IJC
International Journal of Cancer

Frequency and phenotypic spectrum of germline mutations in *POLE* and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas

Isabel Spier[1], Stefanie Holzapfel[1], Janine Altmüller[2,3], Bixiao Zhao[4], Sukanya Horpaopan[1], Stefanie Vogt[1,5], Sophia Chen[4], Monika Morak[6,7], Susanne Raeder[1], Katrin Kayser[1], Dietlinde Stienen[1], Ronja Adam[1], Peter Nürnberg[2], Guido Plotz[8], Elke Holinski-Feder[6,7], Richard P. Lifton[4], Holger Thiele[2], Per Hoffmann[1,9,10], Verena Steinke[1] and Stefan Aretz[1]

[1] Institute of Human Genetics, University of Bonn, Bonn, Germany
[2] Cologne Center for Genomics, University of Cologne, Cologne, Germany
[3] Institute of Human Genetics, University of Cologne, Cologne, Germany
[4] Department of Genetics, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, USA
[5] MVZ Dr. Eberhard & Partner, Dortmund, Germany
[6] Medizinische Klinik—Campus Innenstadt, Klinikum der LMU, Munich, Germany
[7] MGZ—Center of Medical Genetics, Munich, Germany
[8] Medizinische Klinik 1, Biomedical Research Laboratory, University of Frankfurt, Frankfurt, Germany
[9] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany
[10] Division of Medical Genetics, University Hospital Basel and Department of Biomedicine, University of Basel, Basel, Switzerland

In a number of families with colorectal adenomatous polyposis or suspected Lynch syndrome/HNPCC, no germline alteration in the *APC*, *MUTYH*, or mismatch repair (MMR) genes are found. Missense mutations in the polymerase genes *POLE* and *POLD1* have recently been identified as rare cause of multiple colorectal adenomas and carcinomas, a condition termed polymerase proofreading-associated polyposis (PPAP). The aim of the present study was to evaluate the clinical relevance and phenotypic spectrum of polymerase germline mutations. Therefore, targeted sequencing of the polymerase genes *POLD1, POLD2, POLD3, POLD4, POLE, POLE2, POLE3* and *POLE4* was performed in 266 unrelated patients with polyposis or fulfilled Amsterdam criteria. The *POLE* mutation c.1270C>G;p.Leu424Val was detected in four unrelated patients. The mutation was present in 1.5% (4/266) of all patients, 4% (3/77) of all familial cases and 7% (2/30) of familial polyposis cases. The colorectal phenotype in 14 affected individuals ranged from typical adenomatous polyposis to a HNPCC phenotype, with high intrafamilial variability. Multiple colorectal carcinomas and duodenal adenomas were common, and one case of duodenal carcinoma was reported. Additionally, various extraintestinal lesions were evident. Nine further putative pathogenic variants were identified. The most promising was c.1306C>T;p.Pro436Ser in *POLE*. In conclusion, a PPAP was identified in a substantial number of polyposis and familial colorectal cancer patients. Screening for polymerase proofreading mutations should therefore be considered, particularly in unexplained familial cases. The present study broadens the phenotypic spectrum of PPAP to duodenal adenomas and carcinomas, and identified novel, potentially pathogenic variants in four polymerase genes.

Cancer Genetics

Key words: familial colorectal cancer, gastrointestinal polyposis syndromes, adenomatous polyposis, Lynch syndrome, next-generation sequencing

Abbreviations: AFAP: attenuated FAP; BER: base excision repair; CHRPE: congenital hypertrophy of the retinal pigment epithelium; CRC: colorectal cancer; EVS: Exome Variant Server; FAP: Familial Adenomatous Polyposis; HNPCC: hereditary nonpolyposis colorectal cancer; MAF: minor allele frequency; MAP: *MUTYH*-Associated Polyposis; MMR: mismatch repair; PPAP: polymerase proofreading-associated polyposis; TGP: 1000 Genomes Project

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: German Cancer Aid (Deutsche Krebshilfe e.V. Bonn); **Grant number:** 108421, 70-2371, 106244; **Grant sponsor:** Gerok-Stipendium of the University Hospital Bonn; **Grant number:** O-149.0098; **Grant sponsor:** NIH Centers for Mendelian Genomics; **Grant number:** 5U54HG006504

DOI: 10.1002/ijc.29396

History: Received 26 Aug 2014; Accepted 19 Nov 2014; Online 20 Dec 2014

Correspondence to: Isabel Spier, MD, Institute of Human Genetics, University of Bonn, Sigmund-Freud-Str. 25, D-53127 Bonn, Germany, Tel.: +[49-(0)228-287-51020], Fax:+[49-(0)228-287-51011], E-mail: isabel.spier@uni-bonn.de

What's new?

A substantial number of families with adenomatous polyposis and Lynch-like phenotype have no known underlying germline mutations. Recently new mutations in the genes encoding DNA polymerase epsilon and delta were identified. Here the authors characterize the frequency and phenotypic spectrum of this newly described polymerase proofreading-associated polyposis (PPAP) syndrome. They broadened the tumor spectrum to duodenal neoplasias, extraintestinal tumors and multiple colorectal carcinomas, underscoring the clinical relevance of this syndrome beyond adenomatous polyposis. In addition, they identified nine novel, potentially pathogenic variants in four polymerase genes.

Familial colorectal cancer (CRC) is genetically heterogeneous. To date, several monogenic forms have been delineated, including various gastrointestinal polyposis syndromes and Lynch syndrome. However, in a number of families with a gastrointestinal polyposis and families meeting the Amsterdam I or II criteria, which define the hereditary nonpolyposis colorectal cancer (HNPCC) phenotype,[1] no pathogenic germline mutation is identified.

Adenomatous polyposis syndromes of the colorectum are characterized by the occurrence of dozens to thousands of adenomatous polyps, which, if not removed early, invariably result in CRC. So far, two different inherited forms can be delineated by molecular genetic analysis: the autosomal dominant familial adenomatous polyposis (FAP), caused by heterozygous germline mutations in the tumor suppressor gene *APC*,[2,3] and the autosomal recessive *MUTYH*-associated polyposis (MAP), caused by biallelic germline mutations of the base excision repair (BER) gene *MUTYH*.[4,5]

In FAP, the vast majority of mutations are truncating point mutations including deep intronic alterations,[6] and large deletions[7] (www.lovd.nl/APC). In MAP, the mutation spectrum is dominated by missense mutations (www.lovd.nl/MUTYH). Mutation detection rates are dependent on the colorectal phenotype. In classical FAP, a pathogenic *APC* germline mutation is identified in up to 90% of index patients whereas *APC* or *MUTYH* germline mutations are detectable in only 20–30% of index patients with a mild disease course (attenuated FAP, AFAP).[8–10]

In families where the Amsterdam criteria are met and tumor tissue shows evidence for microsatellite instability and loss of mismatch repair (MMR) proteins with immunohistochemical staining, a germline mutation in one of the MMR genes *MLH1, MSH2, MSH6* and *PMS2* or the *EPCAM* gene is present in 70–80%[11] (these cases are designated as Lynch syndrome[1]).

