

Low-level *APC* mutational mosaicism is the underlying cause in a substantial fraction of unexplained colorectal adenomatous polyposis cases

Isabel Spier,^{1,2} Dmitriy Drichel,³ Martin Kerick,^{4,5} Jutta Kirfel,^{2,6} Sukanya Horpaopan,^{1,7} Andreas Laner,^{8,9} Stefanie Holzapfel,^{1,2} Sophia Peters,¹ Ronja Adam,^{1,2} Bixiao Zhao,¹⁰ Tim Becker,^{3,11} Richard P Lifton,¹⁰ Sven Perner,¹² Per Hoffmann,^{1,13,14} Glen Kristiansen,^{2,6} Bernd Timmermann,¹⁵ Markus M Nöthen,^{1,13} Elke Holinski-Feder,^{8,9} Michal R Schweiger,^{4,5} Stefan Aretz^{1,2}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2015-103468>).

For numbered affiliations see end of article.

Correspondence to

Dr Isabel Spier, Institute of Human Genetics, University of Bonn, Sigmund-Freud-Str. 25, Bonn D-53127, Germany; isabel.spier@uni-bonn.de

Received 1 September 2015
Revised 20 October 2015
Accepted 22 October 2015
Published Online First
27 November 2015

ABSTRACT

Background In 30–50% of patients with colorectal adenomatous polyposis, no germline mutation in the known genes *APC*, causing familial adenomatous polyposis, *MUTYH*, causing *MUTYH*-associated polyposis, or *POLE* or *POLD1*, causing polymerase-proofreading-associated polyposis can be identified, although a hereditary aetiology is likely. This study aimed to explore the impact of *APC* mutational mosaicism in unexplained polyposis.

Methods To comprehensively screen for somatic low-level *APC* mosaicism, high-coverage next-generation sequencing of the *APC* gene was performed using DNA from leucocytes and a total of 53 colorectal tumours from 20 unrelated patients with unexplained sporadic adenomatous polyposis. *APC* mosaicism was assumed if the same loss-of-function *APC* mutation was present in ≥ 2 anatomically separated colorectal adenomas/carcinomas per patient. All mutations were validated using diverse methods.

Results In 25% (5/20) of patients, somatic mosaicism of a pathogenic *APC* mutation was identified as underlying cause of the disease. In 2/5 cases, the mosaic level in leucocyte DNA was slightly below the sensitivity threshold of Sanger sequencing; while in 3/5 cases, the allelic fraction was either very low (0.1–1%) or no mutations were detectable. The majority of mosaic mutations were located outside the somatic mutation cluster region of the gene.

Conclusions The present data indicate a high prevalence of pathogenic mosaic *APC* mutations below the detection thresholds of routine diagnostics in adenomatous polyposis, even if high-coverage sequencing of leucocyte DNA alone is taken into account. This has important implications for both routine work-up and strategies to identify new causative genes in this patient group.

INTRODUCTION

Adenomatous polyposis syndromes of the colorectum are precancerous conditions characterised by the presence of dozens to thousands of adenomatous polyps, which, unless detected early and resected, invariably result in colorectal cancer. The phenotypic spectrum ranges from an early-onset manifestation with high numbers of adenomas and

a positive family history to isolated late-onset disease with low polyp burden.

To date, three inherited monogenic forms can be delineated by molecular genetic analyses: (1) autosomal-dominant familial adenomatous polyposis (FAP), caused by heterozygous germline mutations (<http://www.lovd.nl/APC>; <http://www.umd.be/APC>; <http://www.hgmd.cf.ac.uk>) in the tumour suppressor gene *APC*;¹ (2) autosomal-recessive *MUTYH*-associated polyposis, caused by biallelic germline mutations (<http://www.lovd.nl/MUTYH>; <http://www.umd.be/MUTYH>; <http://www.hgmd.cf.ac.uk>) of the base excision repair gene *MUTYH*;² and (3) autosomal-dominant polymerase-proofreading-associated polyposis (PPAP), caused by specific germline missense mutations in the polymerase genes *POLE* and *POLD1*.³ Mutation detection rates are strongly dependent on the colorectal phenotype. In classical FAP, a pathogenic *APC* germline mutation is identified in up to 90% of index patients,^{4,5} whereas *APC* or *MUTYH* germline mutations are detected in only 20–30% of index cases with a mild disease course (attenuated FAP).^{6–8} A PPAP has been reported in up to 7% of families with unexplained multiple colorectal adenomas and carcinomas.⁹

In up to 50% of polyposis patients, no underlying germline mutation is identified, although a hereditary basis is likely. A small fraction of cases might be explained by deep intronic *APC* mutations,¹⁰ rare missense mutations of the *APC* gene¹¹ or mutations in other known cancer predisposition genes.^{12–15} A recent genome-wide analysis of germline copy number variants in 221 patients with unexplained adenomatous polyposis identified a group of genes that are likely to predispose to colorectal adenoma formation.¹⁶

It is well known that the impact of low-level mosaicism in hereditary (tumour) syndromes with a high de novo mutation rate such as *APC*-related FAP is underestimated.^{17–21} Several casuistic reports of somatic and gonadal *APC* mosaicism have been published,^{22–25} one patient was identified by next-generation sequencing (NGS).²⁶ However, only two studies have addressed the issue of mosaicism in a comprehensive manner.^{27,28} These studies detected low-level mutational *APC*



To cite: Spier I, Drichel D, Kerick M, et al. *J Med Genet* 2016;**53**:172–179.

mosaicism in 10–20% of leucocyte DNA among a selected group of unrelated patients with suspected or confirmed *APC* de novo mutation. Prescreening methods such as denaturing high-performance liquid chromatography (DHPLC) or protein truncation test (PTT) were shown to be more sensitive in terms of uncovering mosaicism than Sanger sequencing. This suggests that a number of mosaic cases are likely to be overlooked during current routine diagnostics.^{22–27}

To evaluate the impact of low-level *APC* mutational mosaicism below the detection threshold of Sanger sequencing (10–15% mutated alleles) including mosaicism not visible in tissues originating from the mesodermal germinal sheet such as leucocytes at all, we have chosen a comprehensive approach that took advantage of the more sensitive NGS method^{29–30} and have performed systematic high-coverage sequencing of the *APC* gene in DNA from leucocytes and multiple colorectal tumours (adenomas or carcinomas) in each of 20 unrelated patients with unexplained sporadic adenomatous polyposis.

