron 10	NM_000038.3:c.1409-1G > A, p.?	Kaufmann et al. (2009)
C1867	FAP	Exon 11, intron 10	NM_000038.3:c.1409-1G > A, p.?	Kaufmann et al. (2009)
C2240	FAP	Exon 11, intron 10	NM_000038.3:c.1409-1G > A, p.?	Kaufmann et al. (2009)
C1858	FAP	Exon 11–13 (deletion)	NM_000038.3:c.1419_1902del, p.Ala474GlufsX11	Andreutti-Zaugg et al. (1999)
C1992	FAP	Exon 11, intron 11	NM_000038.3:c.1548 + 1G > A, p.?	Kohoutova et al. (2002)
C1097	FAP	Exon 13	NM_000038.3:c.1660C > T, p.Arg554X	Fodde et al. (1992)
C2321	FAP	Exon 15	NM_000038.3:c.2006T > G, p.Leu669X	This study
C1787	FAP	Exon 15	NM_000038.3:c.2098dup p.Asp700GlyfsX6	This study
C1759	FAP	Exon 15	NM_000038.3:c.2413C > T, p.Arg805X	Dobbie et al. (1996)
C2383	FAP	Exon 15	NM_000038.3:c.2547_2550del, p.Asp849GlufsX11	Ripa et al. (2002)
C2159	FAP	Exon 15	NM_000038.3:c.3180_3184del, p.Gln1062X	This study
C2296	FAP	Exon 15	NM_000038.3:c.3183_3187del, p.Gln1062X	Miyoshi et al. (1992)
C2609	FAP	Exon 15	NM_000038.3:c.3183_3187del, p.Gln1062X	Miyoshi et al. (1992)
C2195	FAP	Exon 15	NM_000038.3:c.3537_3538del, p.Tyr1179X	This study
C1327	FAP	Exon 15	NM_000038.3:c.3927_3931del, p.Glu1309AspfsX4	Miyoshi et al. (1992)
C2643	FAP	Exon 15	NM_000038.3:c.3927_3931del, p.Glu1309AspfsX4	Miyoshi et al. (1992)
C1579	FAP	Exon 15	NM_000038.3:c.4132del, p.Gln1378ArgfsX37	This study
C2574	FAP	Exon 15	NM_000038.3:c.4391_4394del, p.Glu1464ValfsX8	Gismondi et al. (1997)
C1893	FAP	Exon 15	NM_000038.3:c.4570del, p.Ile1524X	This study
C2300	FAP	Exon 15	NM_000038.3:c.4612_4613del, p.Glu1538IlefsX5	Gayther et al. (1994)
C1784	FAP	Exon 15	NM_000038.3:c.4638_4642del, p.Asn1546LysfsX11	Miyoshi et al. (1992)
C1897	FAP	Exon 15	NM_000038.3:c.4655_4656del, p.Glu1552GlyfsX6	Nilbert et al. (2000)
C1807	AFAP	Exon 15	NM_000038.3:c.7709dup, p.Ser2571IlefsX12	This study
C1189	FAP		No mutation detected in APC or MUTYH	
C1369	AFAP		No mutation detected in APC or MUTYH	
C1387	AFAP		No mutation detected in APC or MUTYH	
C1421	AFAP		No mutation detected in APC or MUTYH	
C1427	AFAP		No mutation detected in APC	
C1477	AFAP		No mutation detected in APC	
C1488	AFAP		No mutation detected in APC	
C1491	AFAP		No mutation detected in APC	
C1492	AFAP		No mutation detected in APC	
C1507	AFAP		No mutation detected in APC or MUTYH	
C1651	AFAP		No mutation detected in APC nor MUTYH Y179C or G396D	
C1813	AFAP		No mutation detected in APC	
C2081	AFAP		No mutation detected in APC or MUTYH	
C2110	AFAP		No mutation detected in APC nor MUTYH Y179C or G396D	
C2201	AFAP		No mutation detected in APC nor MUTYH Y179C or G396D	
C2292	AFAP		No mutation detected in APC or MUTYH	
C2341	AFAP		No mutation detected in APC nor MUTYH Y179C or G396D	
C2363	AFAP		No mutation detected in APC nor MUTYH Y179C or G396D	
C476	AFAP		MUTYH NM_001128425.1:[c.1214C > T] + [c.1214C > T], [p.P405 L] + [p.P405 L]	
C1590	AFAP		MUTYH NM_001128425.1:[c.538G > A] + [c.538G > A], [p.G189E] + [p.G189E]	

Abbreviations: AFAP, attenuated familial adenomatous polyposis; APC, adenomatous polyposis coli; FAP, familial adenomatous polyposis; NCBI, National Center for Biotechnology Information.
Genomic localizations according to University of Califonia, Santa Cruz, reference sequence March 2006 (National Centre for Biotechnology Information/hg18).

detected in unaffected individuals from the family, or in healthy controls. To elucidate the significance of this splice variant a panel of 20 normal tissues was analyzed for its presence. The splice variant seemed to be more abundant in heart and brain, though it was present in normal colon mucosa as well (data not shown).

