ORIGINAL ARTICLE

Comprehensive analysis of the *MLH1* promoter region in 480 patients with colorectal cancer and 1150 controls reveals new variants including one with a heritable constitutional *MLH1* epimutation

Monika Morak,[1,2] Ayseguel Ibisler,[3] Gisela Keller,[4] Ellen Jessen,[5] Andreas Laner,[2] Daniela Gonzales-Fassrainer,[2] Melanie Locher,[2] Trisari Massdorf,[1] Anke M Nissen,[2] Anna Benet-Pagès,[2] Elke Holinski-Feder[1,2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2017-104744).

[1]Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, Munich, Bavaria, Germany
[2]Center of Medical Genetics, Medizinisch Genetisches Zentrum, Munich, Bavaria, Germany
[3]Department of Human Genetics, Ruhr University Bochum, Bochum, North Rhine-Westphalia, Germany
[4]Institute of Pathology, Technical University Munich, Munich, Bavaria, Germany
[5]Praxis für Humangenetik, Praxis für Humangenetik, Hamburg, Hamburg, Germany

Correspondence to
Professor Elke Holinski-Feder, Center of Medical Genetics, Medizinisch Genetische Zentrum, Munich 80335, Germany; elke.holinski-feder@mgz-muenchen.de

Received 17 April 2017
Revised 27 October 2017
Accepted 30 October 2017
Published Online First 22 February 2018

Check for updates

To cite: Morak M, Ibisler A, Keller G, *et al*. *J Med Genet* 2018;**55**:240–248.

ABSTRACT

Background Germline defects in *MLH1*, *MSH2*, *MSH6* and *PMS2* predisposing for Lynch syndrome (LS) are mainly based on sequence changes, whereas a constitutional epimutation of *MLH1*(CEM) is exceptionally rare. This abnormal *MLH1* promoter methylation is not hereditary when arising de novo, whereas a stably heritable and variant-induced CEM was described for one single allele. We searched for *MLH1* promoter variants causing a germline or somatic methylation induction or transcriptional repression.

Methods We analysed the *MLH1* promoter sequence in five different patient groups with colorectal cancer (CRC) (n=480) composed of patients with i) CEM (n=16), ii) unsolved loss of MLH1 expression in CRC (n=37), iii) CpG-island methylator-phenotype CRC (n=102), iv) patients with LS (n=83) and v) MLH1-proficient CRC (n=242) as controls. 1150 patients with non-LS tumours also served as controls to correctly judge the results.

Results We detected 10 rare *MLH1* promoter variants. One novel, complex *MLH1* variant c.-63_-58delins18 is present in a patient with CRC with CEM and his sister, both showing a complete allele-specific promoter methylation and transcriptional silencing. The other nine promoter variants detected in 17 individuals were not associated with methylation. For four of these, a normal, biallelic *MLH1* expression was found in the patients' cDNA.

Conclusion We report the second promoter variant stably inducing a hereditary CEM. Concerning the classification of promoter variants, we discuss contradictory results from the literature for two variants, describe classification discrepancies between existing rules for five variants, suggest the (re-)classification of five promoter variants to (likely) benign and regard four variants as functionally unclear.

INTRODUCTION

Tumours with high microsatellite instability and immunohistochemical (IHC) loss of DNA mismatch repair (MMR) protein expression are hallmarks of Lynch syndrome (LS) following an autosomal dominant inheritance mode.[1] The molecular basis of LS is a germline defect in one of the DNA MMR genes *MLH1*, *MSH2*, *MSH6* or *PMS2*, classically due to a nucleotide change such as single nucleotide variants, small insertions or deletions (indels) or larger single/multiple exon deletions.[2 3] In the rare condition of a constitutional epimutation of *MLH1* (CEM), an abnormal *MLH1* promoter methylation in all somatic tissues epigenetically causes a functional *MLH1* defect.[4–7] In addition to germline defects, the group of patients with MSI-H tumours lacking MLH1 protein staining also includes a substantial number of sporadic CRC cases showing CpG-island methylator-phenotype (CIMP) and at least partial biallelic *MLH1* promoter methylation in tumour tissue only.[8]

Aberrant CEM is classically found hemiallelic and conducts transcriptional silencing of *MLH1* and *EPM2AIP1*.[5 9–13] A 'primary' CEM is set-up de novo,[12 14–16] and is not heritable. The aberrant methylation is erased by the epigenetic reprogramming in germline formation,[17] as shown in spermatozoa of CEM carriers[7 13 16] and in family members without methylation on the same allele.[10 12] However, exceptional reports of CEM transmission to the next generation in single families might indicate an underlying genetic cause.[18 19] A 'secondary' CEM can be the consequence of a transcriptional repression, or might be induced by a variant in *cis*.[20] In two cases with genomic deletions including the first exon of *MLH1*,[21 22] an allele-specific methylation of the remaining *MLH1* promoter was found. We previously reported one case of a CEM associated with a large genomic duplication.[23] So far, one *MLH1* promoter variant c.-27C>A in *cis* to variant c.85G>T p.Ala29Ser has been reported for several patients with a CEM.[16 24 25] The allele with the variant showed an incomplete *MLH1* promoter methylation[16 24] and a reduced expression of *MLH1*. The mosaic CEM was reinstated on the variant allele in the next generation,[16 24] and in one family, the accumulation of CRC indicated a dominant trait of inheritance.[25] In a reporter assay, variant c.-27A was designated to be causative for the reduced expression.[16] Furthermore, promoter variants may also have regulatory effects without coincidence of methylation, as assumed, for example, for *MLH1* promoter variants c.-11C>T, c.-42C>T and c.-413_-411delGAG reducing the promoter activity in varying degrees in luciferase reporter assays,[25 26] and for variants c.-28A>G and c.-7C>T found in individuals with a partially reduced *MLH1* gene expression.[27]

Our aim was to investigate the presence and effect of promoter variants that might impair the normal *MLH1* gene function by either inducing a constitutional *MLH1* epimutation in 16 patients with CEM, or by reducing the transcriptional activity in 37 patients with CRC with unsolved MLH1 deficiency in their tumours (H1D). Furthermore, we searched for promoter variants in patients with CIMP tumours and controls. We sequenced the *MLH1* promoter region at least up to *MLH1* c.-667 in a total of 480 patients with CRC divided into five molecular subgroups including controls, and 1150 patients with tumours not associated with LS (nLS) as a control group.