MATERIALS AND METHODS

Patients/data collection

All patients in the present study had unexplained colorectal adenomatous polyposis, that is, no germline mutation in the *APC* or *MUTYH* genes was identified by Sanger sequencing of the coding regions, deletion/duplication analysis using multiplex ligation-dependent probe amplification (MPA), and screening for pathogenic deep intronic *APC* mutations.^{10–31–32} Furthermore, neither of the two hotspot mutations in *POLE* and *POLD1* was present.⁹

The inclusion criteria for all patients enrolled in this study were the presence of at least 20 synchronous or 40 metachronous, histologically confirmed colorectal adenomas. All patients were of central European origin according to family name and self-reporting.

For high-throughput sequencing of leucocyte-derived and adenoma-derived DNA, 20 unrelated patients with an inconspicuous family history were chosen for whom at least two colorectal tumour samples (adenomas or carcinomas) were available, which had to be separated anatomically by a distance of at least 10 cm. For screening of *APC* mosaicism in leucocyte-derived NGS data, 80 additional unselected index patients with sporadic disease were used (same inclusion criteria). The included patients were recruited in the period from 1992 to 2011.

DNA extraction

Genomic DNA was extracted from peripheral EDTA-anticoagulated blood samples using the standard salting-out procedure. DNA of colorectal formalin-fixed and paraffin-embedded (FFPE) tumour tissue was extracted from punches after careful histological re-evaluation of representative sections, as described elsewhere.³³ Alternatively, 10- μ m-thick sections were cut from FFPE tissue blocks. After deparaffinisation, tumour tissue was macrodissected from unstained slides. A previously marked H&E-stained slide served as a reference. Extraction of FFPE embedded tissue DNA was carried out using the BioRobot M48 Robotic Workstation and the corresponding MagAttract DNA Mini M48 Kit (Qiagen, Hilden, Germany), in accordance with the manufacturer's protocol. The tissues were lysed overnight with proteinase K, and the DNA was eluted in 150 μ l Tris-HCl (pH 7.6).

High-throughput sequencing and bioinformatics workflow

In 9 of the 20 polyp-screened patients, exome sequencing of leucocyte and adenoma DNA was performed at the Max Planck

Institute for Molecular Genetics, Berlin, Germany. Library preparation and whole-exome target enrichment was performed according to Agilent's SureSelect protocol (Human All Exon 50Mb v2, 2011) and as previously described.^{9–34–35} Multiplexed paired end sequencing was performed on the Illumina HiSeq 2000 platform in accordance with the manufacturer's protocol. Base calling and demultiplexing were performed using Illumina's CASAVA pipeline v1.7. Raw reads were mapped to GRCh37/hg19 using BWA v0.5.8³⁶ and default parameters. Enrichment statistics were calculated using Agilent's SureSelect target regions. Local realignment, quality value recalibration and variant calling were performed using GATK v2.1–8.³⁷ In-house tools and ANNOVAR³⁸ were used to annotate and filter the variants. Picard (<http://broadinstitute.github.io/picard/>) was used to collect metrics. For leucocyte exome sequencing, the mean on-target coverage of mapped reads was 71 \times , and 85% of bases were covered at $\geq 10\times$ (see online supplementary table S1). On average, 22 148 genetic variants per patient were identified in the coding regions or the canonical splice sites. In the remaining 11 patients, sequencing of colorectal tumour DNA samples was performed at first using the TruSight Cancer Panel (Illumina).

Whole-exome sequencing of the leucocyte screening cohort (n=80) was performed at the Yale Center for Genome Analysis, New Haven, USA, via capture using NimbleGen 2.1M human exome array followed by paired-end sequencing on a HiSeq 2000 instrument, as previously described.^{9–34} Targeted bases were covered by a mean of 67 independent reads, with an average of 94% of all bases covered eight or more times (see online supplementary table S1). Reads were aligned to the hg19 human reference genome using ELAND (Illumina). Single-nucleotide variants along with short insertions and deletions were identified using SAMtools software.

For validation of suspected *APC* mosaic mutations, targeted sequencing using the FAP MASTR Kit (Multiplicom) with high coverage (read depths >1000) for leucocyte DNA was performed on a MiSeq platform (Illumina). In addition, the identified variants were validated by Sanger sequencing of the corresponding region using standard protocols (primer sequences available upon request).

NGS data were filtered for truncating *APC* variants (nonsense mutations, frameshift deletions/insertions and mutations at highly conserved splice sites) present in at least 5% of reads using *Cartagenia Bench Lab NGS v4.0* (Leuven, Belgium) or the *SeqPilot* software (JSI Medical Systems). Detailed visual inspection of variants was done in a read browser (Integrative Genomics Viewer).

RESULTS

The majority of the patients presented with an attenuated colorectal phenotype (late-onset disease and/or <100 colorectal adenomas). The mean age at diagnosis was 45 years (range 14–73 years). None of the patients had a conspicuous family history (sporadic or isolated cases). The family history was classified as inconspicuous if there was no evidence of polyposis or any other early-onset FAP-related tumour in the first-degree or second-degree relatives of the proband. The basic clinical features are summarised in online supplementary table S2.

To screen for *APC* mosaicism, a systematic search was conducted for the presence of loss-of-function *APC* mutations in multiple colorectal tumours from individual patients. For this approach, 53 colorectal tumours (51 adenomas, 2 carcinomas) from 20 unrelated patients were analysed. Mutational *APC* mosaicism was assumed if the same pathogenic loss-of-function

mutation was present and validated in more than one tumour (workflow shown in [figure 1](#)).

In 5 of the 20 polyp-screened patients, *APC* mosaicism could be detected ([table 1](#)). In patient F5018, the same nonsense *APC* mutation (c.1660C>T;p.Arg554*) was recognised in leucocyte DNA (11% of the reads=allelic fraction; read depth 286×) and two adenomas (allelic fraction 25% (4/16) and 60% (3/5), respectively) ([figure 2A](#) and [table 1](#)). To confirm the degree of mosaicism in leucocyte-derived DNA, targeted high-throughput sequencing of the *APC* gene with high coverage was performed using a commercial assay (Multiplicom). Again, a level of around 10% mutated reads was found ([figure 2B](#)). Sanger sequencing revealed only a very discrete peak ([figure 2C](#)). In one of the adenomas, a further *APC* nonsense mutation (c.4348C>T;p.Arg1450*) was detected in 33% (4/12) of reads, which indicate a possible somatic mutation of the wildtype allele (“second hit”) (see online supplementary figure S1).