Determination of expression levels from promoter 1A and 1B in normal tissues from healthy controls using Q–PCR
A panel consisting of 20 different normal tissues was examined for the expression of transcripts from the two promoters (Figure 5). Transcripts from promoter 1B (APC1) were expressed in all tissues. The relatively highest expression for all of the transcripts was detected in brain. In blood and colorectal mucosa, the expression from promoter 1B (APC1) is approximately equal, whereas expression from promoter 1A (APC2 and APC3) in blood is much lower compared with colorectal mucosa. (It is easy to get false impression from Figure 5 that the expression from promoter 1A is higher compared with 1B, as staples are higher, however, this only reflects the fact that the variation in relative expression levels between the tissues is greater for promoter 1A compared with 1B). An estimation of absolute expression levels from the Cq-values (Supple-
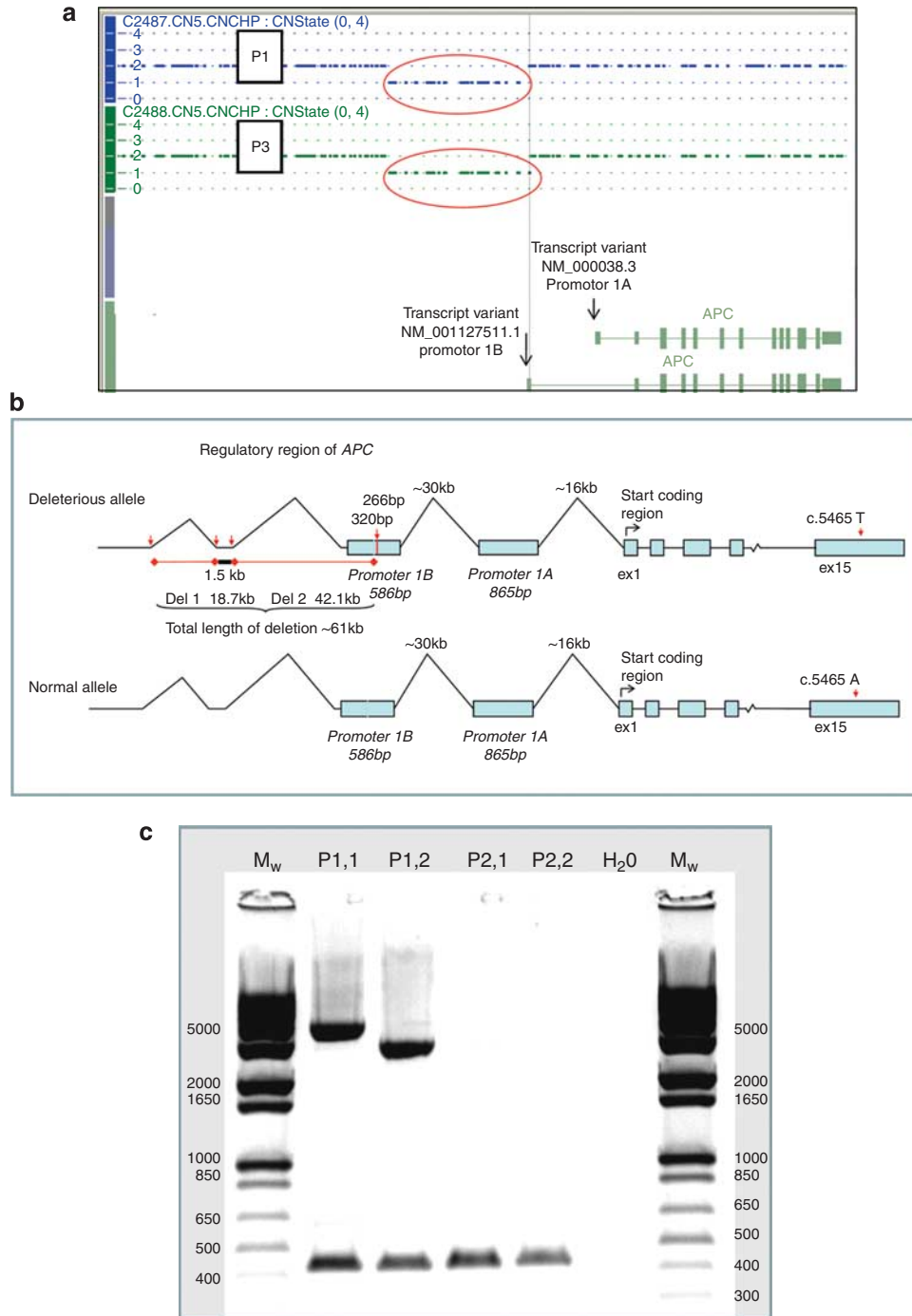
Figure 1 (**a**) The deletion in two affected members, P1 and P3, of Family 1 analyzed with Affymetrix Genome-Wide Human SNP Array 6.0 (analysis output from Genotyping Console; Affymetrix). The deletion (indicated by a red circle), which actually contains two deletions (not possible to discover in this array data), is approximately 61 kb. The transcript variant NM_000038.3 starts from promoter 1A and transcript variant NM_001127511.1 starts from promoter 1B. (**b**) Schematic illustration of the deleterious allele and the normal allele in Family 1. In the deleterious allele, the second deletion (Del 2) includes 320 bp of promoter 1B (586 bp as a whole) of the *APC* gene. The deleterious allele carries a T in the heterozygous SNP position c.5465, whereas the normal allele carries an A. (**c**) Agarose gel after separation of fragments generated from amplification of the region upstream of *APC* where the deletions are located. Sample P1,1 is a mutation carrier in which a fragment of ∼4 kb is obtained after PCR amplification with primers 256 and 343 (Supplementary Table 3), P1,2, an ∼3 kb fragment obtained by amplification of primers 256 and 1428 in the same patient. Samples P2,1 and P2,2; amplification using primers 256 together with 343 and 256 together with 1428 in an unaffected family member which should have given fragments of ∼65 kb and ∼64 kb, respectively, but fragments this long is not possible to amplify with conventional PCR. PCR-amplification of a fragment of known size (450 bp) is carried out simultaneously with the deletion-specific primers to verify the presence of DNA in the reaction tube.
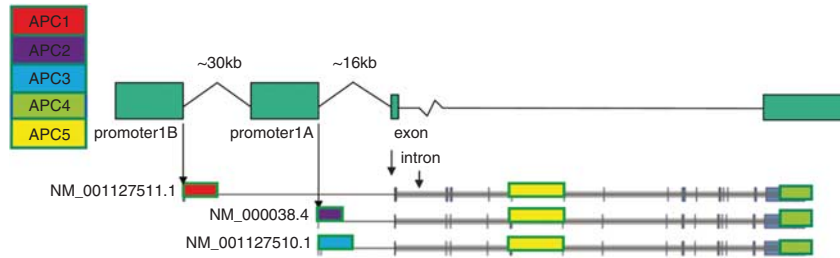
Figure 2 Schematic illustration of the location of the five different TaqMan probes used in the Q–PCR of the *APC* gene. APC1 binds specifically to the reference sequence NM_001127511.1 (probe located in exon 1B, for precise location see http://www.appliedbio-systems.com Hs01568282_m1 Taqman gene expression assay), APC2 binds specifically to the reference sequence NM_000038.4 (probe located in sequence specific for transcripts generated from promoter 1A, Chr. 5:g.112073556_112090642 (March 2006 (National Center for Biotechnology Information/hg18) map)), APC3 binds specifically to the reference sequence NM_001127510.1 (probe located in exon specific for transcripts from 1A, Chr. 5:g.112073579_112074157 (March 2006 (National Center for Biotechnology Information/hg18) map)), APC4 and APC5 binds to all of the transcripts generated from both of the promoters. (For precise location see http://www.appliedbiosystems.com Taqman gene expression assay, Hs01568270_m1 for APC4 and Hs00181051_m1 for APC5.)
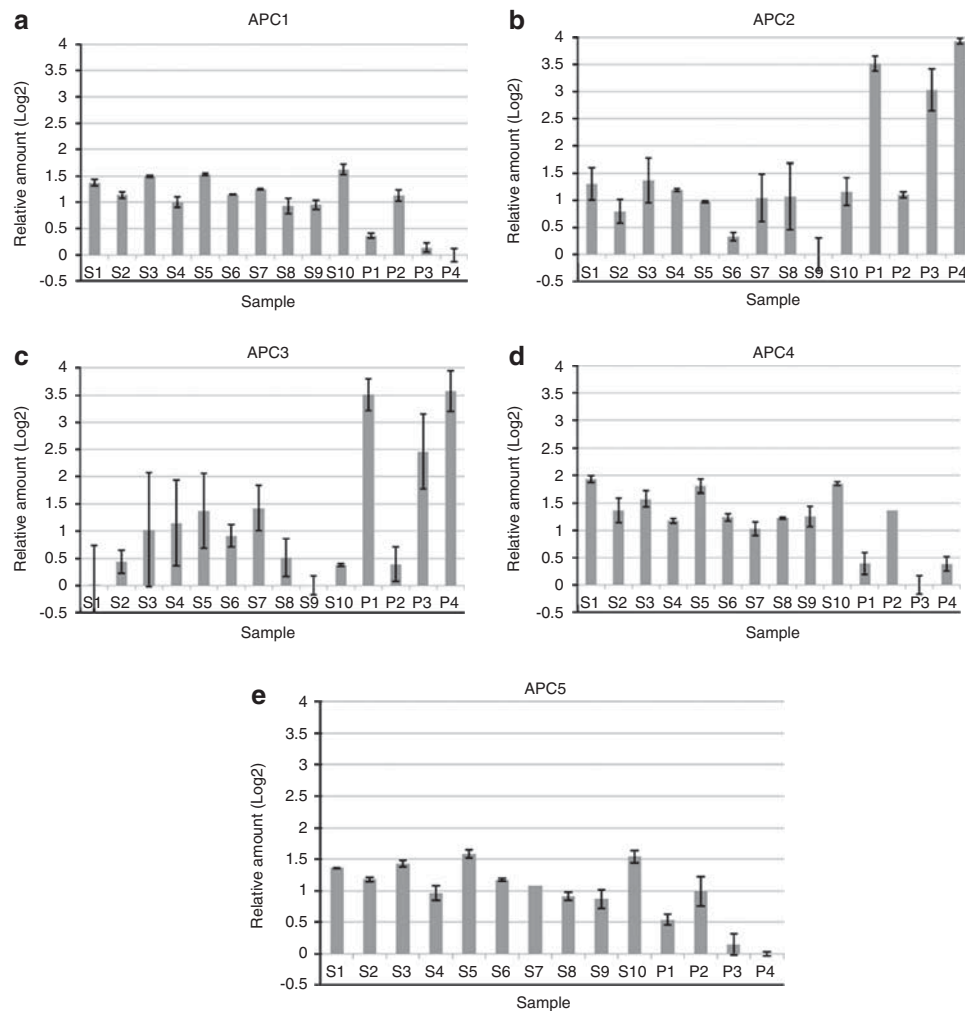