MATERIALS AND METHODS

Recruited patients gave informed consent for the study approved by the ethics committee in Munich. DNA from peripheral blood cells was extracted with the FlexiGene DNA kit (QIAGEN), from buccal cells, normal colon tissue and microdissected colon cancer tissue, the QIAamp DNA Blood Mini kit (QIAGEN) was used. Analyses for germline variants and large deletions/duplications in the genes *MLH1* and *PMS2*, *MSH2*, *EPCAM* and *MSH6* were performed as described previously.[28 29]

We investigated 238 patients with CRC with MLH1-deficient tumours in IHC staining divided into subgroups: i) 16 patients with a CEM (thereof, 12 were published,[12] for details see online supplementary table 1), ii) 37 unsolved patients with MLH1-deficient tumour (H1D) and neither a germline variant in *MLH1* or *PMS2* nor *MLH1* methylation found in blood and tumour DNA, iii) 102 patients showing at least 50% *MLH1* promoter methylation in their tumours (CIMP) and iv) 83 patients with LS with a pathogenic *MLH1* germline variant (class 4 or 5 according to InSiGHT). As controls, we investigated a patient with CRC group V of 242 patients with positive protein staining for MLH1 in their tumours (C-H1P) and 1150 tumour patients not suspicious of having LS (C-nLS) (see table 1).

The promoter analysis was performed by Sanger sequencing from *MLH1* c.-667 to c.116+40 (g.37034372–37035194) as described[12] and was extended for the CEM carriers to a region 5 kb upstream of *MLH1* by Long-Range PCR (TAKARA) to cover also potentially regulatory regions further upstream as the promoter region is not clearly defined. The controls were analysed by next-generation sequencing using the TruSight Rapid Capture and TruSight Cancer Sequencing Panel (Illumina) covering the *MLH1* promoter until c.-667. With MS-MLPA kit ME011 (MRC Holland), the *MLH1* promoter region from c.-659 to c.116+90 was tested for larger genomic deletions, duplications and for methylation. Sodium bisulfite treatment of genomic DNA, methylation-specific PCR amplification of two overlapping fragments in the *MLH1* promoter region from c.-362 to c.-193 and from c.-286 to c.17 spanning 22 CpG dinucleotides and sequencing was performed as published.[12]

For cDNA analyses, total RNA was extracted from peripheral blood cells by the PAXGene Blood RNA and Preparation kit (PreAnalytix), and from lymphocytes cultured after Ficoll separation with and without puromycin incubation to check nonsense-mediated mRNA decay. The cDNAs were generated with iScript select cDNA-Synthesis kit (Bio-Rad) using an oligo(dT)$_{18}$ primer. Biallelic expression of genomically heterozygous variants was investigated for *EPM2AIP1, MLH1* and *LRRFIP2* by PCR amplification followed by digestion with Exo-SAP kit (USB) and Sanger sequencing with Big Dye V.1.1 (Applied Biosystems) on ABI PRISM 3100 Avant using additional primers for sequencing, as we described.[23] The longer transcript of *MLH1* was amplified from c.-148 or c.-113 to c.883 with primers spanning the 5'UTR (untranslated region) to exon 10 by standard procedures with LongAmp Taq (NEB) as described.[23] For cDNA analysis of *EPM2AIP1*, fragments were amplified from c.-84 or c.-227 to c.197, or within the 3'UTR from c.*2470 to c.*2630 using Ampli-Taq Gold (ABI) at standard procedures.[30] In parallel, genomic contamination in cDNA was ruled out by PCR with primers in *MLH1* exon 7 forward and eight reverse spanning a small genomic intron and analysis on a 1% agarose gel, as otherwise, cDNA analysis would be invalid for *EPM2AIP1* due to a lack of introns. The transcript of *LRRFIP2* was amplified from c.1988 to c.*300 using Ampli-Taq Gold (ABI) at standard procedures.[30] By using informative variants we investigated the allelic distribution of *MLH1, EPM2AIP1*

Table 1 Patients groups and results

Group	I) CEM	II) H1D	III) CIMP	IV) LS	V) C-H1P	Control C-nLS
Number of patients	16	37	102	83	242	1150
MSI	MSI-H	MSI-H	MSI-H	MSI-H	106 MSI-H/136 MSS	n.a.
IHC MLH1	neg.	neg.	neg.	neg.	pos.	n.a.
MLH1 germline variant	neg.	neg.	neg.	pos.	183 neg./59 n.a.	n.a.
MLH1 CEM	pos. 50%	neg.	neg.	neg.	106 neg./136 n.a.	n.a.
MLH1 tumour methylation	pos. 50%	neg.	pos.	neg.	n.a.	n.a.
BRAF mutation	neg.	neg.	pos. neg.	neg./n.a.	n.a.	n.a.
CRC	Yes	Yes	Yes	Yes	Yes	No
Rare *MLH1* promoter variants	2 (12.5%): c.-63_-58delins18; c.-269C>G	3 (8.1%): c.-42C>T; c.-269C>G; c.-477T>C	2 (2%): c.-269C>G; c.-369A>G	1 (1.2%): c.-33T>G	8 (7.6%): c.[-28A>G;-7C>T]; c.-28A>G; c.-230G>C; c.-269C>G 4x; c.-593G>C	3 (0.3%): c.[-28A>G;-7C>T] 2x; c.-269C>G

Categorisation of 1630 patients into different groups by molecular characteristics including the status of microsatellite instability (MSI), immunohistochemical staining (IHC) of MLH1 in the tumour (positive: pos., negative: neg.), *MLH1* germline variants, methylation of the *MLH1* promoter in blood (CEM) and in tumour, *BRAF* mutation status in NM_004333.4 c.1799T>A p.Val600Glu in the tumour and diagnosis of colorectal cancer (CRC). Not analysed: n.a. 480 patients with CRC were subdivided into groups I–V, of those, I–IV had MLH1-deficient tumours of different causes: I) constitutional *MLH1* epimutation (CEM), II) unsolved MLH1-deficiency in the tumour (H1D), III) CIMP tumours, IV) patients with Lynch syndrome (LS) with pathogenic *MLH1* germline variants (class 4 or 5 according to InSiGHT). Group V consists of 242 patients with MLH1-proficient tumours (C-H1P) and served as a control group. The second control group (C-nLS) comprises patients with tumours indicating other syndromes, but not LS. The number of rare promoter variants detected in each group (and their percentage in brackets) is given and variant nomenclature is provided in relation to the *MLH1* translation start.