In patient F1243, the same nonsense *APC* mutation (c.3283C>T;p.Gln1095*) was found in both adenomas and an

allelic fraction of 1% in leucocyte DNA ([table 1](#) and online supplementary figure S2). In patient F1676, an identical frameshift *APC* mutation (c.2840_2841delGT;p.Cys947Phefs*15) was identified in both adenomas and with an allelic fraction of 9% in leucocyte DNA. Sanger sequencing revealed only very discrete peaks representing the frameshift ([table 1](#) and online supplementary figure S3). In one of the adenomas, a potential second hit (c.4348C>T;p.Arg1450*) was detected in 18% (6/34) of reads. In patient F5000, an identical frameshift *APC* mutation (c.4127_4128delAT;p.Tyr1376Cysfs*9) was found in all four polyps but not in leucocyte DNA ([table 1](#) and [figure 3](#)). In all patients, the suspected mosaic mutations could be revealed in additional polyp samples by Sanger sequencing ([figure 2D](#) and online supplementary table S3, figure S2C and 3D). Using targeted high-throughput sequencing of the *APC* gene (see above), the levels of mutated reads were confirmed in leucocyte-derived DNA with high coverage in patients F1243 and F1676 ([table 1](#) and online supplementary figure S2B and 3C). In patient F5000, the *APC* mutation, found in seven adenomas, was not detectable in leucocyte-derived DNA in spite of a coverage of 2650×.

Furthermore, a nonsense *APC* mutation (c.1495C>T;p.Arg499*), detected in one adenoma by targeted NGS, could be confirmed in two of three additional polyps by Sanger sequencing (patient F1543, [table 1](#) and online supplementary figure S4). In leucocyte DNA, the mutation was only detectable at a very low-level with high-coverage NGS (allelic fraction 0.1%, 22/17 460 reads, [table 1](#)).

To explore the sensitivity of high-coverage leucocyte-derived NGS alone to uncover low-level *APC* mosaicism, we performed a systematic screen of exome sequencing data in 80 additional patients ([figure 1](#)). In 12 patients, *APC* mosaicism was suspected due to the detection of a truncating *APC* mutation in at least 5% of reads. From 10 of those patients, we were able to receive at least two colorectal tumours for validation by Sanger sequencing (29 adenomas, 1 carcinoma). In two of these 10 patients, the mutation (detected in 8–9% of reads in leucocyte DNA) could be confirmed in 2/3 and 4/4 polyps, respectively (F1252 and F1727, [table 1](#), online supplementary figures S5 and 6). On leucocyte level, the mutations could be confirmed by targeted sequencing with high coverage.

The positions of all seven mosaic mutations in relation to the described somatic mutations and the mutation cluster region (MCR) of the *APC* gene are shown in online supplementary figure S7.

All patients with *APC* mosaicism had an attenuated polyposis phenotype; clinical details are provided in [table 2](#). At the time of the last contact (telephone interview), only one child had a colonoscopy with normal findings at 38 years of age (the age of all 12 children of the mosaic index cases ranges between 6 and 40 years, median 26 years), no symptoms such as gastrointestinal bleedings were reported.

DISCUSSION

In a number of patients with colorectal adenomatous polyposis, no germline mutation in the known causal genes can be identified. Based on previous data from our group and others, it can be hypothesised that a substantial fraction of mosaic cases are overlooked in routine diagnostics using Sanger sequencing of leucocyte DNA.^{22 27} To elucidate the frequency of undiscovered *APC* mosaicism in sporadic polyposis patients, we systematically screened multiple adenomas from a number of patients with unexplained polyposis using NGS.

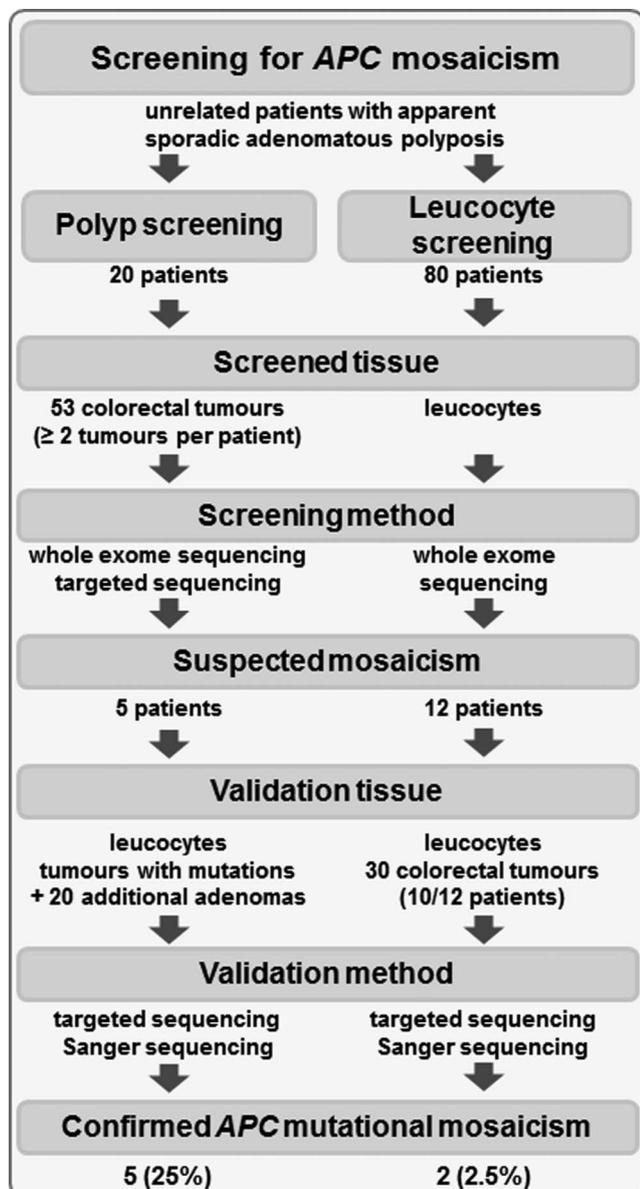


Figure 1 Workflow illustrating the overall procedure of screening for somatic *APC* mutational mosaicism.