Figure 3 Expression in blood from the two promoters in patients and normal controls as measured by Q–PCR (TaqMan) by the five assays, APC1–APC5. Samples S1–S10 are normal controls and P1–P4 are individuals from Family 1 were P1, P3 and P4 are affected individuals (mutation carriers) and P2 is an unaffected individual. The data is normalized to zero for the sample with the lowest expression for each assay. The error bars indicate the standard deviation for the RT replicates. (**a**) APC1 corresponds to transcript NM_001127511.1 from promoter 1B. (**b**) APC2 corresponds to transcript NM_000038.4 from promoter 1A. (**c**) APC3 corresponds to transcript NM_001127510.1 from promoter 1A. (**d**) APC4 corresponds to all transcripts generated from both of the promoters, 1A and 1B. (**e**) APC5 corresponds to all transcripts generated from both of the promoters, 1A and 1B.
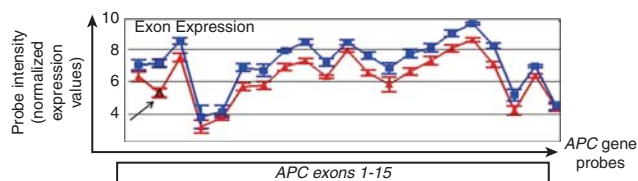
Figure 4 Partek Genomics Suites alternative splice analyses of *APC* exon expression in Family 1 obtained from analyses with Affymetrix GeneChip Human Exon 1.0 ST Arrays. Each blue dot represents the mean expression of an exon in the normal controls (mean value of three) and each red dot represents the mean expression of an exon in samples from affected individuals (mean value of three). For some exons there is more than one probe/dot. A possible splice variant lacking exon 2 is observed in the affected individuals indicated by the arrow in the figure.

mentary Table 1) gives that in blood from normal controls the transcript NM_001127511.1 (APC1) from promoter 1B is expressed at ~100-fold higher levels than the transcript NM_000038.4 (APC2) from promoter 1A, and ~1000-fold higher than transcript NM_001127510.1 (APC3) from promoter 1A.

Determination of expression levels from promoter 1A and 1B in normal colon mucosa and adenoma tissue from Family 1 using Q–PCR
The relative expression of the promoters in adenoma and normal colon mucosa from three mutation carriers and normal colon mucosa from control individuals were determined (Figure 6). The normal colon mucosa from the mutation carriers presented with lower relative expression levels for all transcripts compared with normal mucosa from control individuals. Comparison of all the three adenoma tissues with normal colon mucosa from patients demonstrated lower levels for all transcripts in two of the samples. One of the samples (P3) did not show a clear difference between normal and adenoma tissues. The reason for this could be contamination of normal tissue in the adenoma sample. For APC1, APC4 and APC5, a statistical significance of *P*-values <0.006 was calculated. For APC2 and APC3 (*P*-values <0.01), very low expression levels were observed, and this is a reason why high standard deviations are obtained. From the *C*q-values (Supplementary Table 1), an estimation of the total amount of transcription from promoter 1B (APC1) and 1A (APC2 and APC3) in normal colorectal mucosa was estimated. Transcript NM_001127511.1 (APC1, 1B) was expressed around 250 times higher than transcript NM_001127510.1 (APC3, 1A) and 25 times higher than transcript NM_000038.4 (APC2, 1A) in this tissue.

Determination of methylation status and ASE and characterization of the mechanism for APC inactivation underlying tumor initiation in Family 1
To characterize the mechanisms of inactivation of *APC* underlying tumor initiation, sequencing of the coding region of the *APC* gene including the promoter 1A, analyses of methylation status and determination of ASE in adenoma and normal colorectal mucosa were

carried out. Adenomas were available from four patients and normal colorectal mucosa from three patients.

Sequencing over a heterozygous coding SNP (c.5465 A/T, Figure 1b) on complementary DNA (cDNA) was performed to estimate total ASE of the *APC* gene. The analysis revealed the deleterious T allele to be absent in all four adenomas but still present at levels of 8–9%, in two out of three of the normal colon mucosa tissues as well as in blood lymphocytes from one individual (three adenomas are shown in Figure 7). Using methylation-specific MLPA, the methylation degree of the promoter 1A was determined to be 2, 10, 10 and 21% in the four adenomas, respectively. In normal colon mucosa, only very low degrees of methylation (3%) were observed in two out of the three individuals. The methylation degree of promoters required to have an effect on transcription has been considered to correspond to ~15% of methylated DNA (Judson *et al.*, 2006). Accordingly, only one patient from Family 1 demonstrated functional significant methylation of promoter 1A (P3, 21%) in adenoma tissue. No mutation in the *APC* gene or in the promoter 1A in adenomas was detected.

Discussion

In this study, we report two adjacent large deletions in the regulatory region of *APC* as causative for the reduced expression and development of FAP in Family 1 of the Swedish Polyposis Registry. The deletions include approximately half of the upstream region of promoter 1B. This promoter has not been extensively studied, and has been proposed to have a minor role in the regulation of the *APC* gene compared with promoter 1A (Charames *et al.*, 2008; Segditsas *et al.*, 2008; Tsuchiya *et al.*, 2000). In a recent study by Charames *et al.* (2008), the first deletion of promoter 1A was found in a family with classical FAP (the deletion did not involve the coding region of the gene). A complete silencing of the transcription from the altered allele was caused by the deletion. However, the precise limits of the deletion could not be established, and it is possible that this deletion also encompass promoter 1B. Deletions of promoter 1B solely have so far not been found in any FAP family.