Table 2 Data of all MLH1 promoter variants

Nomenclature	Detected in category	rs	Allelic frequency	Classification	Prediction in Alamut, conservation, TFBS	cDNA expression	Our classifications due to our results and data from literature	Remark
MLH1 c.-63_−58delinsCACGAGGCACGAGCACGA	1x CEM	Novel	Not in ExAC or 1000 Genomes	LOVD NCs	CiV highly positive, nine TFBS lost	*MLH1* and *EPM2AIP1* monoallelic, *LRRFIP2* biallelic	Four by ACMG: PM2+PS3, three in InSiGHT and Liu *et al*[45] (lacking in vitro functional assay/segregation with disease), Liu *et al*: secondary epimutation suspected – *MLH1* and *EPM2AIP1* allelic loss in expression, absence in controls (cosegregation of CEM and promoter variant in two family members is not a criterion yet, and only one MSI-H, MLH1-deficient CRC to reach class 4)	Δ
EPM2AIP1 c.7A>G p.Met3Val (*MLH1* c.-477T>C)	1x H1D	rs 746415556	ExAC: 0.000009, in East Asia 0.00012, no homozygotes; not in 1000 Genomes	LOVD NC	Benign, splice-neutral, CiV positive, two TFBS lost	*MLH1* n.i., *EPM2AIP1* biallelic	Three in InSiGHT, ACMG and Liu *et al* -No data	
EPM2AIP1 c.-102T>C (*MLH1* c.-369A>G)	1x CIMP	Novel	Not in ExAC or 1000 Genomes	LOVD NC	No CiV, one TBFS lost	no cDNA	Three in InSiGHT, ACMG and Liu *et al* -No data	
MLH1 c.-230G>C	1x C-H1P	rs 587782631	Not in ExAC or 1000 Genomes	LOVD NC, class 3 in ClinVar	No CiV, one TBFS lost	no cDNA	Three in InSiGHT, ACMG and Liu *et al* –No data (one H1P case)	
MLH1 c.-33T>G	1x LS	rs 201247839	ExAC: 0.00025, no homozygotes; 1000 Genomes: 0.0002	LOVD NC	Splice-neutral, CiV negative, one TBFS lost	no cDNA	Three in InSiGHT and Liu *et al* (allelic phase of pathogenic variant unknown), ACMG: BP2+BP5+BS3=1 (allelic phase not needed), Liu *et al*: consult InSiGHT database –No data, patient with LS with additional pathogenic *MLH1* variant and allelic phase unknown, AF	Δ
MLH1 c.-42C>T	1x H1D	rs 41285097	ExAC: 0.000025, no homozygotes; 1000 Genomes: 0.0002	LOVD class 3	CiV highly positive, four TFBS lost	*MLH1* and *EPM2AIP1* biallelic	Two in InSiGHT and Liu *et al* (additional argument lacking; 3 CRC MSS/lack of cosegregation), ACMG: BS1+BS3=1 - AF, *MLH1* cDNA no functional defect in our case, but contradictory results: in literature 37% reduced expression in luciferase essay (cosegregation with late-onset CRC)[23 25 26]	Δ ↔ R?
EPM2AIP1 c.123C>G p.= (*MLH1* c.-593G>C)	1x C-H1P	rs 34566456	ExAC: 0.0052, Africa 0.058, 18 homozygotes; 1000 genomes: 0.02, Africa 0.08	LOVD NC	Splice-neutral, CiV negative, no TFBS lost	*MLH1* n.i., *EPM2AIP1* biallelic	Two in InSiGHT and Liu *et al* (founder mutation not excluded, or additional argument in combination with cDNA no functional defect), ACMG: BS1+BS2=1, - AF in Africa, 18 homozygotes	Δ
MLH1 c.-28A>G	1x C-H1P	rs 56198082	ExAC: 0.002, in Finland 0.009, one homozygote; 1000 Genomes: 0.0008	LOVD class 3, also in ClinVar	CiV negative, one TBFS lost	*MLH1* biallelic, n.i. *EPM2AIP1*	One in InSiGHT, ACMG: BS1+BS2+BS3, and Liu *et al* – AF, one homozygote, in >3 cases with H1P, *MLH1* cDNA no functional defect[23 48 49]	R
MLH1 c.-7C>T for variant c.[-28A>G;-7C>T]	1x C-H1P, 2x C-nLS	rs 104894994	ExAC: 0.0015, in Finland 0.0087, one homozygote; 1000 Genomes: 0.0004	LOVD class 3, also in ClinVar	No CiF, two TBFS lost	*MLH1* and *EPM2AIP1* biallelic	One in InSiGHT, ACMG: BS1+BS2+BS3, and Liu *et al* – AF, one homozygote, in >3 cases with H1P, *MLH1* cDNA no functional defect in our case, but in literature reduced expression of the variant allele to 28%–33% in cDNA without other variants detectable, in vitro assays performed[27 50]	↔ R?
EPM2AIP1 c.-202G>C (*MLH1* c.-269C>G)	1x CEM, 1x H1D, 1x CIMP, 4x C-H1P, 1x C-nLS	rs 35032294	1000 Genomes: 0.002	LOVD class 2, also in ClinVar, benign in Invitae	CiV negative, one TBFS lost	*MLH1* and *EPM2AIP1* biallelic	One in InSiGHT and ACMG: BS1+BS3, but Liu *et al*: secondary epimutation suspected (due to 1 CEM case with this variant) – AF, in >3 cases with H1P, *MLH1* cDNA no functional defect, confirmed by Zavodna *et al*[51]	Δ R