Table 1 Detailed histological findings and results of *APC* mutation screening in blood and polyp samples of the seven cases with mosaic *APC* mutations

Patient ID	Sample ID	Histology	Grade of dysplasia	Localisation of polyps	Truncating <i>APC</i> mutation	Sequencing method (number of mutant reads/read depth=% mutant reads)
<i>(A) Mosaic cases in polyp screening cohort (n=20)</i>						
F1243	L				c.3283C>T; p.Gln1095*	Exome (2/213=1%); Multiplicom (33/1875=1.8%) Exome (4/35=11%)
	T1	Tubulovillous adenoma	Low grade	Approximately 50 cm ab ano		
	T2	Tubulovillous adenoma, partly hyperplastic polyp	Low grade	Approximately 40 cm ab ano		
F1543	L				c.1495C>T;p.Arg499*	Multiplicom (22/17 460=0.1%) Sanger Sanger TruSight (65/364=18%)
	T1	Tubular adenoma	Low grade	Rectum		
	T2	Tubular adenoma	Low grade	Caecum		
F1676	L				c.2840_2841delGT; p.Cys947Phefs*15	Exome (19/220=9%); Multiplicom (129/2149=6%) Exome (16/55=29%) Exome (45/130=35%) Exome (6/34=18%)
	T1	Tubular adenoma	Low grade	Transverse colon		
	T2	Tubular adenoma	Low grade	Caecum		
F5000	L				c.4348C>T;p.Arg1450* (second hit?) mutation not detectable	Exome (0/67=0%); Multiplicom (0/2650=0%) Exome (2/3=67%) Exome (23/25=92%)
	T1	Tubular adenoma	Low grade	Descending colon		
	T2	Tubular adenoma	Low grade	13 cm distal from valvula Bauhini		
	T3	Tubular adenoma	Low grade	33 cm distal from valvula Bauhini		
	T4	Tubular adenoma	Low grade	33 cm distal from valvula Bauhini		
F5018	L				c.1660C>T;p.Arg554*	Exome (32/286=11%); Multiplicom (97/1061=9%) Exome (4/16=25%) Exome (3/5=60%) Exome (4/12=33%)
	T1	Tubular adenoma	Low grade	Sigmoid colon		
	T2	Tubular adenoma	Low grade	Rectum		
	T2					
<i>(B) Mosaic cases in leucocyte screening cohort (n=80)</i>						
F1252	L				c.896_897delCT; p.Ser299Leufs*27	Exome (7/77=9%); Multiplicom (344/3643=9%) Sanger Sanger
	T1	Tubulovillous adenoma	Low grade	Sigmoid colon		
	T2	Tubular adenoma	Low grade	Approximately 40 cm ab ano		
F1727	L				c.2240C>G;p.Ser747*	Exome (2/26=8%); Multiplicom (113/2083=5%) Sanger
	T1	Tubulovillous adenoma	Low grade	Approximately 23 cm ab ano		
	T2	Tubulovillous adenoma	Low grade	Approximately 45 cm ab ano		
	T3	Tubulovillous adenoma	Low grade	Approximately 39 cm ab ano		
T4	Tubulovillous adenoma	Low grade	Approximately 32 cm ab ano		Sanger	

L, leucocyte DNA; T, adenoma DNA.

In the previous studies performed by our group²⁷ and Hes *et al.*,²⁸ *APC* mutational mosaicism was detected in 11% and 20% in a selected group of unrelated patients with suspected or confirmed de novo *APC* mutation, respectively. In both studies, a combination of sensitive methods, that is, PTT, DHPLC, denaturing gradient gel electrophoresis (DGGE) and Sanger sequencing was applied.

However, related to the number of *APC* and *MUTYH* mutation-negative index patients included in these study cohorts (450 in our previous study and 295 in the study of Hes *et al.*), the prevalence of uncovered mosaicism is much less. In our current, well-characterised overall cohort of 261 unrelated patients with unexplained adenomatous polyposis,

approximately 80% of patients presented with apparent sporadic disease. When applying this percentage to the mutation negative cohorts of the above-mentioned studies, the frequency of somatic *APC* mosaicism in sporadic unexplained adenomatous polyposis can be estimated to be around 2–4%. These numbers are comparable with our present finding in 80 unselected, sporadic polyposis patients: based on a systematic screening in NGS data of leucocyte-derived DNA and subsequent sequencing of the respective *APC* exons in adenoma DNA, low-level *APC* mosaicism was detected in two cases (2.5%).

In contrast, the multiple adenoma approach identified somatic mosaicism for a pathogenic *APC* mutation in 25% (5/20) of patients with sporadic disease in whom no *APC* and *MUTYH*