However, in about 30% of families meeting the Amsterdam criteria, the tumor tissue is MMR proficient and in up to 50% of families with unexplained adenomatous polyposis no underlying germline mutation is identified even though a genetic cause is likely in both groups.[11–14] Most mutation-negative polyposis cases are sporadic or associated with an uncertain family history, and are characterized by a more attenuated or atypical disease course.[9,14,15] A recent study identified specific germline missense mutations in the genes *POLE* (c.1270C>G;p.Leu424Val; NM_006231) and *POLD1* (c.1433G>A;p.Ser478Asn; NM_002691) as rare, highly penetrant causative events in patients with MMR proficient multiple colorectal adenomas and early-onset CRC.[16] Currently, the condition is termed polymerase proofreading-associated polyposis (PPAP).[17]

The PPAP cases identified so far displayed striking inter- and intrafamilial variability in terms of the number, and the age of onset of colorectal adenomas and carcinomas. The phenotypic pattern ranged from large (>2 cm in diameter) or multiple (≥5) conventional adenomas resembling AFAP or MAP to a clustering of CRCs or early-onset (≤40 years of age) CRCs compatible with a HNPCC phenotype. The tumors were characterized by microsatellite stability and chromosomal instability, a high rate of missense mutations and a somatic mutation spectrum suggestive of a classical pathway of tumorigenesis.

In the meantime, a *de novo* occurrence of the *POLE* mutation c.1270C>G;p.Leu424Val has been described in a female patient diagnosed with CRC and >35 colonic polyps at age 28 years. In another family, a novel mutation in *POLD1* (c.1421T>C;p.Leu474Pro) was found.[18] In families with *POLD1* mutations, a predisposition to endometrial cancer has been suggested.[16,18] However, no study to date has reported convincing evidence for an increased risk of other extracolonic manifestations.

POLE and *POLD1* encode the catalytic and proofreading subunits of the polymerase enzyme complexes ϵ (Pol ϵ) and δ (Pol δ), respectively. Both polymerase enzymes are heterotetramers which contain accessory subunits (POLE2/3/4 and POLD2/3/4). These subunits are involved in synthesis regulation and co-factor binding.[17] Somatic exonuclease domain mutations in *POLE* have been identified in CRC and endometrial cancer patients, and showed association with hypermutability and microsatellite-stability. In contrast, corresponding somatic mutations in *POLD1* are rare and the causal relevance is unknown so far.[17,19–21]

The aim of the present study was to further evaluate the clinical relevance and phenotypic spectrum of polymerase germline mutations in more detail. A large and well-characterized cohort of unrelated index patients with unexplained colorectal adenomatous polyposis or multiple colorectal adenomas and carcinomas was screened for point mutations in eight polymerase genes using whole-exome and targeted high-throughput sequencing approaches.

Subjects and Methods

Patients/data collection

The inclusion criteria were: (*i*) the presence of at least 20 synchronous or 40 metachronous, and histologically confirmed colorectal adenomas irrespective of inheritance pattern and no germline mutation in the *APC* and *MUTYH* genes (unexplained polyposis cases) or (*ii*) fulfilment of the Amsterdam I or II criteria[22,23] and the presence of microsatellite stable tumors (familial CRC cases) (for details see Supporting Information).

All unrelated index patients with a colorectal adenomatous polyposis were screened for germline mutations in the *APC* and *MUTYH* genes including multiplex ligation-dependent probe amplification (MLPA) analysis and screening for pathogenic deep intronic *APC* mutations as previously described.[6,7] In all index patients with HNPCC, microsatellite analysis and/or immunohistochemical analysis of the tumor tissue for MMR proteins was performed.[11]

Information on medical and family history was obtained during genetic counseling sessions, from medical records, and comprehensive structured telephone interviews. Nearly all of the included index patients were of Caucasian/Western European origin according to family name and self-report. Relatives were only classified as being affected if their medical records confirmed fulfilment of the inclusion criteria. The study was approved by the local ethics review boards (including Ethics Committee no. 224/07 and 115/09 of the Medical Faculty, University of Bonn). All patients gave informed consent to participate in the study.

Next generation sequencing

In total, 266 unrelated index patients were sequenced. In 191 patients (144 polyposis patients and 47 familial CRC patients), mutation screening of the polymerase genes *POLD1, POLD2, POLD3, POLD4, POLE, POLE2, POLE3* and *POLE4* was performed using targeted next generation sequencing. In another 75 polyposis patients, exome sequencing was done. The inclusion criteria regarding a polyposis were the same for the targeted and the exome sequencing group.

For targeted next generation sequencing, the oligonucleotide probes were designed with the Illumina DesignStudio based on the TruSeq enrichment protocol (Illumina, San Diego, USA). A total of 1 μg genomic DNA was extracted from leukocytes and fragmented using sonication technology (Bioruptor, Diagenode Liège, Belgium). Fragments were end repaired and adapter ligated using Illumina's TruSeq® DNA HT Sample Preparation Kit. Custom capture of targeted regions was performed on pools of 12 indexed libraries using Illumina's TruSeq enrichment protocol. The captured DNA was sequenced using an Illumina HiSeq2000 sequencer with 2 × 100 bp paired-end reads. A 30× coverage was achieved for at least 96% of the targeted bases (Supporting Information Table S2). For data analysis, the *Varbank* pipeline version 2.6 and interface was used. Primary data were filtered according to signal purity using the Illumina Realtime Analysis software version 1.8. Subsequently, the reads were mapped to the human genome reference build GRCh37 using the *BWA* version alignment algorithm,[24] *GATK* version 1.6,[25] *SAMtools* version 0.1.7[26] and *Dindel* version 1.01[27] were used to mark duplicated reads, perform local realignment around short insertion and deletions, recalibrate the base quality scores and call single nucleotide polymorphisms (SNPs) and short Indels. For filtering of the targeted sequencing data, data scripts developed at the *Cologne Center for Genomics* were applied to detect protein changes, affected donor and acceptor splice sites and overlaps with known variants. Acceptor and donor splice site mutations were analyzed using a Maximum Entropy model,[28] and then filtered for effect changes. To exclude pipeline-related artifacts (minor allele frequency [MAF] <0.02), filtering was performed against an inhouse-database of variants from 511 exomes that had been obtained from epilepsy patients.

Whole-exome sequencing was performed at the Yale Center for Genome Analysis *via* capture using NimbleGen 2.1M human exome array followed by paired-end sequencing on a HiSeq 2000 instrument, as previously described.[29] Targeted bases were covered by a mean of 67 independent reads, with an average of 94% of all bases covered eight or more times (Supporting Information Table S2). 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. To filter the exome sequencing data for potential pathogenic variants in the eight polymerase genes (frameshift, stop gain/loss or start loss, mutations at highly conserved splice sites, and nonsynonymous mutations), the Cartagenia BENCHlab NGS platform version 3.0.4 (Leuven, Belgium) was used.

Subsequently, the targeted and exome sequencing data were filtered for high-quality, rare (MAF <0.01) variants using allele frequencies from *dbSNP*, the *1000 Genomes* database (TGP) and the *Exome Variant Server* (EVS). The remaining missense variants were selected manually for MAF <0.001 and deleterious effect as predicted by at least two of three *in silico* analysis tools (PolyPhen-2, MutationTaster and SIFT).

Sanger sequencing

Variant validation and segregation analysis was performed using Sanger sequencing and standard protocols. Briefly, primers were designed using Primer3 v.0.4.0 (primers available on request). Genomic DNA was extracted from leukocytes and used to amplify the genomic region of the respective variant. PCR products were purified and sequenced on an ABI 3500xl Genetic Analyzer (Applied Biosystems, Darmstadt, Germany).