Summary of data of all *MLH1* promoter variants detected in our study including category of patient or control in which the variant was detected, rs if known, the allelic frequency (AF) in ExAC or 1000 Genomes database, existence of homozygotes, variant classification in InSiGHT-LOVD or other database, in silico predictions in Alamut including splicing, the level of nucleotide conservation in vertebrates (CiV) from UCSC and loss of transcription factor binding sites (TFBS) predicted by ALGGEN-PROMO. The expression was investigated for all variants with cDNA available, and 'biallelic' indicates that the transcript showed a heterozygous variant in the transcript with normal allelic distribution, n.i. indicates cDNA analysis was 'not informative' due to a lack of heterozygous variants in the transcript. All variants are regarded as class 3 by default. Our classification of the variants applying the InSiGHT[43] and ACMG[44] guidelines and proposal by Liu *et al* [45] due to our results are given including the arguments and literature. ACMG arguments in detail were: PS3: abrogated mRNA expression, PM2: absence in controls, BS1: allele frequency higher than expected for disorder, BS2: observed in a health adult, BS3: in vivo functional studies (here: cDNA analyses) show no damaging effect on splicing, BP2: observed in *trans* or in *cis* with a pathogenic variant in any inheritance pattern, BP5: variant found in a case with an alternate molecular basis for disease (= control group). In brackets, additional information not sufficient as arguments are provided. In the remark, a discordance in classification between InSiGHT and ACMG is depicted by Δ, contradictory results in literature are marked with ↔ and suggested reclassification by R.

CRC, colorectal cancer; CEM, constitutional *MLH1* epimutation; H1P, MLH1-proficient colorectal cancer.

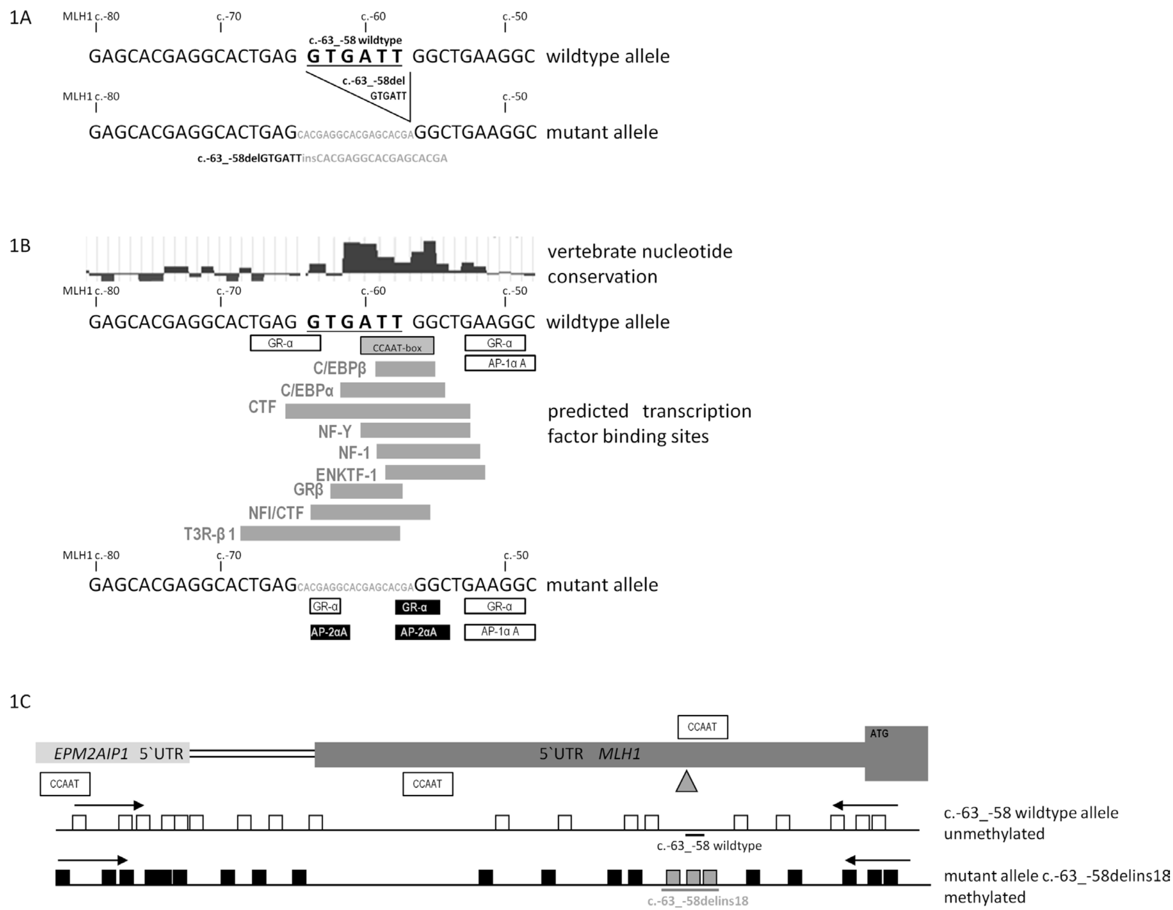

Figure 1 (A) Schematic diagram of the genomic alleles of the *MLH1* promoter in patient CEM-15 harbouring the complex heterozygous variant c.-63_-58 delGTGATTinsCACGAGGCACGAGCACGA or c.-63_-58delins18 in the 5'UTR of the *MLH1* transcript. The wild-type allele above shows six bases in bold, which are deleted in the mutant allele below. Instead, the mutant allele contains an insertion of 18 nucleotides (written in grey). (B) Diagram of vertebrate nucleotide conservation (UCSC Genome Browser), predicted transcription factor binding sites (TFBS, by ALGGEN-PROMO), a CCAAT-box in the complementary reverse strand of the wild-type allele in bold and changes in the mutant allele (written in grey) below. Grey filled boxes indicate lost TFBS due to the variant, white boxes show preserved TFBS and black boxes depict newly generated TFBS in the mutated allele. (C) Schematic illustration of bisulfite sequences. The germline promoter methylation of 50% was also investigated with methylation-specific primers in bisulfite-converted DNA of the patient. Selection for unmethylated alleles (open boxes) presented only the wild-type allele (underlined in black). Sequencing of the methylated fragments below detected complete methylation (filled boxes) in all 15 CpG dinucleotides analysed in the fragment, which specifically show only the variant allele c.-63_-58delins18 (depicted and underlined in grey).