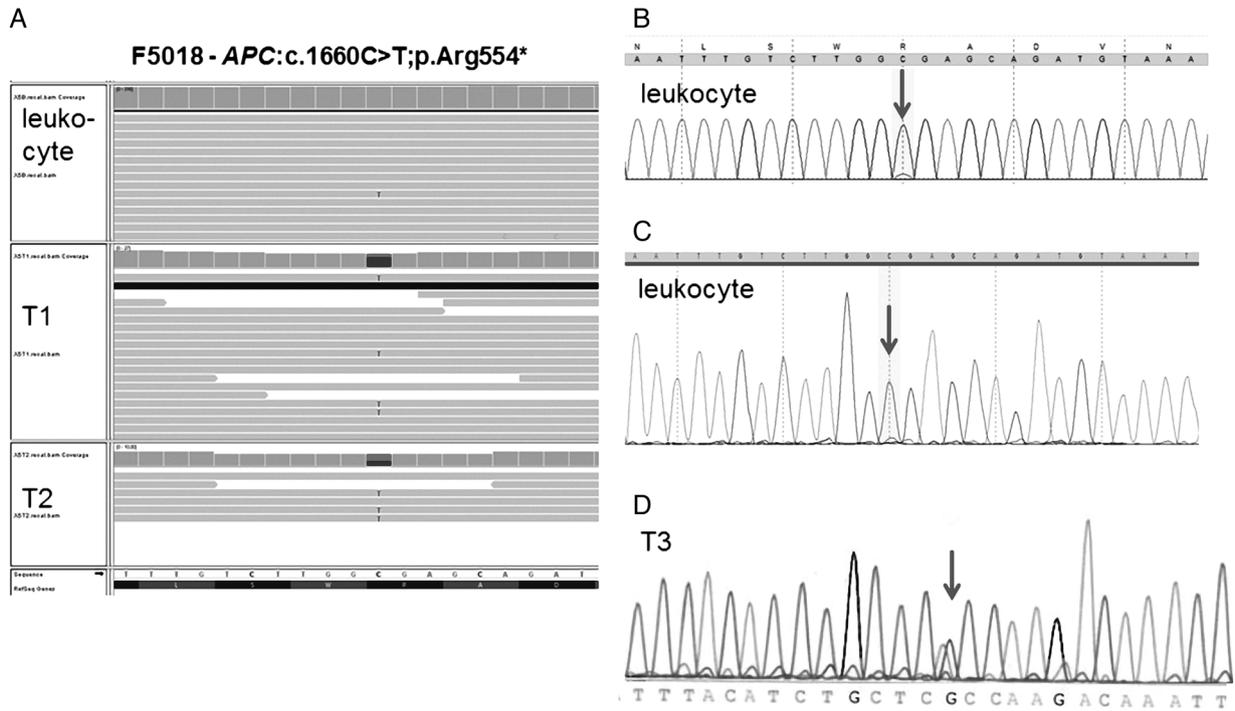


Figure 2 *APC* mutation c.1660C>T;p.Arg554* in mosaic state (patient F5018): (A) Exome sequencing data of leukocyte DNA (11%; 32 mutant reads, coverage 286×) and of adenoma DNA (T1: 25%; 4 mutant reads, coverage 16×, and T2: 60%; 3 mutant reads, coverage 5×) (Integrative Genomics Viewer) (B) Targeted next-generation sequencing data of leukocyte DNA (9%; 97 mutant reads, coverage 1061×); *FAP MASTR Kit* (Multiplicom). The mutation is shown in the ‘peak view’ of the analysis software *SeqPilot* (JSI Medical Systems), where the height of the peaks corresponds to the number of reads. (C) Sanger sequencing of the corresponding region (leukocyte DNA). (D) Sanger sequencing of the corresponding region (additional adenoma sample, reverse sequence).

mutations had been identified in routine diagnostics. According to this result, pathogenic somatic *APC* mosaicism is considerably higher than previous data might have been suggested.

All mosaic mutations were validated by two different methods excluding allele drop-out and technical artefacts. Each tumour-derived DNA sample originated from a single

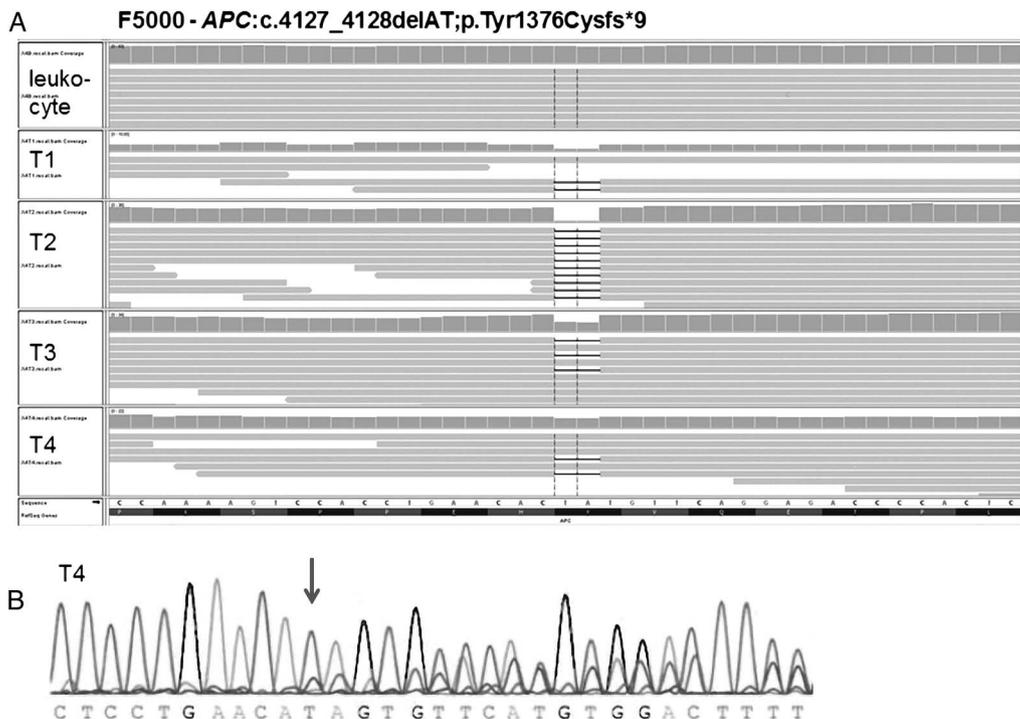


Figure 3 *APC* mutation c.4127_4128delAT;p.Tyr1376Cysfs*9 in mosaic state (patient F5000): (A) Exome sequencing data of adenoma DNA (T1: 67%; 2 mutant reads, coverage 3×, T2: 92%; 23 mutant reads, coverage 25×, T3: 43%; 12 mutant reads, coverage 28× and T4: 17%; 2 mutant reads, coverage 12×) (Integrative Genomics Viewer). (B) Sanger sequencing of the corresponding region (adenoma DNA of T4, reverse sequence).

Table 2 Phenotype of patients with *APC* mosaicism (all had an attenuated polyposis and appeared to be sporadic cases)

Patient ID	Age at diagnosis (years)—adenomas	No. of colorectal adenomas	Distribution of adenomas	Colorectal cancer (position—age at diagnosis)	Colectomy (age)	Duodenal phenotype	Extracolonic phenotype/tumours
<i>(A) Patients of polyp screening cohort</i>							
F1243	48	51–100	Whole colon, distal pronounced	Yes (sigmoid colon, 48 years)	Sigmoid resection (48 years)	Spigelman stage II	Unremarkable
F1543	60	21–50	Whole colon	No	No	Normal	Unremarkable
F1676	55	51–100	Whole colon	No	No	Normal	Gastric fundic gland polyps, gastric foveolar adenoma
F5000	49	100–500	Whole colon	Yes (rectum, 48 years)	Proctocolectomy (48 years)	Normal	Unremarkable
F5018	61	100–500	Whole colon	No	Proctocolectomy (62 years)	Spigelman stage III	Osteoma?
<i>(B) Patients of leucocyte screening cohort</i>							
F1252	40	51–100	Distal pronounced	No	No	Normal	Unremarkable
F1727	30	51–100	Whole colon	No	No	Normal	Unremarkable

independent adenoma. At least two adenomas with the same *APC* mutation were separated anatomically by a distance of >10 cm, the majority even grew in different colonic segments (table 1). None of the mosaic patients reported a conspicuous family history regarding siblings, parents or second-degree ancestors, which is consistent with the assumption of a de novo event.

Interestingly, six of the seven mosaic mutations (five from polyp-screened and two from leucocyte-screened patients) were located outside the somatic MCR (codons 1286–1513) of the *APC* gene.³⁹ In contrast, the two presumed second hits, both of which were identified in a single adenoma (F1676 and F5018), affect the same codon (c.4348C>T;p.Arg1450*), which represents a mutational hotspot within the MCR (see online supplementary figure S7).