To demonstrate the significance of promoter 1B, we have examined the activity of this promoter in relation to promoter 1A in normal tissues from controls. We have shown that significant levels of transcripts are generated from promoter 1B in blood and in colorectal mucosa, as well as in a number of additional normal tissues. We estimated absolute levels of expression from promoter 1B to be 25- and 250-fold higher compared with two different transcripts generated from 1A in normal colorectal mucosa and 100- and 1000-fold higher for each transcript in blood. The relatively highest levels of transcription from promoter 1B were observed in blood, brain and small intestine. In a recent study by Hosoya *et al.* (2009), transcription from promoter 1B in human normal gastric mucosa was estimated to be 15-fold higher as compared with levels from the
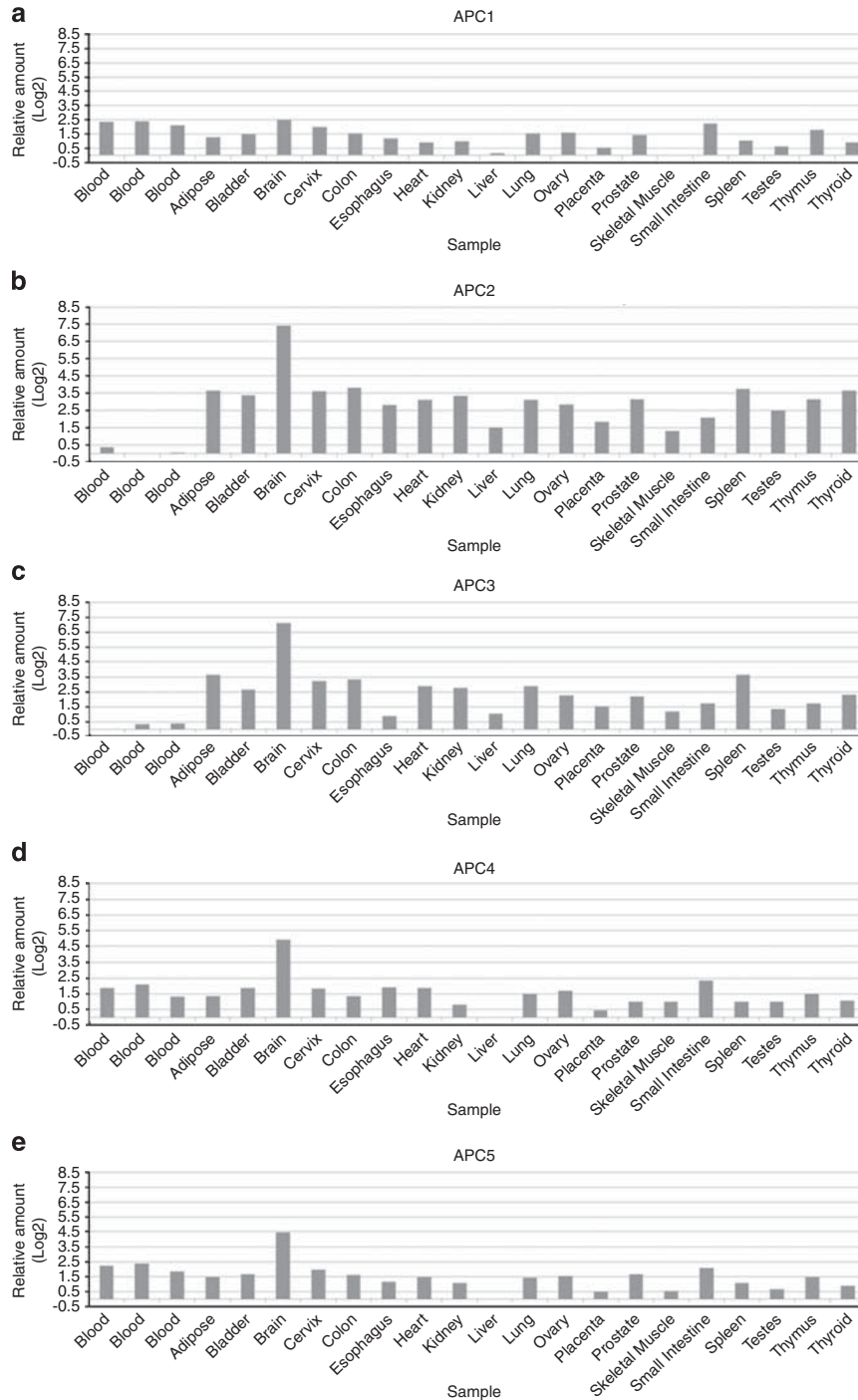
a

b

c

d

e

Figure 5 Expression from the two promoters in a panel of normal tissues as measured by Q–PCR (TaqMan) by the five assays, APC1–APC5. The data is normalized to zero for the sample with the lowest expression for each assay. (**a**) APC1 corresponds to transcript NM_001127511.1 from promoter 1B. (**b**) APC2 corresponds to transcript NM_000038.4 from promoter 1A. (**c**) APC3 corresponds to transcript NM_001127510.1 from promoter 1A. (**d**) APC4 corresponds to all transcripts generated from both of the promoters, 1A and 1B. (**e**) APC5 corresponds to all transcripts generated from both of the promoters, 1A and 1B.

promoter 1A. It is thus established that promoter 1B has an important role in the regulation of *APC* in a variety of normal tissues.

Reduced ASE of the *APC* gene has been shown to be associated with a predisposition to FAP. The degree of expression reduction for one allele has been variable, ranging from around 50% to complete silencing

(Venesio *et al.*, 2003; Renkonen *et al.*, 2005; Kanter-Smoler *et al.*, 2008; Castellsague *et al.*, 2010b). There also seems to be a phenotypic variation due to the changes in gene expression, and a threshold for *APC* product being required to suppress intestinal tumorigenesis (Solomon *et al.*, 1987; Yan *et al.*, 2002). To further assess the consequence of the deleted promoter region,
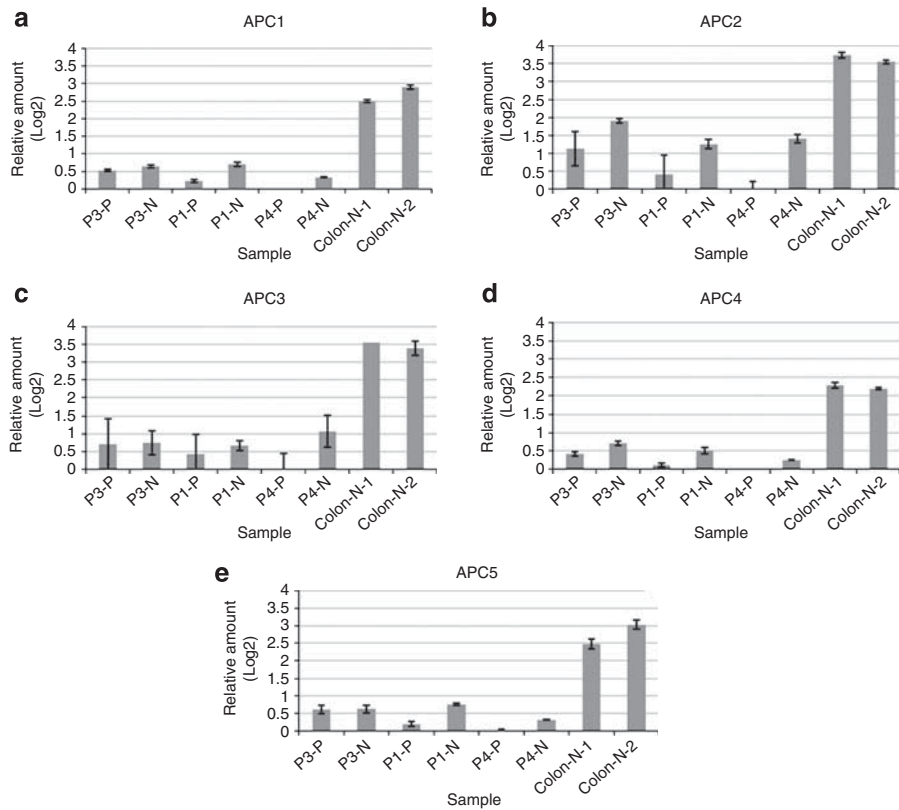
Figure 6 Expression from the two promoters measured in normal colon mucosa (P1-N, P3-N and P4-N) and adenoma (polyps, P1-P, P3-P and P4-P) from mutation carriers and normal colon mucosa from healthy controls (colon-N-1 and colon-N-2) as measured by Q–PCR by the five assays, APC1–APC5. The data is normalized to zero for the sample with the lowest expression for each assay. The error bars indicate the standard deviation for the RT replicates. (**a**) APC1 corresponds to transcript NM_001127511.1 from promoter 1B. (**b**) APC2 corresponds to transcript NM_000038.4 from promoter 1A. (**c**) APC3 corresponds to transcript NM_001127510.1 from promoter 1A. (**d**) APC4 corresponds to all transcripts generated from both of the promoters, 1A and 1B. (**e**) APC5 corresponds to all transcripts generated from both of the promoters, 1A and 1B.
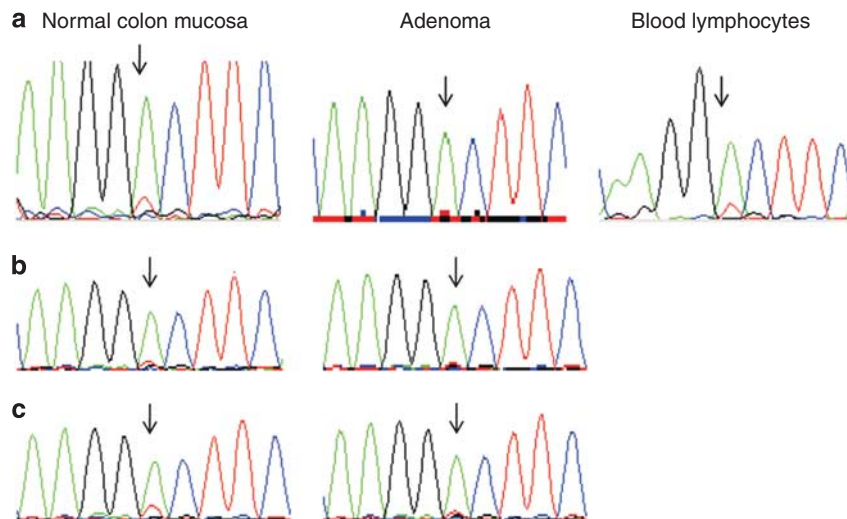