and *LRRFIP2* transcripts in presence of the promoter variant. Primer sequences are available on request.

For sequence analysis, the Mutation Surveyor V.3.1 (SoftGenetics) software was used. For annotation we refer to the RefSeq transcripts NM_000249.3, NG_007109.2 for *MLH1*, NM_014805.3, NG_008418.1 for *EPM2AIP1* and NM_006309.3, NC_000003.11 for *LRRFIP2* on Chr.3 (GRCh37); nomenclature is given according to HGVS standard recommendations V.2016 (http://varnomen.hgvs.org/recommendations)[31] referring to the genomic positions in hg19. Alamut V.2.6.1 was used for variant interpretation, as well as allelic frequencies in different populations (ExAC browser and 1000 Genomes). The evolutionary nucleotide conservation in vertebrates was derived from UCSC Genome Browser. For the prediction of transcription factor binding sites abolished by promoter variants, we applied the ALGGEN-PROMO tool V.3.0[32] with preselection to only human factors and human sites in a 'Search Promotor Sites' mode using standard parameters for sequences including 10 nucleotides around the promoter variant and compared the wild-type with the variant. Additionally, the generation of new translational start codons by promoter variants was ruled out. The *MLH1* promoter variants identified have been submitted to the InSiGHT MMR gene variant database LOVD3 (http://www.insight-group.org/variants/database/).[3]

RESULTS

By sequencing the *MLH1* promoter in a total of 1630 individuals including patients with CRC and controls, we detected 10 different rare *MLH1* promoter variants (table 1). The heterozygous allelic presence of the frequent variant c.-93G>A oscillated between 33% and 51% depending on the group analysed.

Promoter variants in patients with constitutional *MLH1* epimutation

Within the 16 patients with LS due to a constitutional *MLH1* epimutation (CEM), we detected two rare *MLH1* promoter variants: the novel, complex variant c.-63_-58delGTGATTinsCACGAGGCACGAGCACGA (from now on referred to as c.-63_-58delins18) was found in patient CEM-15 (see figure 1A and online supplementary figure 1A), and variant c.-269C>G in patient CEM-6 (see table 1 and online supplementary table

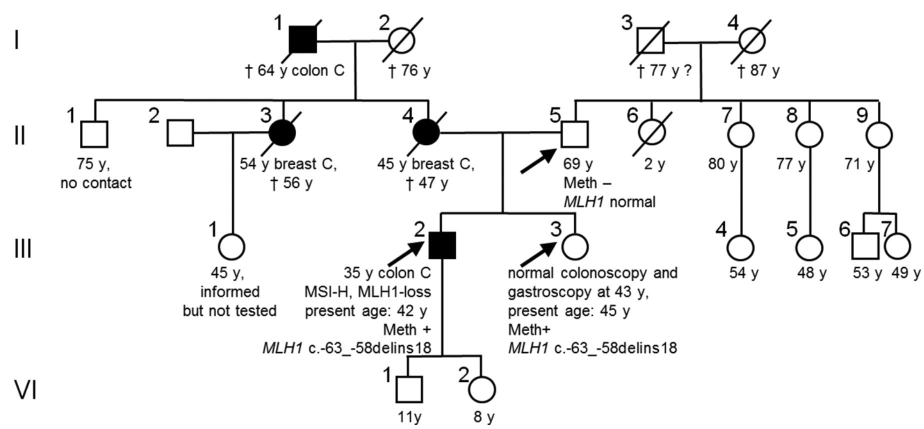

Figure 2 Family pedigree of patient CEM-15. Tumours (C=cancer) are given at age of diagnosis in years (y), and age at death is indicated by †. Blood samples were available from the index patient III-2, his sister and his father (indicated by arrows). The constitutional *MLH1* epimutation (Meth+) is linked to the *MLH1* variant c.-63_–58delins18 and is present in the index patient and the sister, whereas the father reveals the normal promoter sequence and no methylation (Meth-).

1, and details in table 2). Variant c.-269C>G was also found in seven further patients without CEM (see below). The common variant c.-93G>A was present heterozygously in eight patients with CEM, and homozygously in one patient with CEM. By MS-MLPA, no CNV was detected in the promoter region for 14 CEM cases, 1 large duplication was described previously[23] and 1 case could not be analysed.

For patient CEM-15, the *MLH1* promoter methylation of approximately 50% of alleles in all CpG dinucleotides investigated by MS-MLPA was found in peripheral blood cells, buccal cells, normal colon tissue and in his colon adenocarcinoma diagnosed at the age of 35 years, as well as in peripheral blood cells and buccal cells of his sister. The father was tested negative for both CEM and the promoter variant in peripheral blood, and no tumour diagnoses were reported in the paternal line of the family (for pedigree and clinical data see figure 2). No DNA was available from the mother who died at the age of 47 years after diagnosis with breast cancer, the maternal aunt with breast cancer diagnosed at the age of 54 years or the maternal grandfather who died from colon cancer at the age of 64 years.