In 2/5 cases, the mutation level in leucocyte-derived DNA was slightly below the detection threshold of Sanger sequencing (10–15% mutated alleles), which explains why the mutation was not found during routine diagnostics. Apart from NGS, these cases might have been identified with other sensitive methods such as PTT, DHPLC, DGGE, high-resolution melting or pyrosequencing.^{27–28, 40} However, the use of these techniques is decreasing in clinical practice and even in those cases, additional non-mesoderm tissue is usually required to confirm mosaicism.

In 3/5 cases, the mutation level (allelic fraction) in leucocyte-derived DNA was very low (0.1–1%) or the mutation was not detectable at all despite deep sequencing. It can be presumed that in cases like patient F5000 an early postzygotic mutation must have occurred after mesoderm and endoderm specification.¹⁷ In all three patients, the mutation was present in at least three colorectal adenomas, a finding that is strongly suggestive of mosaicism. These mosaic mutations are undetectable even in high-coverage NGS data from leucocyte-derived DNA alone since screening for mosaic mutations below a level of 5% mutated reads reveals too much false positive results and is thus not feasible in clinical practice. This is also demonstrated in our screening cohort of 80 patients, where 8/10 potential low-level mosaic mutations could not be confirmed by examination of tumour-derived DNA. As a consequence, such mutations, which represent a magnitude of around 60% of mosaic cases in our cohort of 20 patients, can only be uncovered by screening of additional tissues and usually adenomas are easiest

to get and the most promising tissue to identify mosaic mutations.

Hence, our data confirm theoretical considerations of a high frequency of somatic mosaicism not detectable in leucocyte DNA. Although a small number of such cases have been reported in recent years by Hes *et al*,^{28, 40} to date, no systematic investigation of the clinical relevance of *APC* mosaicism has been conducted. Given the number of cases in this study, the results are likely to be representative. The inclusion of multiple adenomas into the routine workup is a promising way to identify those cases and to increase the diagnostic sensitivity considerably.

While somatic mosaicism can contribute to deviations from the predicted phenotype, the level of mosaicism in leucocytes shows no consistent correlation with disease severity. In the present seven patients with somatic mosaicism, the polyps were not restricted to a particular colonic segment; in fact, in two cases even the upper gastrointestinal tract was affected. Nonetheless, all of our mosaic cases presented with an attenuated colorectal disease according to age at onset and/or polyp burden, despite the fact that the position of the mutation in the *APC* gene would have been expected to result in classical (typical) FAP. As a consequence, the offspring of mosaic patients may develop a more severe phenotype than the affected parent. This is relevant for genetic counselling and the decision as to when surveillance should commence.

The diagnosis of mosaicism is also important in terms of estimating the recurrence risk in relatives. While siblings and parents will not be affected, the risk for children is up to 50%, depending on the distribution of the mutation in different tissues. Patients in whom mosaicism is restricted to the endoderm are thought to have a low risk of transmitting the mutation since the primordial germ cells should not be affected.¹⁷ A yet unknown polyposis in children of our index cases cannot be excluded for sure, although there is no evidence for this so far.

To the best of our knowledge, this is the first study that systematically evaluates the impact of *APC* mosaicism using multiple tumour samples in a sufficient number of patients with unexplained polyposis. In conclusion, this work suggests that low-level *APC* mosaic mutations below the threshold of routine diagnostics contribute significantly to the aetiology of adenomatous polyposis. While the frequency of detectable somatic *APC*

mosaicism in leucocyte-derived DNA alone may not exceed 3–4%, even when more sensitive methods are applied, a multiple adenoma approach can increase the diagnostic yield to 20–30%. Our findings may have important implications for both routine diagnostics and research strategies. The inclusion of *APC* mutation screening in multiple (≥ 2) colorectal adenomas can clarify the diagnosis in up to one-quarter of mutation-negative patients, and thus should be considered in all cases with proven (> 20 adenomas) and sporadic disease (inconspicuous family history) in whom no pathogenic *APC*, *MUTYH*, *POLE* or *POLD1* mutation is detected. It also might be reasonable to carefully exclude *APC* mutational mosaicism prior to admittance of this patient group in studies that aim to identify new causative highly penetrant genes.

DATABASES/URLS

APC locus-specific mutation databases: <http://www.lovd.nl/APC>; <http://www.umd.be/APC>; <http://www.hgmd.cf.ac.uk>
IGV (Integrative Genomics Viewer): <http://www.broadinstitute.org/igv/>
MUTYH locus-specific mutation databases: <http://www.lovd.nl/MUTYH>; <http://www.umd.be/MUTYH>; <http://www.hgmd.cf.ac.uk>
Picard: <http://broadinstitute.github.io/picard/>
Primer3 v.0.4.0: <http://frodo.wi.mit.edu/primer3/input.htm>

Author affiliations

- ¹Institute of Human Genetics, University of Bonn, Bonn, Germany
- ²Center for Hereditary Tumor Syndromes, University of Bonn, Bonn, Germany
- ³German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
- ⁴Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany
- ⁵Cologne Center for Genomics (CCG), University of Cologne, Cologne, Germany
- ⁶Institute of Pathology, University of Bonn, Bonn, Germany
- ⁷Department of Anatomy, Faculty of Medical Science, Naresuan University, Phitsanulok, Thailand
- ⁸Medizinische Klinik—Campus Innenstadt, Klinikum der LMU, Munich, Germany
- ⁹MGZ—Center of Medical Genetics, Munich, Germany
- ¹⁰Departments of Genetics, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut, USA
- ¹¹Institute of Medical Biometry, Informatics, and Epidemiology, University of Bonn, Bonn, Germany
- ¹²Section for Prostate Cancer Research, Institute of Pathology, Center for Integrated Oncology Cologne/Bonn, University Hospital of Bonn, Bonn, Germany
- ¹³Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany
- ¹⁴Division of Medical Genetics, University Hospital Basel and Department of Biomedicine, University of Basel, Basel, Switzerland
- ¹⁵Next Generation Sequencing Group, Max Planck Institute for Molecular Genetics, Berlin, Germany

Acknowledgements The authors thank the patients and their families for participating in the study. They are grateful to Siegfried Uhlhaas and Dietlinde Stienen for their excellent technical support.