Haplotype analysis

Haplotype analysis of families with a segregating *POLE* mutation was performed using a panel of five microsatellite

markers. These flanked the *POLE* region and spanned ~3.3 Mb. The order of markers on chromosome 12q was: CEN–D12S1045–D12S343–D12S1723–*POLE*–D12S1638–D12S2342–TEL.

Structural and conservation analysis

For the analyses, 40 nonredundant, full-length sequences of the POLD1 or POLE protein from diverse organisms were retrieved from the NCBI RefSeq protein database and aligned using T-Coffee.[30] The alignments were used to create Weblogo diagrams of local sequence conservation.[31] For structural analysis, the crystal structures of yeast POLD1 (3IAY) and POLE (4M8O) proteins were used, which share 51% and 55% identity with the corresponding human proteins within the crystallized sequence, respectively.[32,33] Structural figures were created using PyMOL v.1.4.1 (Schrödinger LLC).

Results

The study cohort (266 unrelated index patients) comprised 219 patients with a histopathologically confirmed colorectal adenomatous polyposis in whom no *APC* or *MUTYH* germline mutations could be identified, and 47 patients who met the Amsterdam criteria for HNPCC and exhibited microsatellite stable tumors (familial CRC patients). The mean age at diagnosis was 45 years in the polyposis patients (range 12–78 years), and 52 years in the familial CRC patients (range 30–86 years). The vast majority of polyposis patients presented with a more attenuated colorectal phenotype (late onset disease and/or <100 colorectal adenomas) with no evident extracolonic lesions and an unremarkable family history. Five of the familial CRC patients met the Amsterdam II criteria, and the remainder met the Amsterdam I criteria. The clinical features of the study cohort are summarized in Supporting Information Table S1.

Frequency and phenotype of the POLE mutation p.Leu424Val

By targeted next generation and exome sequencing, all 76 exons of *POLE* and *POLD1* were covered (average read depth 478× and 40×, respectively) (Supporting Information Table S2). The previously described *POLE* missense mutation c.1270C>G;p.Leu424Val was found in four unrelated index patients. About the family history of one patient (F1505), only insufficient information was available, however in the maternal family several members died because of cancer at a rather young age. Three patients had a positive family history that was suggestive of autosomal dominant inheritance. From the three respective families, ten further mutation carriers were identified among relatives for whom DNA was available. Thus a total of 14 mutation carriers were identified (six females, eight males, Table 1). In all ten relatives, the mutation segregated with the phenotype (Table 1, Fig. 1, Supporting Information Figs. 1A–1C). Subsequent haplotype analysis using microsatellite markers around the *POLE* locus showed that all three families exhibited differing haplotypes. The index patients of the two parts of family F156/F381 were primarily referred as independent patients. The relationship was confirmed by haplotype analysis and extended pedigree information provided by family members.

In families F156/F381, F354 and F1505, the colorectal phenotype of the mutation carriers was that of a more attenuated polyposis with either a late-onset manifestation or with early-onset occurrence of <100 polyps. The fourth family (H1427) met the Amsterdam I criteria. Here, the index patient and the mutation carrying mother had <10 and the mutation carrying cousin had >20 colorectal polyps. In all four families, the polyps were mainly adenomas, although four patients showed also few hyperplastic polyps.

The mean age at diagnosis in all 14 patients was 35 years (range 16–63 years). The diagnosis in a 16-year old patient (son of index patient F354) was assigned after a surveillance colonoscopy; the earliest age at diagnosis at symptomatic stage was 27 years (index patients F156 and F354).

CRC occurred in 10/14 (71%) of the mutation carriers and affected all families. The age at diagnosis ranged between 27 and 63 years. In all index patients, CRC was diagnosed at a symptomatic stage, while a proportion of their relatives had been diagnosed with CRC as a result of presymptomatic surveillance. The CRCs occurred throughout the colon, with no predilection site. Five mutation carriers (39% of the 14 mutation carriers or 56% of the nine CRC patients) had synchronous and/or metachronous CRCs. All five patients underwent a partial colorectal resection. In three of the five patients, a subsequent proctocolectomy was performed due to the development of a polyposis phenotype . In three of the ten CRC patients (30%), or in three of all 17 CRC of known histology (18%), a mucinous adenocarcinoma was described. No other Lynch syndrome-associated histological findings such as tumor-infiltrating lymphocytes, signet ring histology or a medullary growth pattern were reported. In four patients, carcinoma tissue was investigated for MMR deficiency and only microsatellite stability was found.

All 14 mutation carriers had undergone at least one gastroduodenoscopy. Three had a history of gastric fundic gland polyps; in none of the three patients chronic use of proton pump inhibitors was reported. In 8/14 patients (57%) duodenal adenomas were described. Of these, one individual (F381, III:1) was diagnosed with a duodenal adenocarcinoma at age 45 years (Table 1). In one patient (F156, III:5) jejunal polyps were found.

Further malignant tumors were reported in two mutation carriers. One (H1427, III:8) had an ovarian carcinoma at age 33 years, classified histologically as solid clear cell malignant epithelial ovarian tumor (in addition to adnexectomy a hysterectomy was done). In another (F354, II:5), a neuroendocrine carcinoma of the ascending colon was diagnosed at age 41 years and a glioblastoma at age 47 years. In addition, a child of F156, III:3 was reported by the family to have had an adrenal neuroblastoma, and a child of F156, III:4 was reported to have had a glioblastoma. These children died at

Table 1. Phenotype of patients with *POLE* mutations in the proofreading domain

Family ID	POLE mutation	Family History	Relationship/ pedigree ID	Age at diagnosis (years)	Phenotype	No. of colorectal adenomas	Histology polyps	CRC—age at diagnosis (years)	Position CRC	Histology CRC	Duodenal phenotype[1]	Extracolonic phenotype/ further tumors/ diseases	Mismatch repair analysis
F156/F381	p.Leu424Val	F	F156 III:5 (index)	27	Typical	51–100	Adenomas	27	Left colic flexure	Adenocarcinoma	Moderate (adenomas), also 4 in jejunum	Pilomatricomas, primary biliary cirrhosis	
			F156 III:1	35	Atypical	>100	Adenomas, HP	No			Moderate (adenomas)	Numerous gastric fundic gland polyps	
			F156 III:3	32	Attenuated	Multiple	Adenomas	42	Caecum	Not available	Moderate (adenomas)	Retroperitoneal fibrosis	
			F156 III:4	30	Attenuated	51–100	Adenomas, HP	30 (4×), +46	1) Left and right colic flexure, transverse colon (4 synchronous tumors), 2) rectum	Adenocarcinomas (left colic flexure: mucinous)	Normal	Some gastric fundic gland polyps	
			F381 II:6 (index)	36	Attenuated	51–100	Adenomas	36 (2×) +56 + 66	1) Caecum, rectum, 2) transverse colon, 3) rectum	Adenocarcinoma (caecum: mucinous)	Moderate–severe (adenomas)	Unconfirmed self report of CHRPE	MSS
			F381 III:1	34	Attenuated	21–50	Adenomas	36	Right colic flexure	Adenocarcinoma	Duodenal cancer (45 years)	Sarcoidosis	MSS (duodenal cancer), IHC no loss[2]
			F381 III:2	37	Attenuated	21–50	Adenomas	No			Normal	Nothing reported	IHC no loss (adenoma)
			F381 III:5	32	Attenuated	11–20	Adenomas	No			Mild (adenoma)	Some gastric fundic gland polyps	
F354	p.Leu424Val	F	II:5 (index)	27	Typical	51–100	Adenomas	27 + 38+39	1) Rectum, 2 + 3) ascending colon (carcinoma in situ)	Adenocarcinoma	Moderate (adenomas)	Fibromas, neuroendocrine carcinoma ascending colon (41 years), glioblastoma (47 years)	
			III:5	16	Typical	11–20	Adenomas, some HP	No			Normal	Nothing reported	
F1505	p.Leu424Val	S	Index	38	Attenuated	51–100	Adenomas	38	Transverse colon	Adenocarcinoma	Normal	One lipoma	MSS, IHC no loss
H1427	p.Leu424Val	F	III:1 (index)	46	Amsterdam I	<10	Adenomas	46 + 48	Rectum, caecum	Adenocarcinoma (rectum: mucinous)	Normal	Nothing reported	MSS, IHC no loss
			II:2	63	Amsterdam I	<10	adenomas	63	Rightside	Not available	Normal	Nothing reported	
			III:8	45	Attenuated	21–50	Adenomas, some HP	45 (2×)+53	1) Sigmoid colon, 1 + 2) transverse colon	Adenocarcinoma	Mild	Ovarian carcinoma (33 years)	
F1228	p.Pro436Ser	S/F	III:4 (index)	31	Attenuated	21–50	Adenomas, HP	31 + 41	Rectum, caecum	Adenocarcinoma (rectum: mucinous)	Moderate-severe (adenomas), also jejunal polyps	Nothing reported	