No segregation analysis was possible for patient CEM-6 with *MLH1* promoter variant c.-269C>G, and nine other CEM cases. With the segregation and transmission analyses of the four other CEM index patients (details in online supplementary table 1), we found a de novo methylation of the maternal allele in two families, and observed the erasure of the CEM from the paternal allele when transmitted to their sons in two other families. CEM was previously shown to be inherited in one family in association with a large duplication.[23]

After bisulfite conversion and amplification specifically for methylated alleles, the patient CEM-15 and his sister showed a 100% complete methylation of all 15 CpG dinucleotides covered, which were specific for the *MLH1* variant allele c.-63_-58delins18 (schematic figure 1C, sequence electropherograms in online supplementary figure 2B). In the PCR for unmethylated alleles of patient CEM-15 and his sister, only the *MLH1* promoter wild-type allele was represented (see online supplementary figure 2C). In the father with wild-type at c.-63_-58, we could amplify two unmethylated alleles, as these were informative for the heterozygous promoter variant c.-93G>A (see online supplementary figure 2C). Bisulfite sequencing in patient CEM-6 revealed complete methylation in all CpG dinucleotides analysed, which was specific for the variant allele *MLH1* c.-269G, whereas the wild-type allele c.-269C was unmethylated.

We performed cDNA analyses for *MLH1, EPM2AIP1* and *LRRFIP2* transcripts from RNA isolated from PAXgene and short-term lymphocyte cultures with and without puromycine incubation prior to RNA-isolation from patient CEM-15 and his sister. For *MLH1* only the wild-type allele was expressed (see online supplementary figure 1C), whereas the variant c.-63_-58delins18 located in the 5'UTR of the *MLH1* transcript was not detectable in any cDNA of the two siblings. *EPM2AIP1* also showed a monoallelic expression of only the T allele in cDNAs of both siblings (see online supplementary figure 1C) in the genomically heterozygous variant c.*2570G>T in the 3'UTR. For *LRRFIP2,* a biallelic expression was found in patient CEM-15 and his sister by analysing the heterozygous variant c.*272G>A in the 3'UTR. For patient CEM-6 harbouring variant *MLH1* c.-269C>G, no RNA was available, but cDNA analyses were performed in another person with this variant diagnosed with a CIMP tumour (results in section: "Promoter variants in other patients with CRC and controls" and table 2).

Promoter variants in other patients with CRC and controls

Sequencing the *MLH1* promoter in 1614 further individuals, we identified nine rare *MLH1* promoter variants in 17 cases: c.-593G>C, c.-477T>C, c.-369A>G, c.-269C>G, c.-230G>C, c.-42C>T, c.-33T>G, c.-28A>G, c.[-28A>G;-7C>T] (table 1, details in online supplementary figure 1A). These variants were found in three patients with unsolved MLH1 deficiency (H1D), one patient with a pathogenic *MLH1* germline variant c.793C>T (LS), two CIMP cases, eight patients with MLH1-proficient CRC (C-H1P) and three controls without LS-tumours (C-nLS) (see table 2). Four of these promoter variants were previously listed in LOVD, and assigned *MLH1* c.-42C>T, c.-28A>G and c.-7C>T as class 3, and c.-269C>G as class 2 (table 2). The allele c.[-28A>G;-7C>T] was represented in three control cases. Methylation in the *MLH1* promoter was absent in blood DNA of all 17 patients with one of the nine rare *MLH1* promoter variants. To rule out tissue-specific methylation set-up in colon, colonic normal mucosa was investigated and found methylation-negative for all six variants c.-477T>C, c.-369A>G, c.-269C>G, c.-230G>C, c.-42C>T and c.-33T>G in seven cases with tissue available.

To investigate the effect of promoter variants on mRNA generation on the respective allele in vivo, we analysed the allelic balance of genomically heterozygous variants in the transcripts of *MLH1* and *EPM2AIP1* by sequencing. For six of these promoter variants RNA of patients with at least one informative variant in *MLH1* or *EPM2AIP1* could be obtained (see online supplementary figure 1C). A biallelic expresssion of *EPM2AIP1* was found for *MLH1* promoter variants c.-593G>C and c.-477T>C, which are both located in the coding region of *EPM2AIP1*, whereas *MLH1* was not analysable due to a lack of informative variants. A biallelic expression was found in cDNA analyses for *MLH1* promoter variants c.-269C>G c.-42C>T and c.[-28A>G;-7C>T] for both *EPM2AIP1* and *MLH1* in all cDNAs, and for variant c.-28A>G without c.-7C>T in *cis* for *MLH1*, without informative variants for *EPM2AIP1*. For the frequent promoter variant c.-93G>A, we performed cDNA analyses of three homozygote and 15 heterozygote controls, and found a biallelic expression of both transcripts *EPM2AIP1* and *MLH1* in all patient cDNAs. No RNA was available from patients with *MLH1* promoter variants c.-369A>G, c.-230G>C and c.-33T>G.

DISCUSSION

MLH1 and *EPM2AIP1* are bidirectionally paired genes with a shared promoter region. The core promoter region for *MLH1* has been defined from c.-184 to c.-132,[33] whereas additional *cis* elements and essential protein binding sites were defined from c.-301 to c.-76.[34] The transcriptional activity of the *MLH1* promoter strongly depends on two CCAAT boxes located in c.-278_-282 and c.-145_-141 from the *MLH1* translation start,[34–36] for *EPM2AIP1* a complementary reverse CCAAT box is located in the *MLH1* 5'UTR at c.-56_-60. Furthermore, the regions from c.-250 to c.-151 bp[36 37] and from c.-273 to c.-4[38] are described as critical for the regulation of *MLH1* transcription. Promoter variants have the potential to reduce or abrogate the transcriptional activity either with[16 24 25] or without[25] an allele-specific promoter methylation, and might be tissue-specific.