Contributors IS, DD and MK contributed equally. Study design: SA, IS and MRS. NGS experiments: DD, TB, MK, BT, MRS, BZ and RPL. NGS analysis: DD, MK, IS, SHor and SPet. Sanger sequencing: IS. Bioinformatics input: DD, TB, MK and BZ. Patient recruitment, collection of clinical and familial data: IS, SHol, RA, AL and EH-F. Analysis of clinical data: IS, SHol and RA. DNA extraction out of FFPE tumour tissue: JK, SPer, GK and MRS. Literature search and figures: IS and SA. Preparing manuscript: IS and SA. Critical review of the paper: PH, SHol, SHor, RA, SPet, EH-F, DD, MK, MRS and MMN. Supervision and expert advice: SA, PH, MRS and MMN. IS had full access to all study data and had final responsibility for the decision to submit the manuscript for publication.

Funding This work was supported by the German Cancer Aid (Deutsche Krebshilfe e.V. Bonn, Grant number 108421); the Gerok-Stipendium of the University Hospital Bonn (grant no. O-149.0098); the NIH Centers for Mendelian Genomics (5U54HG006504); the Federal Ministry of Education and Research (0316190A); and the Volkswagenstiftung (Lichtenberg Program to MRS).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Medical Faculty of the University of Bonn, ethics review board no. 224/07.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. *Am J Gastroenterol* 2006;101:385–98.
- 2 Mazzei F, Viel A, Bignami M. Role of *MUTYH* in human cancer. *Mutat Res* 2013;743–744:33–43.
- 3 Palles C, Cazier J-B, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Almeida EG, Salguero I, Sherborne A, Chubb D, Carvajal-Carmona LG, Ma Y, Kaur K, Dobbins S, Barclay E, Gorman M, Martin L, Kovac MB, Humphray S, Thomas HJW, Maher E, Evans G, Lucassen A, Cummings C, Stevens M, Walker L, Halliday D, Armstrong R, Paterson J, Hodgson S, Homfray T, Side L, Izatt L, Donaldson A, Tomkins S, Morrison P, Goodman S, Brewer C, Henderson A, Davidson R, Murday V, Cook J, Haites N, Bishop T, Sheridan E, Green A, Marks C, Carpenter S, Broughton M, Greenhalge L, Suri M, Donnelly P, Bell J, Bentley D, Mcvean G, Ratcliffe P, Taylor J, Wilkie A, Donnelly P, Broxholme J, Buck D, Cazier J-B, Cornall R, Gregory L, Knight J, Lunter G, Mcvean G, Taylor J, Tomlinson I, Wilkie A, Buck D, Gregory L, Humphray S, Kingsbury Z, Mcvean G, Donnelly P, Cazier J-B, Broxholme J, Grocock R, Hassen A, Hollings CC, Hughes L, Humburg P, Kanapin A, Lunter G, Murray L, Rimmer A, Lucassen A, Holmes CC, Bentley D, Donnelly P, Taylor J, Petridis C, Roylance R, Sawyer EJ, Kerr DJ, Clark S, Grimes J, Kearsley SE, Thomas HJW, Mcvean G, Houlston RS, Tomlinson I. Germline mutations affecting the proofreading domains of *POLE* and *POLD1* predispose to colorectal adenomas and carcinomas. *Nat Genet* 2013;45:136–44.
- 4 Friedl W, Aretz S. Familial adenomatous polyposis: experience from a study of 1164 unrelated german polyposis patients. *Hered Cancer Clin Pract* 2005;3:95–114.
- 5 Lagarde A, Rouleau E, Ferrari A, Noguchi T, Qiu J, Briaux A, Bourdon V, Remy V, Gaildrat P, Adelaide J, Birnbaum D, Lidereau R, Sobol H, Olschwang S. Germline *APC* mutation spectrum derived from 863 genomic variations identified through a 15-year medical genetics service to French patients with FAP. *J Med Genet* 2010;47:721–2.
- 6 Aretz S, Uhlhaas S, Goergens H, Siberg K, Vogel M, Pagenstecher C, Mangold E, Caspari R, Propping P, Friedl W. *MUTYH*-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. *Int J Cancer* 2006;119:807–14.
- 7 Grover S, Kastrinos F, Steyerberg EW, Cook EF, Dewanwala A, Burbidge LA, Wenstrup RJ, Syngal S. Prevalence and phenotypes of *APC* and *MUTYH* mutations in patients with multiple colorectal adenomas. *JAMA* 2012;308:485–92.
- 8 Nielsen M, Hes FJ, Nagengast FM, Weiss MM, Mathus-Vliegen EM, Morreau H, Breuning MH, Wijnen JT, Tops CMJ, Vasen HFA. Germline mutations in *APC* and *MUTYH* are responsible for the majority of families with attenuated familial adenomatous polyposis. *Clin Genet* 2007;71:427–33.
- 9 Spier I, Holzapfel S, Altmüller J, Zhao B, Horpaopan S, Vogt S, Chen S, Morak M, Raeder S, Kayser K, Stienen D, Adam R, Nürnberg P, Plotz G, Holinski-Feder E, Lifton RP, Thiele H, Hoffmann P, Steinke V, Aretz S. Frequency and phenotypic spectrum of germline mutations in *POLE* and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int J Cancer* 2015;137:320–31.
- 10 Spier I, Horpaopan S, Vogt S, Uhlhaas S, Morak M, Stienen D, Draaken M, Ludwig M, Holinski-Feder E, Nöthen MM, Hoffmann P, Aretz S. Deep intronic *APC* mutations explain a substantial proportion of patients with familial or early-onset adenomatous polyposis. *Hum Mutat* 2012;33:1045–50.
- 11 Azzopardi D, Dallosso AR, Eliason K, Hendrickson BC, Jones N, Rawston E, Colley J, Moskvina V, Frye C, Sampson JR, Wenstrup R, Scholl T, Cheadle JP. Multiple rare nonsynonymous variants in the adenomatous polyposis coli gene predispose to colorectal adenomas. *Cancer Res* 2008;68:358–63.
- 12 Lefevre JH, Bonilla C, Colas C, Winney B, Johnstone E, Tonks S, Day T, Hutnik K, Boumertit A, Soubrier F, Midgley R, Kerr D, Parc Y, Bodmer WF. Role of rare variants in undetermined multiple adenomatous polyposis and early-onset colorectal cancer. *J Hum Genet* 2012;57:709–16.
- 13 Will O, Carvajal-Carmona LG, Gorman P, Howarth KM, Jones AM, Polanco-Echeverry GM, Chinalong J-A, Günther T, Silver A, Clark SK, Tomlinson I. Homozygous *PMS2* deletion causes a severe colorectal cancer and multiple adenoma phenotype without extraintestinal cancer. *Gastroenterology* 2007;132:527–30.
- 14 Rio Frio T, Lavoie J, Hamel N, Geyer FC, Kushner YB, Novak DJ, Wark L, Capelli C, Reis-Filho JS, Mai S, Pastinen T, Tischkowitz MD, Marcus VA, Foulkes WD. Homozygous *BUB1B* mutation and susceptibility to gastrointestinal neoplasia. *N Engl J Med* 2010;363:2628–37.
- 15 de Voer RM, Geurts van Kessel A, Weren RDA, Ligtenberg MJL, Smeets D, Fu L, Vreede L, Kamping EJ, Verwiel ET, Hahn MM, Ariaans M, Spruijt L, Van Essen T, Houge G, Schackert HK, Sheng JQ, Venselaar H, van Ravenswaaij-Arts CM, Van Krieken JH, Hoogerbrugge N, Kuiper RP. Germline mutations in the spindle assembly checkpoint genes *BUB1* and *BUB3* are risk factors for colorectal cancer. *Gastroenterology* 2013;145:544–7.