Figure 7 cDNA sequencing over the cSNP c.5465 A/T in affected family members of Family 1. Allele-specific expression was determined by visual estimation. The A allele (normal) is green and the T allele (deleterious) is red. (**a**) Patient P4, normal colorectal mucosa; 100% A, ~8% T, adenoma; 100% A, ~0% T, blood lymphocytes; 100% A, ~9% T. (**b**) P1, normal colorectal mucosa; 100% A, ~0% T, adenoma; 100% A, ~0% T. (**c**) P3, normal colorectal mucosa; 100% A, ~9% T, adenoma; 100% A, ~0% T.

we have made a thorough study of the effect the deletion has on the transcription of *APC* in mutation carriers. We have used two different approaches for this study.

Analyses of relative expression levels of specific transcripts generated from each promoter separately and also from the total expression from both of the

promoters in combination was carried out by Q–PCR (TaqMan). We also used sequencing over a heterozygous cDNA polymorphism (c.5465 A/T) to estimate the overall ASE in mutations carriers. Expression levels were measured in blood, normal colorectal mucosa and adenoma from mutation carriers. In the normal colon mucosa from mutation carriers, a level of $\sim 10\%$ expression of the deleterious allele was determined using cDNA sequencing. No detectable expression of the deleterious allele could be observed in the adenomas indicating allele loss. The reason for the loss of deleterious allele in one of the normal colorectal mucosa samples is probably because of the fact that this sample was taken close to the adenoma. The Q–PCR results indicate markedly lowered levels of expression from promoter 1B in mutation carriers. Unexpectedly in blood, elevated levels of expression from promoter 1A are observed. The reason for this could be tissue-specific compensation of 1A transcription due to the damaged promoter 1B (Fiering et al., 1999). However, this compensation is not seen in samples from normal colon mucosa and adenoma from mutation carriers. As we have described above, estimation of absolute expression levels for promoter 1B in blood from normal controls indicate that 1B is expressed in considerably higher amounts compared with 1A, which is also obvious when relative expression levels from both promoters (APC4 and APC5, Figure 3) are examined, as those levels are more equal to expression from 1B than 1A. In view of these observations, it is straightforward to state that the inactivation of promoter 1B, caused by the deletion we identified in this study, is a cause of FAP in carriers. In Family 1, we have ruled out other possible mechanisms of inactivation of APC, such as mutations in the coding and promoter regions and also partly in the intronic regions. Involvement of mutations affecting mRNA and inactivation caused by promoter methylation has also been excluded as well as chromosomal aberrations.

Apparently no loss of heterozygosity of the normal allele (A allele) is present in the adenomas. If there is a concordance with the study by Renkonen et al. (2005), a complete loss of heterozygosity of the normal allele, not involved in the germline alteration, should be expected in tumors. We were not able to obtain tumors from Family 1 for this study, as colorectal cancer is rare in affected members living today. Aberrant Wnt signaling with nuclear accumulation of β-catenin, following homozygous inactivation of APC, has been the governing model for tumor initiation in FAP, and germline mutations are believed to dictate the second hit in APC in FAP, and thus, indirectly affect the severity of the disease. Two or three out of seven 20-amino-acid repeats in the central domain of APC, which give an optimal level of β-catenin required for tumor formation, seem to be required for tumorigenesis. Consequently, in FAP patients with whole gene deletions, all second somatic hits should be in the mutation cluster region in exon 15 leaving two or three of the repeats intact. However, sequence analyses of the adenomas in this family did not reveal any inactivating mutation at all in this region of the gene or in the entire gene including promoter 1A

(mutations could though be present in tumors). Emerging evidence points to an alternative model for adenoma initiation in which APC mutations induces adenoma formation through dysregulation of C-terminal binding protein-1 (Hamada and Bienz, 2004, Sierra et al., 2006) rather than activation of β-catenin (Phelps et al., 2009a, b). It has been demonstrated that nuclear localization of β-catenin is not present in the early adenomas (Amos-Landgraf et al., 2007, Blaker et al., 2003), and that β-catenin requires mutations of additional genes for nuclear accumulation and activation of the cell-proliferating genes. The fact that we, in adenomas from mutation carriers, cannot detect homozygous inactivation of APC support this model, as the initiation of adenoma formation in the carriers occurs independently of the favored number of 20-amino-acid repeats, and, therefore, is independent of β-catenin levels. Instead additional mutations in other genes might be necessary.

Hypermethylation of CpG islands in the promoters 1A, but not 1B, of APC has been reported as a cause of gene silencing in sporadic tumors and tumors from FAP patients, but no methylation has been detected in normal colon mucosa (Hiltunen et al., 1997; Esteller et al., 2000; Segditsas et al., 2008). We examined the adenomas for methylation status, and as all adenomas had low-grade dysplasia, we would expect very low methylation levels, which was the case in all but one sample. However, this might be a random event, and as this was just one sample from one individual no further conclusion can be made.