CHRPE = congenital hypertrophy of the retinal pigment epithelium, CRC = colorectal cancer, F = familial, HP = hyperplastic polyps, IHC = immunohistochemistry, MSS = microsatellite stable tumor, S = sporadic.
[1]All patients had at least one gastroduodenoscopy.
[2]MSH6 not clearly interpretable, no mutation in *MSH6*.

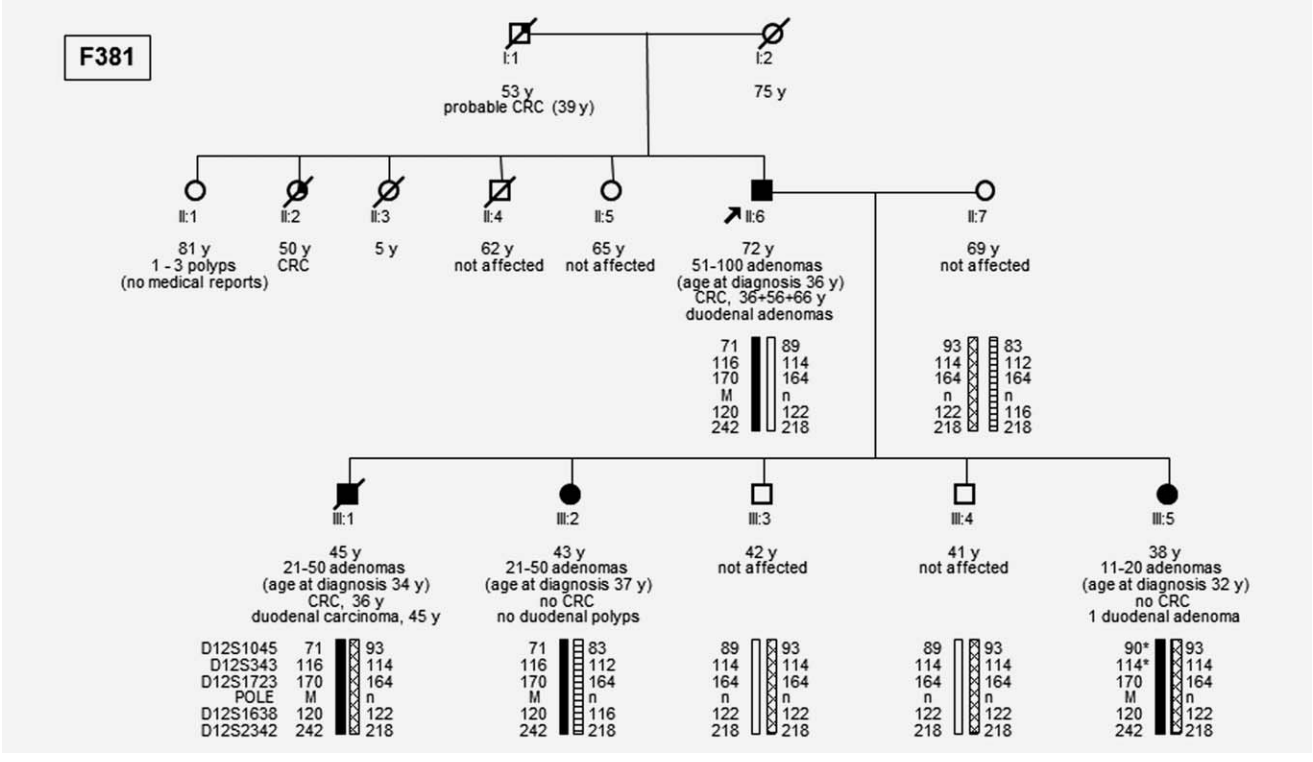

Figure 1. Pedigree of family F381 with the *POLE* mutation c.1270C>G;p.Leu424Val. The age stated below of each individual is the age at last contact or the age at death. The cumulative number of colorectal adenomas, the occurrence of CRC and duodenal polyps together with the respective age at diagnosis are indicated. The index patient is indicated with an arrow. Filled symbol, affected; M = mutation; n = wild type allele. The results of the haplotype analysis using a panel of five microsatellite markers flanking the *POLE* region is shown. All affected individuals share the same haplotype, as marked with a black bar (same haplotype as F156, Supporting Information Fig. 1A). *The allele combination of markers D12S1045 and D12S343 in patient III: 5 points to a previous recombination event.

the age of 4 and 5 years, respectively. No medical records were available to confirm these diagnoses.

Benign extraintestinal manifestations were reported in five patients including a patient with fibromas (F354, II:5), one patient with a lipoma (F1505), another case with pilomatricomas and his sister with retroperitoneal fibrosis following a colectomy (F156, III:5 and III:3). She had received no chemotherapy. One patient (F381, II:6) reported a history of congenital hypertrophy of the retinal pigment epithelium (CHRPE).

Novel germline mutations in four polymerase genes

By applying stringent filter criteria, a further nine different unique or rare (MAF <0.1% in population-based controls, based on dbSNP, the EVS and TGP data) germline variants were identified in a total of four polymerase genes (*POLD1, POLD3, POLE, POLE2*). These comprised one frameshift mutation and eight heterozygous rare missense mutations (Table 2). All putative missense variants were predicted to be pathogenic by at least two of the three applied *in silico* analysis tools and are located in highly conserved regions; none were predicted to generate new or activate cryptic splice sites. Two different missense variants were detected in *POLD1* and *POLD3*, respectively, and three in *POLE*. In *POLE2*, a frameshift and a missense variant were found. With the exception of the *POLE* missense mutation c.2683G>A;p.Ala895Thr, each variant was identified in one patient only. One patient (H2399) carried two variants (in *POLD1* and *POLD3*, respectively).