We searched for causative *MLH1* promoter changes within at least *MLH1* c.-667 and c.116+40 in 16 patients with CEM and 37 patients with unsolved H1D. To better judge the variants, we also analysed 102 patients with CIMP tumours, 83 patients with LS, 242 patients with H1P and 1150 patients with non-LS tumours (nLS). We detected a total of 10 different rare *MLH1* promoter variants in 1630 individuals (table 1). One variant is associated with a heritable CEM. Six of our 10 variants were not listed in LOVD before, and 2 were regarded as novel. The incidence rates of promoter variants differed between 12.5% and 0.3%, but the association of a variant with CEM or unsolved MLH1-deficient patients with CRC did not reach a statistical significance due to small case numbers.

Promoter variants associated with an epimutation

A CEM is regarded not to be heritable for cases in which the methylation was set-up de novo by chance. However, *cis*-acting germline variants may induce a stably inherited CEM. So far, only one *MLH1* allele c.-27C>A in *cis* with c.85G>T was found in several families association with a heritable mosaic CEM and reduced transcriptional expression.[16 24 25] In our 16 patients with CEM, we sequenced the *MLH1* promoter and regulatory region up to 5 kb upstream of the *MLH1* transcription start and identified two rare variants in one patient each.

We report the novel *MLH1* promoter variant c.-63_-58delins18 in 1 of 16 patients with CEM, which is for the first time associated with a complete promoter methylation in tissues from ectodermal and mesodermal germ layers and transcriptional silencing of the variant allele for *MLH1* and *EPM2AIP1* in the index patient and his sister. This variant was probably inherited from the mother who died of breast cancer and had a family history for breast and colon cancer. We have evidence that *MLH1* promoter variant c.-63_-58delins18 causes a secondary CEM following a stable, autosomal-dominant inheritance, and offered predictive testing for related family members at the age of majority.

So far, only one *MLH1* promoter variant c.-27C>A[16 24 25 27] and one large genomic duplication[23] were reported to induce a heritable, secondary CEM, but both were associated with mosaic methylation.

The mechanism causing the heritable CEM in our patient remains unclear, as no in vitro assays have been performed. Methylation can be induced as a consequence of transcriptional silencing, as shown for *MSH2* in patients with *EPCAM* deletions,[39] and for genomic deletions including *MLH1* exon 1.[21 22] Taking into account the high evolutionary conservation, the predicted loss of nine TFBS including NF-Y in a critical regulatory region,[38] and the loss of a CCAAT-box consensus sequence (see online supplementary figure 1B), the transcriptional silencing of either *MLH1* or *EPM2AIP1* by regulatory effects might have induced methylation of the shared promoter region as a secondary consequence. Alternatively, a variant-directed methylation of the DNA could be hypothesised,[20] or a variant-specific change of the histone modifications compacting the chromatin conformation could be taken into account.[40 41] However, this effect is limited to *MLH1* and *EPM2AIP1* as reflected by a monoallelic expression in cDNA analyses, while *LRRFIP1* downstream of *MLH1* shows a biallelic expression in our patient.

For another epimutation carrier, we detected the *MLH1* variant c.-269C>G. Even though the variant allele was methylated here, in another case in literature the wild-type allele was methylated,[25] arguing against a variant-specific CEM induction. Furthermore, we detected c.-269C>G also in seven individuals without CEM.

The heritability of CEM was not investigable for this and nine other CEM cases. In four families a primary constitutional *MLH1* methylation is suspected, as a de novo set-up of methylation or an erasure of methylation in children could be demonstrated. For one family a secondary CEM in combination with a large duplication was published previously,[23] whereas no CNV was detected in 14 CEM cases.

Promoter variants without epimutation

MLH1 promoter variants may have the potential to abrogate TFBS, generate transcriptional repressors or change the chromatin status and by these means have an impact on the transcription, but do not necessarily induce methylation. In literature, *MLH1* promoter variants c.-11C>T, c.-27C>A, c.-42C>T, c.-413_-411delGAG and c.-435_-432delAAAG were reported in patients without CEM, but alleles c.-11T, c.-27A and c.-42T significantly reduced the promoter activity in luciferase assays.[25]

In 17 individuals without CEM, we detected 9 *MLH1* promoter variants (c.-593G>C, c.-477T>C, c.-369A>G, c.-269C>G, c.-230G>C, c.-42C>T, c.-33T>G, c.-28A>G and c.[-28A>G;-7C>T]) (tables 1 and 2), which seem to be quite infrequent in LOVD and literature.[3 42] We regard the biallelic gene expression of *MLH1* but not *EPM2AIP1* as a significant argument for the classification of promoter variants in patients suspected of having LS. The *MLH1* cDNA analyses informative for four variants showed no evidence of a reduced transcriptional activity.

Promoter variant prediction, classification and interpretation

So far, no specific rules for the classification of promoter variants have been described by the current InSiGHT or ACMG scheme, neither in presence nor in absence of CEM,[43 44] but the necessity is already underscored in a pioneer publication.[45] Only with criteria such as a high population frequency and/or homozygous state in healthy controls it is possible to reach a benign or likely benign classification, as it was the case for our *MLH1* promoter variant c.-593G>C. The exception is one promoter variant in *MSH2* c.-78_-77delGT, which was graded as class 4—likely pathogenic in LOVD.[46] For the classification of our 10 rare promoter variants identified, we apply the five-tiered InSiGHT scheme,[43] ACMG guidelines[44] and the guide proposed by Liu *et al*[45] (results compiled in table 2), and suggest amendments specific for promoter variants. The loss of TFBS and the nucleotide conservation in vertebrates (depicted in online supplementary figure 1B) are listed in table 2, but have no consented predictive value. We therefore only used our cDNA results, allelic frequency and presence in patients with CRC with or without MMR defect as criteria for the classification.