The absence of desmoid tumors in the large Family 1 is striking, and except for osteomas and lipomas other extra-intestinal manifestations seem to be rare. Less frequent occurrence of extra-colonic manifestations in FAP patients with reduced expression have also been reported in other studies (Moisio et al., 2002; Renkonen et al., 2005). The fact that mutation carriers in Family 1 retain $\sim 10\%$ expression of the deleterious allele and demonstrate an increased expression from promoter 1A in blood but not in colorectal mucosa could be of importance. It could well be that elevated expression levels from promoter 1A also are present in other tissues in Family 1, which could be an explanation for less frequent manifestations in other organs in the family.

We used the recently introduced Affymetrix Genome-Wide Human SNP 6.0 CNV arrays, with an average spacing of 0.7 kb between probes in this study. This resolution is required to identify CNVs of the sizes we describe. Even if the resolution of this array is adequate to detect CNVs as small as this, it is difficult to discover very small deviations of single or a few probes, which is exemplified by the fact that we could not detect the retained region (1.5 kb) between the two deletions in Family 1 in the array analyses. We suggest that in families displaying a classical FAP phenotype with no deleterious mutation identified, screening for CNVs in the regulatory regions of APC should follow the initial screening of the coding region as mutations located upstream of promoter 1A could be a possible causative alteration.

Materials and methods

Patients and material

We analyzed 54 FAP and AFAP families in this study. The majority of families are registered in the Swedish Polyposis Registry. Family 1 was the first family enrolled in the Registry in the late 1950s. It is by far the largest pedigree in the registry, including seven generations (150 individuals), with 59 verified FAP cases, of whom 24 were probands with a mean age at diagnosis of 38.3 (range, 26–71) years. The clinical manifestations of the individuals are described in Supplementary Table 2. Of the 35 call-up patients, with a mean age of 19 (range, 8–36) years at diagnosis, none had colorectal cancer at surgery. In patients examined with gastroscopy, duodenal adenomas were diagnosed in 93% of the cases for which we had data. One patient had a duodenal cancer and four have been prophylactically operated with duodenectomy. Benign extra-intestinal manifestations were exclusively osteomas and lipomas, only one malignant extra-intestinal manifestation has so far been discovered, and none of the patients have developed desmoid tumors. Adenoma and normal colon mucosa (localized close to the adenoma) were examined from four patients in this family. All of the adenomas were classified as low-grade dysplasia. Blood donors were used as healthy controls. All patients gave their consent, and the study has been approved by the local ethics committee at the Karolinska Institute in Stockholm and at the University of Gothenburg.

All index patients included in this study have been screened for germline mutations in the coding region of *APC* (NM_000038.3) as described by Kanter-Smoler *et al.* (2008). The FAP/AFAP-mutation-negative patients were in some cases screened for the two most frequent mutations in *MUTYH* (NM_001128425.1) or the whole gene, see Table 1. The Human Gene Mutation Database database and The Colon Cancer Gene Variant Database (http://www.LOVD.nl/APC) were used to control if the identified mutations in APC had been reported previously. Family 1 has been extensively screened for mutations including sequencing of promoter regions and also ∼50% of the intron sequences, as well as analyses of promoter methylation (in blood lymphocytes). To detect second-hit mutations in adenomas, the *APC*-coding region and promoter 1A was also sequenced. To rule out any chromosomal aberrations linked to the deleterious genotype, two affected members were also karyotyped.

DNA, RNA, cDNA preparation and karyotyping

Genomic DNA was extracted using the BioRobot EZ1 (Qiagen, Hilden, Germany) with the EZ1 DNA Blood 350 µl kit (Qiagen). DNA was extracted from tissues from adenomas and normal colon mucosa using QiaAmp DNA microkit (Qiagen). Blood samples were collected into PAXgene Blood RNA tubes (Qiagen) and RNA was extracted using PAXgene Blood RNA kit (Qiagen) and from tissue samples using the RNeasy Mini kit (Qiagen). The RNA samples used from normal tissues (except blood) were the Human Total RNA Master Panel II (Clontech Laboratories, Mountain View, CA, USA), which was used to evaluate the aberrant transcript lacking exon 2 and the FirstChoice Human Total RNA Survey Panel (Ambion, Applied Biosystems), which was used in the Q–PCR analyses together with a separate sample from normal colorectal mucosa from the same distributor. cDNA was synthesized using Superscript system for reverse transcriptase–PCR (Invitrogen GmBH, Karlsruhe, Germany). Spreads from metaphase chromosomes were prepared from cultured lymphocytes as described in Johannesson *et al.* (1991).

Amplification, sequence analyses and MLPA

Amplification, purification and sequencing were carried out as described previously (Kanter-Smoler *et al.*, 2008). Primers used for direct sequencing were identical to those used in the amplification reactions. The primer sequences for the exons and promoter regions are given in Supplementary Table 3, and the cDNA primers are described in De Rosa *et al.* (2007). CN detection by MLPA was performed according to the protocol provided by the supplier using the SALSA MLPA kit P043 APC (version B1, MRC-Holland, Amsterdam, the Netherlands). The MLPA data were analyzed using GeneMapper 4.0 genotyping software (Applied Biosystem) and SeqPilot version 3.3.2 (JSI medical systems, GmBH, Kippenheim, Germany).

Expression analyses using Affymetrix Gene Chip Human Exon 1.0 ST arrays

Quantification of RNA was made using Nanodrop ND-1000 UV-spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, IL, USA), and RNA quality was assessed using RNA 6000 Nanochips with Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). RNA (1 µg) from samples with RNA integrity numbers between 7 and 10 were analyzed. Three technical replicates were used for each sample. Ribosomal RNA was removed using RiboMinus Human Transcriptome Isolation (Invitrogen). Samples were prepared according to standard conditions (Affymetrix). Washing and staining of arrays were performed on the GeneChip Fluidics station 450 (Affymetrix).

Data analysis

Analysis of raw data was performed with the expression console software from Affymetrix. A principle component analysis was conducted to identify outliers in the samples. The control spikes and hybridization metrics were evaluated as described by Affymetrix.

Analyses of the data was also carried out using the Partek Genomics Suites 6.4 software (Partek Inc., St Louis, MO, USA). The robust multichip average algorithm was used for probe set intensity analysis. The data was filtered to include only the core probe set. Analysis of variance in the Partek software was used for calculation of *P*-values to identify expression fold changes and splicing events. A decrease in fold change by a factor 2 corresponds to approximately a 50% downregulation in gene expression.

CNV detection using Affymetrix Genome-Wide Human SNP Array 6.0

The GeneChip 6.0 platform includes about 906 600 SNP and about 900 000 CN probes, covering the whole genome with an average spacing of ∼0.7 kb. Samples were prepared according to standard conditions (Affymetrix). Purification of PCR products was performed using Magnetic Beads (Agencourt Bioscience Corporation, Beverly, MA, USA). Hybridization, washing and staining of arrays was performed as described by the supplier (Affymetrix).

Data analysis

Affymetrix Genotyping Console 3.0.1 was used for CN analysis. The Genotyping Console Contrast Quality Control metric was used to filter out low-quality samples using the default threshold of >0.4. For the CN analysis the Median Absolute Pairwise Difference value threshold was set to ⩽0.35. The Affymetrix HapMap 270 Reference Model was used for the analysis. Statistical analysis was carried out with

the default settings in Genotyping Console, including regional GC correction and the Hidden Markov model.