Two of the three *POLE* missense mutations (c.1306C>T;p.Pro436Ser and c.1337G>A;p.Arg446Gln) are located in the exonuclease domain (Fig. 2*b*). The variant c.1306C>T had occurred *de novo* and was identified in a patient (F1228) with attenuated polyposis (>40 colorectal polyps, mainly adenomas, but also hyperplastic polyps), early-onset, mucinous, metachronous CRC (ages 31 and 41 years) and duodenal adenomas (Table 1, Supporting Information Fig. 1D). The mother presented with few polyps (<10, histologically adenomas and hyperplastic polyps) and an age of diagnosis of >60 years. A maternal uncle had been diagnosed with CRC at age 49 years.

The exonuclease *POLE* variant c.1337G>A;p.Arg446Gln was found in a sporadic case (F708) with a classical phenotype (age at diagnosis 28 years, > 100 polyps). The rare *POLE* variant c.2683G>A;p.Ala895Thr, which is located outside the exonuclease domain, was identified in two sporadic polyposis index patients (F1275 and F1495) with an attenuated phenotype. No DNA from relatives was available for segregation analysis.

Table 2. Potential pathogenic polymerase mutations

Patient ID	Gene	Mutation	Frequency[1]	Designation (dbSNP)	Mutation prediction			Somatic mutations (tumor)[2]	Phenotype	Family history	Segregation
					PolyPhen-2	Mutation taster	SIFT				
H2399	POLD1	c.55C>T;p.Arg19Cys	0.008%	rs368033860	Benign	Disease causing	Damaging	NA	Amsterdam II	F	Detected in affected mother (EC)
F1798	POLD1	c.245C>T;p.Pro82Leu	NA	rs201006221	Probably damaging	Disease causing	Damaging	NA	Attenuated	S	NA
F5039	POLD3	c.956A>G;p.Lys319Arg	0.015%	rs201575276	Probably damaging	Disease causing	Tolerated	NA	Attenuated	S	NA
H2399	POLD3	c.1154G>A;p.Arg385His	0.1%	rs200646605	Probably damaging	Disease causing	Damaging	1× (EC)[3]	Amsterdam II	F	Detected in affected mother (EC)
F1228	POLE	c.1306C>T;p.Pro436Ser	NA	NA	Probably damaging	Disease causing	Damaging	1× (CRC)[4]	Attenuated	S/F	*de novo*
F708	POLE	c.1337G>A;p.Arg446Gln	0.05%	rs151273553	Probably damaging	Disease causing	Tolerated	NA	Typical	S	NA
F1275 F1495	POLE	c.2683G>A;p.Ala895Thr	0.03%	rs201115064	Probably damaging	Disease causing	Tolerated	NA	Attenuated	S	NA
F291	POLE2	c.823C>A;p.Leu275Ile	0.03%	rs141483427	Probably damaging	Disease causing	Damaging	NA	Typical	F	Presumed segregation
F1467	POLE2	c.1328_1329insT; p.Leu443Phefs*17	NA	NA	NA	NA	NA	1× (CRC)[5]	Attenuated	S/F	NA

EC = endometrial cancer, CRC = colorectal cancer, F = familial, NA = not available, S = sporadic.
[1]According to Exome Variant Server and the 1000Genomes Project.
[2]Previously described somatic mutations in the same codon/region; according to COSMIC and The Cancer Genome Atlas Network.[20]
[3]c.1153C>G;p.Arg385Gly; sample TCGA-BS-A0TI-01.
[4]c.1307C>G;p.Pro436Arg; sample TCGA-AA-A010.
[5]Frameshift mutation in the same region: c.1370_1371delAT;p.Tyr457fs*9; mutation COSM1369919.