- 16 Horpaopan S, Spier I, Zink AM, Altmüller J, Holzapfel S, Laner A, Vogt S, Uhlhaas S, Heilmann S, Stienen D, Pasternack SM, Keppler K, Adam R, Kayser K, Moebus S, Draaken M, Deegenhardt F, Engels H, Hofmann A, Nöthen MM, Steinke V, Perez-Bouza A, Herms S, Holinski-Feder E, Fröhlich H, Thiele H, Hoffmann P, Aretz S. Genome-wide CNV analysis in 221 unrelated patients and targeted high-throughput sequencing reveal novel causative candidate genes for colorectal adenomatous polyposis. *Int J Cancer* 2015;136:E578–89.
- 17 Tuohy TM, Burt RW. Somatic mosaicism: a cause for unexplained cases of FAP? *Gut* 2008;57:10–12.
- 18 Macosko EZ, McCarroll SA. Exploring the variation within. *Nat Genet* 2012;44:614–16.
- 19 Moyhuddin A, Baser ME, Watson C, Purcell S, Ramsden RT, Heiberg A, Wallace AJ, Evans DG. Somatic mosaicism in neurofibromatosis 2: prevalence and risk of disease transmission to offspring. *J Med Genet* 2003;40:459–63.
- 20 Gammon A, Jasperson K, Pilarski R, Prior TW, Kuwada S. PTEN mosaicism with features of Cowden syndrome. *Clin Genet* 2013;84:593–5.
- 21 Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, Nagamani SC, Erez A, Bartnik M, Wiśniowiecka-Kowalik B, Plunkett KS, Pursley AN, Kang SH, Bi W, Lalani SR, Bacino CA, Vast M, Marks K, Patton M, Olofsson P, Patel A, Veltman JA, Cheung SW, Shaw CA, Vissers LE, Vermeesch JR, Lupski JR, Stankiewicz P. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am J Hum Genet* 2014;95:173–82.
- 22 Necker J, Kovac M, Attenhofer M, Reichlin B, Heinemann K. Detection of APC germ line mosaicism in patients with de novo familial adenomatous polyposis: a plea for the protein truncation test. *J Med Genet* 2011;48:526–9.
- 23 Baert-Desurmont S, Piton N, Bou J, Tinat J, Guimbaud R, Selves J, Frébourg T. A remarkable APC mosaicism with two mutant alleles in a family with familial adenomatous polyposis. *Am J Med Genet A* 2011;155A:1500–2.
- 24 Filipe B, Albuquerque C, Bik E, Lage P, Rodrigues P, Vossen R, Tops C, Nobre Leitão C. APC somatic mosaicism in a patient with Gardner syndrome carrying the E1573X mutation: report of a case. *Dis Colon Rectum* 2009;52:1516–20; discussion 1520–1.
- 25 Schwab AL, Tuohy TM, Condie M, Neklason DW, Burt RW. Gonadal mosaicism and familial adenomatous polyposis. *Fam Cancer* 2008;7:173–7.
- 26 Yamaguchi K, Komura M, Yamaguchi R, Imoto S, Shimizu E, Kasuya S, Shibuya T, Hatakeyama S, Takahashi N, Ikenoue T, Hata K, Tsurita G, Shinozaki M, Suzuki Y, Sugano S, Miyano S, Furukawa Y. Detection of APC mosaicism by next-generation sequencing in an FAP patient. *J Hum Genet* 2015;60:227–31.
- 27 Aretz S, Stienen D, Friedrichs N, Stemmler S, Uhlhaas S, Rahner N, Propping P, Friedl W. Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). *Hum Mutat* 2007;28:985–92.
- 28 Hes FJ, Nielsen M, Bik EC, Konvalinka D, Wijnen JT, Bakker E, Vasen HF, Breuning MH, Tops CM. Somatic APC mosaicism: an underestimated cause of polyposis coli. *Gut* 2008;57:71–6.
- 29 Freed D, Stevens EL, Pevsner J. Somatic mosaicism in the human genome. *Genes (Basel)* 2014;5:1064–94.
- 30 Rohlin A, Wernersson J, Engwall Y, Wiklund L, Björk J, Nordling M. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum Mutat* 2009;30:1012–20.
- 31 Aretz S, Stienen D, Uhlhaas S, Pagenstecher C, Mangold E, Caspari R, Propping P, Friedl W. Large submicroscopic genomic APC deletions are a common cause of typical familial adenomatous polyposis. *J Med Genet* 2005;42:185–92.
- 32 Michils G, Tejpar S, Thoelen R, Cutsem E, Vermeesch JR, Fryns J-P, Legius E, Matthijs G. Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): a Belgian study. *Hum Mutat* 2005;25:125–34.
- 33 Schweiger MR, Kerick M, Timmermann B, Albrecht MW, Borodina T, Parkhomchuk D, Zatloukal K, Lehrach H. Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. *PLoS ONE* 2009;4:e5548.
- 34 Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 2009;106:19096–101.
- 35 Kerick M, Isau M, Timmermann B, Sülthmann H, Herwig R, Krobitch S, Schaefer G, Verdorfer I, Bartsch G, Klocker H, Lehrach H, Schweiger MR. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics* 2011;4:68.
- 36 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–60.
- 37 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
- 38 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
- 39 Kohler EM, Derungs A, Daum G, Behrens J, Schneikert J. Functional definition of the mutation cluster region of adenomatous polyposis coli in colorectal tumours. *Hum Mol Genet* 2008;17:1978–87.
- 40 Out AA, van Minderhout IJHM, van der Stoep N, Van Bommel LSR, Kluijff I, Aalfs C, Voorendt M, Vossen RHAM, Nielsen M, Vasen HFA, Morreau H, Devilee P, Tops CMJ, Hes FJ. 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.