Isolation and verification of the deleted region
Primer walking was used to map the two breakpoints in Family 1, the primers can be obtained from the author upon request. The PCR primers used for the verification of the deletions were forward primer: 5′-CCAGGGAACTCAAT CAAGGA-3′ (chr5:111980705–111980724) and reverse primer: 5′-CTACCGCTTTGATCCCACAT-3′ (chr5:112044642–112044661). These primers amplified a fragment of ∼3 kb in deletion carriers and was absent in non-carriers. The PCR fragment was sequenced with internal primers (can be obtained from the author upon request) to reveal the break points.

Promoter methylation analysis by sequencing and MLPA
The presence of DNA methylation at CpG dinucleotides in the *APC* promoter 1A (accession number U02509) and 1B (accession number D13981) was determined using bisulphite sequencing. The DNA samples were treated with the Epitec bisulphite kit (Qiagen). The primers and conditions were according to Romero-Gimenez *et al.* (2008). Methylation-specific MLPA analysis was performed using the SALSA MLPA kit ME001 Tumor suppressor-1 (MRC-Holland). This kit contains one methylation-specific probe in promoter 1A of the *APC* gene. The probe is located 72 bp upstream of exon 2 (NM_000038.4). The MLPA results were analyzed using GeneMapper 4.0 genotyping software (Applied Biosystems) and SeqPilot version 3.3.2. Methylation dosage ratio was calculated according to Joensuu *et al.* (2008).

Q–PCR performed using TaqMan gene expression assays
Five different TaqMan probes were obtained from Applied Biosystems. The APC 1 probe (Hs01568282_m1) is specific for transcript NM_001127511.1 expressed from promoter 1B, APC2 is specific for NM_000038.4 expressed from promoter 1A, APC3 is specific for NM_001127510.1 also expressed from promoter 1A, whereas APC4 (Hs01568270_m1) and APC5 (Hs00181051_m1) bind to all transcripts generated from both promoters. The APC2 and APC3 probes were designed by us using the Custom TaqMan Expression Assays (Applied Biosystems). Q–PCR was performed in 384-well plates using the Applied Biosystems 7900HT fast real-time PCR system (Applied Biosystems) or the Roche LightCycle 480 (Hoffman-La Roche, Basel, Switzerland). Amplifications were carried out in triplicates or in duplicates for the blood controls (S1–S10), patient samples (P1–P4) and the colorectal mucosa and in singlets for all the other tissues. The standard curve method was used for the Q–PCR analyses. The relative expression levels for each sample was calculated from the mean Cq-values and CN were expressed as a fold difference relative to the control samples. All samples were normalized against glyceraldehyde 3-phosphate dehydrogenase (GAPDH), beta-glucuronidase (GUSB) and TATA-binding protein (TBP). The

t-test was used to calculate statistical significance. It is possible to estimate absolute expression levels in Q–PCR by using the fact that Q–PCR by TaqMan is an exponential reaction in which a 10-cycle difference corresponds to a 1000-fold difference in the amount of transcripts in a correctly designed assay functioning properly in terms of specificity and efficiency. The Cq-values obtained were compared between the assays, and an estimate of absolute levels was made from this raw data of the Q–PCR reactions. The amounts calculated can serve as a guideline regarding the amount of transcripts. The Q–PCR analyses were carried out at the TATAA Biocenter AB and the Genomics Core Facility at the Sahlgrenska Academy, University of Gothenburg (Gothenburg, Sweden).

Bioinformatics
To identify evolutionary conserved regulatory elements in the deleted regions of Family 1, an alignment to 43 species using the University of California, Santa Cruz database and the ENSEMBL (European Molecular Biology Laboratory-European Bioinformatics Institute and the Sanger Institute) database was performed. The March 2006 (National Center for Biotechnology Information/hg18) map was used for the alignments in the University of California, Santa Cruz database. Conserved regions were further analyzed in detail to find regulatory elements including transcription factor binding sites, enhancer elements, expressed sequence tags and open reading frames with open reading frame finder (http://www.ncbi.nlm.nih.gov/projects/gorf/), TRANSFAC (http://www.gene-regulation.com/pub/databases.html) for prediction of transcription binding sites (Matys *et al.*, 2006) and Enhancer element locator (Palin *et al.*, 2006; http://sbz-17.cs.helsinki.fi/∼tkt_bsap/EELweb/). The region was also compared with RIKEN data for CAGE (cap analysis gene expression) transcripts in the region (Kawaji *et al.*, 2009).

Conflict of interest

The authors declare no conflict of interest.

References

Amos-Landgraf JM, Kwong LN, Kendziorski CM, Reichelderfer M, Torrealba J, Weichert J et al. (2007). A target-selected Apc-mutant rat kindred enhances the modeling of familial human colon cancer. *Proc Natl Acad Sci USA* **104**: 4036–4041.

Andreutti-Zaugg C, Couturier A, Chappuis P, Hutter P. (1999). Detection of protein truncating mutations in exons 1–14 of the APC gene using an *in vivo* fusion protein assay. Mutations in brief no. 214. Online. *Hum Mutat* **13**: 170–171.

Bisgaard ML, Fenger K, Bulow S, Niebuhr E, Mohr J. (1994). Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. *Hum Mutat* **3**: 121–125.

Blaker H, Scholten M, Sutter C, Otto HF, Penzel R. (2003). Somatic mutations in familial adenomatous polyps. Nuclear translocation of

beta-catenin requires more than biallelic APC inactivation. *Am J Clin Pathol* **120**: 418–423.

Castellsague E, Gonzalez S, Guino E, Stevens KN, Borras E, Raymond VM *et al.* (2010a). Allele-specific expression of APC in adenomatous polyposis families. *Gastroenterology* **139**: 439–447, 447 e431.

Castellsague E, Gonzalez S, Guino E, Stevens KN, Borras E, Raymond VM *et al.* (2010b). Allele-specific expression of APC in adenomatous polyposis families. *Gastroenterology* **139**: 439–447, 447.e1.

Charames GS, Ramyar L, Mitri A, Berk T, Cheng H, Jung J *et al.* (2008). A large novel deletion in the APC promoter region causes gene silencing and leads to classical familial adenomatous polyposis in a Manitoba Mennonite kindred. *Hum Genet* **124**: 535–541.

Cowie S, Drmanac S, Swanson D, Delgrosso K, Huang S, du Sart D *et al.* (2004). Identification of APC gene mutations in colorectal cancer using universal microarray-based combinatorial sequencing-by-hybridization. *Hum Mutat* **24**: 261–271.

De Rosa M, Morelli G, Cesaro E, Duraturo F, Turano M, Rossi GB *et al.* (2007). Alternative splicing and nonsense-mediated mRNA decay in the regulation of a new adenomatous polyposis coli transcript. *Gene* **395**: 8–14.

Dobbie Z, Spycher M, Mary JL, Haner M, Guldenschuh I, Hurliman R *et al.* (1996). Correlation between the development of extracolonic manifestations in FAP patients and mutations beyond codon 1403 in the APC gene. *J Med Genet* **33**: 274–280.

Esteller M, Sparks A, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA *et al.* (2000). Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. *Cancer Res* **60**: 4366–4371.

Fiering S, Bender MA, Groudine M. (1999). Analysis of mammalian cis-regulatory DNA elements by homologous recombination. *Methods Enzymol* **306**: 42–66.

Fodde R, van der Luijt R, Wijnen J, Tops C, van der Klift H, van Leeuwen-Cornelisse I *et al.* (1992). Eight novel inactivating germ line mutations at the APC gene identified by denaturing gradient gel electrophoresis. *Genomics* **13**: 1162–1168.