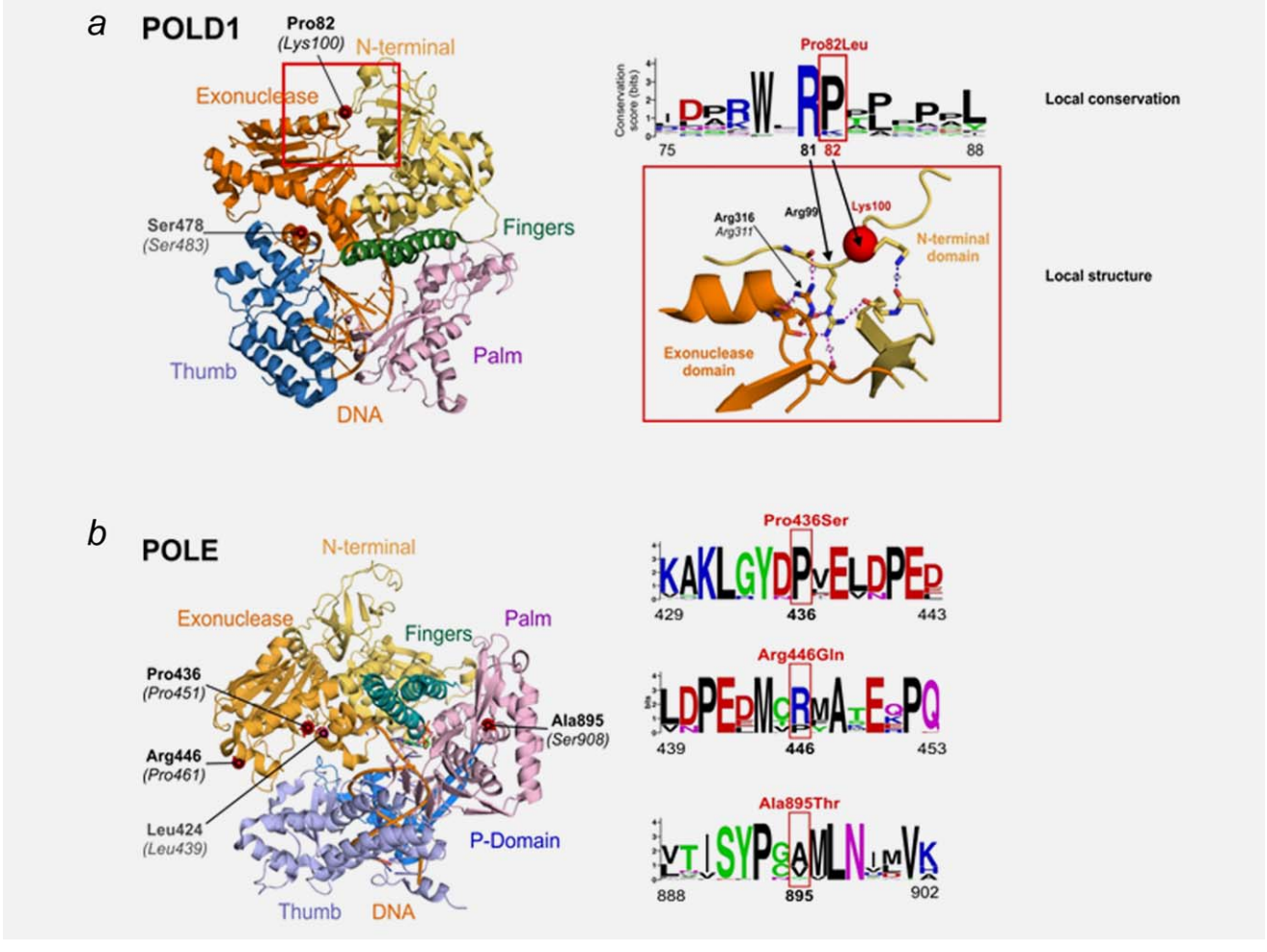

Figure 2. Structural analysis of the POLD1 (*a*) and POLE (*b*) proteins. The amino acid positions of the novel germline mutations c.245C>T;p.Pro82 Leu in *POLD1* (*a*, detailed view on the right) and c.1306C>T;p.Pro436Ser, c.1337G>a;p.Arg446Gln and c.2683G>a;p.Ala895Thr in *POLE* (*b*) are indicated. Furthermore, the amino acid positions of the previously described mutations c.1433G>a;p.Ser478Asn in *POLD1* and c.1270C>G;p.Leu424Val in *POLE* are shown. The corresponding positions in yeast are depicted in brackets. On the right side, the degree of local conservation is shown. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The *POLE2* variant c.823C>A;p.Leu275Ile (NM_002692.3) was identified in a patient with 80–100 polyps and CRC at age 31 years (F291, Supporting Information Fig. 1E). Gastric fundic gland polyps were also described. The *POLE2* variant was found in the father (CRC at age 60 years, few hyperplastic polyps reported), the two sisters (colonoscopy revealed one adenoma and one sessile serrated adenoma in one sister and no pathological findings in the other) and in the daughter (no polyps, last colonoscopy at age 22 years). The variant was not detected in the healthy mother (no history of colonoscopy) or the healthy son (last colonoscopy at age 20 years was unremarkable).

In a family meeting, the Amsterdam II criteria, the index patient (H2399) and his mother carried two missense variants each (*POLD1*:c.55C>T;p.Arg19Cys and *POLD3*:c.1154 G>A;p.Arg385His; NM_006591.2). The index patient was diagnosed with CRC at age 44 years, while the mother was affected by endometrial cancer at age 66 years (no information concerning her colorectal phenotype was available). In addition, the maternal uncle and grandmother had CRC at age 55 and 81 years, respectively (Supporting Information Fig. 1F).

The remaining patients with rare variants in *POLD1*, *POLD3* and *POLE2* (F1798, F5039 and F1467, respectively) exhibited an attenuated polyposis phenotype and were sporadic cases or showed an unclear family history (Table 2). No parental DNA was available for segregation analysis.

Discussion

To explore the frequency and phenotypic spectrum of the recently described germline missense mutations in *POLE* and *POLD1* and to look for variants in additional polymerase genes, we performed a systematic mutation screening in a large and well characterized cohort of 266 unrelated patients with clinically confirmed and genetically unexplained multiple colorectal adenomas or familial CRC. The clinical features

and family histories of the present cohort are consistent with published data for other mutation negative samples.[12–14]

Impact and clinical characteristics of the POLE mutation p.Leu424Val

The *POLE* mutation, c.1270C>G;p.Leu424Val was identified in four unrelated families, in whom thereby a PPAP was diagnosed. Although this corresponds to only 1.5% of index patients in the whole cohort (4/266), the frequency is considerably higher in selected subgroups, in particular, in families with a positive family history (4%; 3/77) and familial polyposis cases (7%; 2/30). This is consistent with the original study, which identified the mutation in one of 15 selected patients (>10 colorectal adenomas by age 60 years) but in only 0.3% of a large validation cohort (12/3,805). The latter comprised 1,560 unselected CRC cases in addition to patients with familial colorectal tumors, multiple adenomas and early-onset disease.[16] As with the present cohort, the majority of mutation carriers had a positive family history that was suggestive of autosomal dominant inheritance and high-penetrance. The fact that the alteration was found to segregate with different haplotypes support the view that the mutation POLE:c.1270C>G;p.Leu424Val is a mutational hot spot rather than a founder mutation.

The colorectal phenotype in the four mutation-carrying families from the present report is compatible with previous observations. Some individuals presented with ≥100 adenomas or early-onset disease, while others had ≤10 adenomas and a late-onset. In around 30% of cases, hyperplastic polyps were reported. Wide variation in the phenotype was observed, even within families. As would be expected from the study inclusion criteria, the frequency of CRC was high. However, a remarkably high proportion of the patients had multiple syn- or metachronous carcinomas, and the percentage of tumors with one criteria of microsatellite instability-high (MSI-H) histology (mucinous adenocarcinomas) was comparable to that observed in Lynch syndrome patients, irrespective of whether the index patient had been recruited as a polyposis or HNPCC case. In contrast to Lynch syndrome or AFAP, no predilection in CRC location was evident. As in the previous study by Palles *et al.*,[16] all investigated tumors showed microsatellite stability and normal immunochistochemical results.

Detailed phenotyping enabled us to broaden the spectrum of extracolonic features in patients with a PPAP. As in FAP, the present data indicate that both duodenal adenomas, ranging from a mild to a severe phenotype, and gastric fundic gland polyps are frequent manifestations of this condition. Based on the occurrence of a duodenal carcinoma in one patient, an increased risk for duodenal cancer can be presumed. The prevalence and severity of duodenal tumors was higher in polyposis patients compared to individuals meeting the Amsterdam criteria only. The observation of one patient with jejunal polyps indicates that adenomas might also occur in distal parts of the small bowel.

In 43% of the patients, extraintestinal lesions were described. Of particular interest are the pilomatricomas, retroperitoneal fibrosis, possible CHRPEs, fibromas and a lipoma that were reported in a single patient respectively, since these are typical manifestations of the established hereditary syndromes FAP and MAP. This suggests a significant clinical overlap between these conditions, even beyond the gastrointestinal tract. Early-onset ovarian cancer, reported in one patient, was also described in a family from the initial report (at age 35 years, in addition to two CRC at age 50 years; testing for the familial mutation was not possible).[16] Furthermore, the occurrence of two glioblastomas in our cohort is remarkable (one in childhood where genetic testing was impossible, and one at age 47 years in a confirmed mutation carrier). Palles *et al.* reported an astrocytoma in one carrier of the *POLD1* p.Ser478Asn alteration at age 26 years and in one relative of a family with the *POLE* p.Leu424Val alteration at age 67 years (the relative had three CRCs between 39 and 54 years of age; testing for the familial mutation was impossible).[16] Glioblastomas belong to the tumor spectrum of Lynch syndrome and ovarian cancer to both Lynch syndrome and MAP.[34,35] These findings emphasize the overlap in the tumor spectrum in hereditary tumor syndromes caused by mutations in DNA repair genes. Whether neuroendocrine carcinomas and adrenal neuroblastomas are causally related to the *POLE* mutation remains unclear.

Naturally, it cannot be excluded that some of the findings occurred by chance; however, the number of extracolonic tumors in our and other studies is striking and suggests an association with the POLE mutation. Nonetheless, more mutation carriers with detailed phenotype data are needed to confirm these observations.

Given the limited clinical data and small number of families with polymerase germline mutations described to date, current surveillance recommendations must be considered preliminary. However, the early-onset manifestation of colorectal adenomas described in the second decade, the increased risk and multiplicity of CRCs from the third decade of life and the high prevalence of duodenal tumors suggest that frequent endoscopic examinations similar to those recommended for FAP and Lynch syndrome are indicated in polymerase germline mutation carriers. These recommendations will also be influenced by national guidelines and should include colonoscopies every 1–2 years starting at age 18–20 years, and gastroduodenoscopies at least every three years (in the case of normal findings) from around age 25–30 years.[36,37] The timing and extent of prophylactic surgery has yet to be determined. In none of our female *POLE* mutation carriers, endometrial cancer is reported; however, no systematic screening has been performed in any of these patients. Based on the limited data available so far, no specific surveillance regarding endometrial cancer seems to be warranted.

Spectrum of additional polymerase gene variants

Palles *et al.* detected the *POLD1* mutation c.1433G>A;p.Ser478Asn in 2/15 selected patients but in only one further patient from a large validation cohort of selected and

unselected CRC cases (1/3,805 = 0.03%).[16] Furthermore, they described three rare heterozygous *POLD1* variants, each of which was found in a single patient. The variant c.981C>G;p.Pro327Leu seems to be of particular functional relevance. Recently, a patient with a novel *POLD1* mutation in the proofreading domain (c.1421T>C;p.Leu474Pro) was identified in a cohort of 858 familial/early onset CRC and polyposis cases.[18] Together with the absence of *POLD1* mutations in our cohort, those mutations seem to occur at a lower frequency than the *POLE* mutation in families selected for colorectal adenomas and carcinomas.

Using a comprehensive approach to screen all known subunits of the polymerases δ and ϵ, we identified nine additional different rare heterozygous germline variants in a total of four polymerase genes. All are novel or rare (MAF<0.1%) and predicted to be pathogenic by at least two prediction analysis tools (Table 2).

The *POLE* missense mutation c.1306C>T;p.Pro436Ser was identified in a patient whose phenotype is similar to patients with the known *POLE* mutation. The index patient shows the most pronounced gastrointestinal phenotype in the family. Thus, the *de novo* status of the mutation seems to be compatible with the family history. This variant is located also in the exonuclease domain of the protein and lies in a disordered loop, which becomes ordered on DNA binding.[16] Furthermore, it is structurally located in direct vicinity to the previously described mutation p.Leu424Val (Fig. 2*b*). An amino acid change at the same position (p.Pro436Arg) has been described previously as a somatic mutation in sporadic CRC.[20] Based on this information it is likely that the germline mutation is pathogenic. However, it is possible that there are further genetic risk factors for CRC in this family.

The *POLE* variant c.1337G>A;p.Arg446Gln leads to a substitution of a highly conserved arginine. However, it is located at the outer surface of the exonuclease domain (Fig. 2*b*) and thus, it remains unclear yet, whether or not the function of the exonuclease is impaired. Interestingly, another rare variant at the same codon, which is predicted to be probably-damaging by three *in silico* analysis tools, is listed in the EVS data (c.1336C>T;p.Arg446Trp, rs200403177, frequency 0.008%). Nonetheless, the polymerases are highly conserved proteins in general and thus, pathogenic changes may be over-called.

The co-occurrence of the variants *POLD1*:c.55C>T; p.Arg19Cys and *POLD3*:c.1154G>A;p.Arg385His in two patients of the same family is interesting since it points to a possible functional interaction (see below). In addition the occurrence of endometrial cancer is remarkable, since, endometrial cancers have been reported recently in four carriers of the *POLD1* mutation p.Ser478Asn, as well as in three further family members for whom no testing was possible, and in two carriers of the *POLD1* mutation p.Leu474Pro.[16,18]

The variant p.Arg19Cys lies outside the proofreading domain of *POLD1*, however, it is part of a potential bipartite nuclear localization signal, which could lead to a deficiency of the polymerase. Another amino acid change at the same position (c.56G>A;p.Arg19His) occurs with a frequency of 0.05% in population-based cohorts (EVS). Remarkably, a somatic mutation in the same codon affected by the *POLD3* variant p.Arg385His has been reported in an endometrial carcinoma (*COSMIC*), and a common intronic polymorphism in *POLD3* has been associated with a moderately increased risk of CRC.[38] For *POLD3,* only polymerase domains are described. However, the subunit POLD3 (p66) seems to be necessary for accumulation of the POLD1 (p125) catalytic core at sites of local damage.[39] Interestingly, Lemmens *et al.*[40] demonstrated that a highly conserved region between amino acids 384 and 399 of p66 is phosphorylated, and that this is associated with the co-localization of p66 with the p125 subunit. A plausible hypothesis therefore is that mutations in *POLD3* impair the interaction with POLD1 and thus the proofreading function of POLD1. Whether the *POLD3* variant p.Arg385His affects protein function in this way is unknown. However, the occurrence of one variant in both *POLD1* and *POLD3* in a single patient and his affected mother points to a putative impaired interaction between the two proteins, resulting in digenic inheritance. Since *POLD1* and *POLD3* are located on different chromosomes (19 and 11, respectively), mutations in these genes are inherited independently.

The recurrence of the rare *POLE* variant c.2683G>A; p.Ala895Thr (frequency 0.03% in the general population) in two unrelated polyposis patients indicates that the mutation might be of functional relevance to the disease. This variant and *POLD1*:c.245C>T;p.Pro82Leu lie outside the exonuclease domain of the respective genes (Fig. 2). Nevertheless, the variant Pro82 Leu could affect the function of the exonuclease, since the binding of the exonuclease domain to the N-terminal domain is conveyed by the proximate Arg81 (Fig. 2*a*).

If causative, the segregation data for the *POLE2* variant c.823C>A;p.Leu275Ile point to a reduced penetrance or rather mild phenotypic expression. For *POLE2,* only polymerase domains are described. Such mutations may lead to insufficient polymerase function, involving incorporation of mispaired bases. Furthermore, it can be assumed in general, that mutations in the accessory subunits of the polymerases result in impaired function of the proofreading subunits, as described above.[39]

One of the rare variants identified in *POLE2* is a frameshift mutation (c.1328_1329insT;p.Leu443Phefs*17). A somatic *POLE2* frameshift mutation in the same region has also been identified in CRC (c.1370_1371delAT;p.Tyr457fs*9; COSM1369919, *COSMIC*). According to current knowledge, truncating mutations are not typical disease causing events in PPAP patients. However, a recent study described a sporadic CRC patient (diagnosed at age 26 years) with a frameshift *POLE* mutation (c.5621_5622delGT).[41] According to the EVS data, only two of the 148 listed coding variants in *POLE2* are frameshift mutations and both occur at a low frequency (0.008%, respectively), indicating that frameshift mutations are rare events in *POLE2*.

In conclusion, the present study identified the previously described *POLE* mutation c.1270C>G;p.Leu424Val in a

substantial number of patients with an autosomal dominant inheritance pattern and multiple colorectal adenomas, adenomatous polyposis, or familial CRC. Our data confirm the highly variable colorectal phenotype of PPAP, and broaden the phenotypic spectrum to include a high frequency of duodenal tumors and various extraintestinal lesions that are characteristic of established hereditary CRC syndromes. PPAP seems to be characterized by a high prevalence of multiple CRCs. The broad clinical overlap between PPAP, Lynch syndrome and MAP, both in terms of gastrointestinal manifestations and extraintestinal tumors, seems to be based on a similar pathogenesis involving reduced DNA repair.

Of the nine novel variants, the *POLE* mutations and the two mutations identified in two individuals from the same family (*POLD1*, *POLD3*) are likely to be pathogenic. However, since *in silico* programs are often inconsistent and the polymerases are highly conserved in general, conclusions regarding the functional impact should be drawn with caution unless more data is available. To verify the causal relevance of these variants and identify recurrent mutations, studies of larger cohorts and functional analyses are warranted.