Gayther SA, Wells D, SenGupta SB, Chapman P, Neale K, Tsioupra K *et al.* (1994). Regionally clustered APC mutations are associated with a severe phenotype and occur at a high frequency in new mutation cases of adenomatous polyposis coli. *Hum Mol Genet* **3**: 53–56.

Gismondi V, Bafico A, Biticchi R, Pedemonte S, Molina F, Heouaine A *et al.* (1997). Characterization of 19 novel and six recurring APC mutations in Italian adenomatous polyposis patients, using two different mutation detection techniques. *Hum Mutat* **9**: 370–373.

Hamada F, Bienz M. (2004). The APC tumor suppressor binds to C-terminal binding protein to divert nuclear beta-catenin from TCF. *Dev Cell* **7**: 677–685.

Hiltunen MO, Alhonen L, Koistinaho J, Myohanen S, Paakkonen M, Marin S *et al.* (1997). Hypermethylation of the APC (adenomatous polyposis coli) gene promoter region in human colorectal carcinoma. *Int J Cancer* **70**: 644–648.

Horii A, Nakatsuru S, Ichii S, Nagase H, Nakamura Y. (1993). Multiple forms of the APC gene transcripts and their tissue-specific expression. *Hum Mol Genet* **2**: 283–287.

Hosoya K, Yamashita S, Ando T, Nakajima T, Itoh F, Ushijima T. (2009). Adenomatous polyposis coli 1A is likely to be methylated as a passenger in human gastric carcinogenesis. *Cancer Lett* **285**: 182–189.

Joensuu EI, Abdel-Rahman WM, Ollikainen M, Ruosaari S, Knuutila S, Peltomaki P. (2008). Epigenetic signatures of familial cancer are characteristic of tumor type and family category. *Cancer Res* **68**: 4597–4605.

Johannesson T, Holmqvist D, Martinsson T, Wahlstrom J. (1991). An improved technique for chromosome preparations from human lymphocytes. *Hereditas* **115**: 295–297.

Judson H, Stewart A, Leslie A, Pratt NR, Baty DU, Steele RJ *et al.* (2006). Relationship between point gene mutation, chromosomal abnormality, and tumour suppressor gene methylation status in colorectal adenomas. *J Pathol* **210**: 344–350.

Kanter-Smoler G, Fritzell K, Rohlin A, Engwall Y, Hallberg B, Bergman A *et al.* (2008). Clinical characterization and the mutation spectrum in Swedish adenomatous polyposis families. *BMC Med* **6**: 10.

Kaufmann A, Vogt S, Uhlhaas S, Stienen D, Kurth I, Hameister H *et al.* (2009). Analysis of rare APC variants at the mRNA level: six pathogenic mutations and literature review. *J Mol Diagn* **11**: 131–139.

Kawaji H, Severin J, Lizio M, Waterhouse A, Katayama S, Irvine KM *et al.* (2009). The FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation. *Genome Biol* **10**: R40.

Kohoutova M, Stekrova J, Jirasek V, Kapras J. (2002). APC germline mutations identified in Czech patients with familial adenomatous polyposis. *Hum Mutat* **19**: 460–461.

Lambertz S, Ballhausen WG. (1993). Identification of an alternative 5′ untranslated region of the adenomatous polyposis coli gene. *Hum Genet* **90**: 650–652.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A *et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* **34**: D108–D110.

Miyoshi Y, Ando H, Nagase H, Nishisho I, Horii A, Miki Y *et al.* (1992). Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients. *Proc Natl Acad Sci USA* **89**: 4452–4456.

Moisio AL, Jarvinen H, Peltomaki P. (2002). Genetic and clinical characterisation of familial adenomatous polyposis: a population based study. *Gut* **50**: 845–850.

Mutoh M, Akasu T, Takahashi M, Niho N, Yoshida T, Sugimura T *et al.* (2006). Possible involvement of hyperlipidemia in increasing risk of colorectal tumor development in human familial adenomatous polyposis. *Jpn J Clin Oncol* **36**: 166–171.

Nilbert M, Fernebro J, Kristoffersson U. (2000). Novel germline APC mutations in Swedish patients with familial adenomatous polyposis and Gardner syndrome. *Scand J Gastroenterol* **35**: 1200–1203.

Palin K, Taipale J, Ukkonen E. (2006). Locating potential enhancer elements by comparative genomics using the EEL software. *Nat Protoc* **1**: 368–374.

Phelps RA, Broadbent TJ, Stafforini DM, Jones DA. (2009a). New perspectives on APC control of cell fate and proliferation in colorectal cancer. *Cell Cycle* **8**: 2549–2556.

Phelps RA, Chidester S, Dehghanizadeh S, Phelps J, Sandoval IT, Rai K *et al.* (2009b). A two-step model for colon adenoma initiation and progression caused by APC loss. *Cell* **137**: 623–634.

Renkonen ET, Nieminen P, Abdel-Rahman WM, Moisio AL, Jarvela I, Arte S *et al.* (2005). Adenomatous polyposis families that screen APC mutation-negative by conventional methods are genetically heterogeneous. *J Clin Oncol* **23**: 5651–5659.

Ripa R, Bisgaard ML, Bulow S, Nielsen FC. (2002). *De novo* mutations in familial adenomatous polyposis (FAP). *Eur J Hum Genet* **10**: 631–637.

Romero-Gimenez J, Dopeso H, Blanco I, Guerra-Moreno A, Gonzalez S, Vogt S *et al.* (2008). Germline hypermethylation of the APC promoter is not a frequent cause of familial adenomatous polyposis in APC/MUTYH mutation negative families. *Int J Cancer* **122**: 1422–1425.

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. (2002). Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* **30**: e57.

Segditsas S, Sieber OM, Rowan A, Setien F, Neale K, Phillips RK *et al.* (2008). Promoter hypermethylation leads to decreased APC mRNA expression in familial polyposis and sporadic colorectal

tumours, but does not substitute for truncating mutations. *Exp Mol Pathol* **85**: 201–206.

Sierra J, Yoshida T, Joazeiro CA, Jones KA. (2006). The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. *Genes Dev* **20**: 586–600.

Solomon E, Voss R, Hall V, Bodmer WF, Jass JR, Jeffreys AJ *et al*. (1987). Chromosome 5 allele loss in human colorectal carcinomas. *Nature* **328**: 616–619.

Stenson PD, Mort M, Ball EV, Howells K, Phillips AD, Thomas NS *et al*. (2009). The Human Gene Mutation Database: 2008 update. *Genome Med* **1**: 13.

Tsuchiya T, Tamura G, Sato K, Endoh Y, Sakata K, Jin Z *et al*. (2000). Distinct methylation patterns of two APC gene promoters in normal and cancerous gastric epithelia. *Oncogene* **19**: 3642–3646.

Venesio T, Balsamo A, Rondo-Spaudo M, Varesco L, Risio M, Ranzani GN. (2003). APC haploinsufficiency, but not CTNNB1 or CDH1 gene mutations, accounts for a fraction of familial adenomatous polyposis patients without APC truncating mutations. *Lab Invest* **83**: 1859–1866.

Yan H, Dobbie Z, Gruber SB, Markowitz S, Romans K, Giardiello FM *et al*. (2002). Small changes in expression affect predisposition to tumorigenesis. *Nat Genet* **30**: 25–26.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)