The results of both the previous reports and the present study suggest that screening for *POLE* germline mutations in exons encoding the proofreading domain should be seriously considered, particularly in unexplained familial cases of colorectal adenomas and microsatellite-stable CRCs. More extensive clinical data are required before valid surveillance recommendations can be formulated. However, according to present knowledge, regular gastroduodenoscopies at intervals similar to those recommended for FAP seem to be warranted in addition to colonoscopies and prophylactic surgery. An increased risk for endometrial cancer in *POLD1* mutation carriers seems likely; if this observation can be confirmed in future studies, the surveillance recommendations might be extended accordingly.

Databases/URLs

APC locus-specific mutation database: www.lovd.nl/APC
dbSNP (The Single Nucleotide Polymorphism database): https://www.ncbi.nlm.nih.gov/SNP/
COSMIC (Catalogue of somatic mutations in cancer): http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/
EVS (Exome Variant Server): http://evs.gs.washington.edu/EVS/
MutationTaster: www.mutationtaster.org
MUTYH locus-specific mutation database: www.lovd.nl/MUTYH
NCBI: http://www.ncbi.nlm.nih.gov/
PolyPhen-2: http://genetics.bwh.harvard.edu/pph2/
Primer3 v.0.4.0: http://frodo.wi.mit.edu/primer3/input.htm
SIFT (Sorting Intolerant From Tolerant): http://sift.jcvi.org/
TGP (1000 Genomes Project): http://www.1000genomes.org
VARBANK: https://varbank.ccg.uni-koeln.de

Acknowledgement

Authors thank the patients and their families for participating in the study.

References

1. Kravochuck SE, Kalady MF, Burke CA, et al. Defining HNPCC and Lynch syndrome: what's in a name? *Gut* 2014;63:1525–6.
2. Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. *Am J Gastroenterol* 2006;101:385–98.
3. Knudsen AL, Bisgaard ML, Bulow S. Attenuated familial adenomatous polyposis (AFAP). a review of the literature. *Fam Cancer* 2003;2:43–55.
4. Sampson JR, Jones N. MUTYH-associated polyposis. *Best Pract Res Clin Gastroenterol* 2009;23:209–18.
5. Aretz S, Uhlhaas S, Goergens H, et al. MUTYH-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. *Int J Cancer* 2006;119:807–14.
6. Spier I, Horpaopan S, Vogt S, et al. Deep intronic APC mutations explain a substantial proportion of patients with familial or early-onset adenomatous polyposis. *Hum Mutat* 2012;33:1045–50.
7. Aretz S, Stienen D, Uhlhaas S, et al. 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.
8. Grover S, Kastrinos F, Steyerberg EW, et al. Prevalence and phenotypes of APC and MUTYH mutations in patients with multiple colorectal adenomas. *JAMA* 2012;308:485–92.
9. 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.
10. Filipe B, Baltazar C, Albuquerque C, et al. APC or MUTYH mutations account for the majority of clinically well-characterized families with FAP and AFAP phenotype and patients with more than 30 adenomas. *Clin Genet* 2009;76:242–55.
11. Steinke V, Holzapfel S, Loeffler M, et al. Evaluating the performance of clinical criteria for predicting mismatch repair gene mutations in lynch syndrome: a comprehensive analysis of 3,671 families. *Int J Cancer* 2013;3:28650.
12. Hes FJ, Ruano D, Nieuwenhuis M, et al. Colorectal cancer risk variants on 11q23 and 15q13 are associated with unexplained adenomatous polyposis. *J Med Genet* 2014;51:55–60.
13. Mongin C, Coulet F, Lefevre J, et al. Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists. *Clin Genet* 2012;81:38–46.
14. Thirlwell C, Howarth KM, Segditsas S, et al. Investigation of pathogenic mechanisms in multiple colorectal adenoma patients without germline APC or MYH/MUTYH mutations. *Br J Cancer* 2007;96:1729–34.
15. Renkonen ET, Nieminen P, Abdel-Rahman WM, et al. Adenomatous polyposis families that screen APC mutation-negative by conventional methods are genetically heterogeneous. *J Clin Oncol* 2005;23:5651–9.
16. Palles C, Cazier JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 2013;45:136–44.
17. Briggs S, Tomlinson I. Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. *J Pathol* 2013;230:148–53.
18. Valle L, Hernandez-Illan E, Bellido F, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Hum Mol Genet* 2014;18:18.
19. Yoshida R, Miyashita K, Inoue M, et al. Concurrent genetic alterations in DNA polymerase proofreading and mismatch repair in human colorectal cancer. *Eur J Hum Genet* 2011;19:320–5.
20. TCGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487:330–7.
21. Church DN, Briggs SE, Palles C, et al. DNA polymerase epsilon and delta exonuclease domain mutations in endometrial cancer. *Hum Mol Genet* 2013;22:2820–8.
22. Vasen HF, Mecklin JP, Khan PM, et al. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum* 1991;34:424–5.
23. Vasen HF, Watson P, Mecklin JP, et al. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by

the International Collaborative group on HNPCC. *Gastroenterology* 1999;116:1453–6.
24. Li H, Durbin R. Fast and accurate short read alignment with Burrows-wheeler transform. *Bioinformatics* 2009;25:1754–60.
25. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20: 1297–303.
26. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;25:2078–9.
27. Albers CA, Lunter G, MacArthur DG, et al. Dindel: accurate indel calls from short-read data. *Genome Res* 2011;21:961–73.
28. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004;11: 377–94.
29. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 2009; 106:19096–101.
30. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000;302: 205–17.
31. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* 2004;14:1188–90.
32. Swan MK, Johnson RE, Prakash L, Prakash S, Aggarwal AK. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat Struct Mol Biol* 2009;16:979–86.
33. Hogg M, Osterman P, Bylund GO, Ganai RA, Lundstrom EB, Sauer-Eriksson AE, Johansson E. Structural basis for processive DNA synthesis by yeast DNA polymerase varepsilon. *Nat Struct Mol Biol* 2014;21:49–55.
34. Vogt S, Jones N, Christian D, et al. Expanded extracolonic tumor spectrum in MUTYH-associated polyposis. *Gastroenterology* 2009;137: 1976–85.
35. Engel C, Loeffler M, Steinke V, et al. Risks of less common cancers in proven mutation carriers with lynch syndrome. *J Clin Oncol* 2012;30:4409–15.
36. National Comprehensive Cancer Network. Genetic/Familial High-Risk Assessment: Colorectal (Version 1.2014). Available at: http://www.nccn.org/professionals/physician_gls/pdf/genetics_colon.pdf. Last accessed on April 16, 2014.
37. Vasen HF, Moslein G, Alonso A, et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP). *Gut* 2008;57:704–13.
38. Dunlop MG, Dobbins SE, Farrington SM, et al. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet* 2012;44:770–6.
39. Ogi T, Limsirichaikul S, Overmeer RM, et al. Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol Cell* 2010;37:714–27.
40. Lemmens L, Urbach S, Prudent R, et al. Phosphorylation of the C subunit (p66) of human DNA polymerase delta. *Biochem Biophys Res Commun* 2008;367:264–70.
41. Smith CG, Naven M, Harris R, et al. Exome resequencing identifies potential tumor-suppressor genes that predispose to colorectal cancer. *Hum Mutat* 2013;34:1026–34.