In our attempt to classify the *MLH1* promoter variant c.-63_-58delins18 we detected in two siblings with an epimutation, we would reach class 4—likely pathogenic with the ACMG criteria, but only class 3—uncertain significance as CEM is not included in the current InSiGHT classification rules and Liu *et al*[45] due to the lack of functional tests (table 2). This outstanding work addressing the problem of promoter variant classification and interpretation is regarding a CEM as secondary and heritable if a promoter variant is detected.[45] We nevertheless suggest to investigate the heredity of CEM for each case, also in absence of promoter variants. For a class 4 classification of a promoter variant like ours associated with a CEM, we would recommend to add to the InSiGHT guidelines: 'In constitutional promoter methylation carriers, the allele with the promoter variant has to be proven to be methylated allele and/or an allele-specific transcriptional silencing has to be demonstrated and/or a segregation with the promoter variant and a constitutional *MLH1* promoter methylation in a family is shown'. To reach a class 5 for promotor variants, we would agree with Liu *et al*[45] to add: 'For regulatory defects, the allele-specific silencing has to be confirmed in an in vitro functional assay'.

The nine further *MLH1* promoter variants c.-593G>C, c.-477T>C, c.-369A>G, c.-269C>G, c.-230G>C, c.-42C>T, c.-33T>G, c.-28A>G and c.[-28A>G;-7C>T] were not causatively associated with CEM in 17 individuals. The AF of variant c.-593G>C in Africans allowed a classification to class 2—likely benign. With the intention to classify the other promoter variants, we interpreted the biallelic cDNA expression of *MLH1* equivalent to the InSiGHT argument 'with no associated mRNA aberration' (table 2), but to reach a class 2, an additional argument is needed. Alternatively, it could be discussed whether promoter variants can be put equal to synonymous substitutions, as both do not change the coding transcript, and reach class 2 only by showing absence of mRNA aberration. However, Liu *et al* suggest an additional criterion in combination with 'absence of allelic loss in vivo',[45] which we strongly support to reach class 1.

Based on our normal cDNA results and presence in controls with MLH1-proficient CRC, we would suggest a reclassification from class 3 to class 1 for the three *MLH1* promoter variants (c.-269C>G, c.-28A>G and c.[-28A>G;-7C>T]), and to class 2 for one variant (c.-42C>T). Four variants (c.-477T>C, c.-369A>G, c.-230G>C and c.-33T>G) with insufficient data remain class 3—of uncertain significance.

The classification of our promoter variants gave divergent results between InSiGHT and ACMG for five variants (table 2). Our impression is that with the ACMG rules a meaningful classification can be reached more easily—and might be subject to revision, whereas the InSiGHT guidelines require more evidence and are quite definite.

For the *MLH1* promoter variants c.-42C>T and c.[-28A>G;-7C>T], contradictory findings are reported in the literature.[26 27] We observed a normal biallelic representation of heterozygous variants in cDNA by PCR and Sanger sequencing, which is not a quantitative method, but is capable to show major allelic imbalances, as well as allelic losses.[23] We cannot explain the conflicting results reported as partial allelic imbalance and reduced expression in luciferase assays,[25–27] which also exist for variant c.-93,[47] and might be attributable to different isoforms, or additional pathogenic variants. So far, no clear procedure is provided for contradictory results by the InSiGHT classification rules, and a threshold definition for a reduced promoter activity in functional assays or a reduced cDNA expression and their interpretation is also claimed by Liu *et al*.[45] Furthermore, the necessity to investigate expression analyses in colon mucosa has to be discussed on international level, as this is one of the target tissues for LS, and might also attribute for splicing analyses of variants in general.

To sum up, we describe 10 different rare *MLH1* promoter variants. The novel *MLH1* promoter variant c.-63_-58delins18 was found in one of our 16 epimutation carriers. We report the second *MLH1* promoter variant associated with a secondary, heritable CEM in in two siblings, and the first variant showing full methylation and complete transcriptional silencing. Promoter variant c.-27C>A was previously reported in association with an incomplete, mosaic promoter methylation and a reduced *MLH1* gene expression.[16 24 27] For nine other *MLH1* promoter variants identified in patients and controls, no variant-specific promoter methylation was detected, and in informative cases a normal *MLH1* transcription was found for four variants.

For variant c.-63_-58delins18 associated with a heritable CEM, we would suggest class 4, assign four variants to class 3, and based on our results we would (re-)classify five *MLH1* promoter variants c.-593G>C, c.-269C>G, c.-42C>T, c.-28A>G and c.[-28A>G;-7C>T] as class 1 or 2.

Variants with an impact on transcription or inducing a stably heritable CEM are only rarely identified in the *MLH1* promoter region. However, internationally approved rules are needed for a standardised classification of MMR promoter variants, which have to be discussed and amended on international level by the InSiGHT interpretation committee.

Acknowledgements The authors thank the German Cancer Aid (Deutsche Krebshilfe) and the Wilhelm Sander-Stiftung for their support of this work. The authors also thank all patients for their participation in this study, as well as their respective doctors for contributing materials and clinical information.

Contributors EH-F and MM designed the study, wrote the manuscript and are responsible for the content of the study. EH-F, AI, GK, EJ and DG-F provided patients clinical and molecular data and samples. Experiments were mainly performed by TM, supervised by MM; tumour analyses were also carried out by ML and AL. AB-P was responsible for NGS techniques and supervised AMN for bioinformatical analyses. EH-F, MM, AL, AB-P and AN were involved in data analyses and interpretation. EH-F and MM wrote the manuscript and are responsible for the content of the study. MM submitted the article.

Funding This work was supported by grants from the German Cancer Aid (Deutsche Krebshilfe) (#111222) and the Wilhelm Sander-Stiftung (#2012.081.1).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Ethikkommission der Medizinischen Fakultät der LMU München.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* 1999;36:801–18.
2. Mangold E, Pagenstecher C, Friedl W, Mathiak M, Buettner R, Engel C, Loeffler M, Holinski-Feder E, Müller-Koch Y, Keller G, Schackert HK, Krüger S, Goecke T, Moeslein G, Kloor M, Gebert J, Kunstmann E, Schulmann K, Rüschoff J, Propping P. Spectrum and frequencies of mutations in MSH2 and MLH1 identified in 1,721 German families suspected of hereditary nonpolyposis colorectal cancer. *Int J Cancer* 2005;116:692–702.
3. Plazzer JP, Sijmons RH, Woods MO, Peltomäki P, Thompson B, Den Dunnen JT, Macrae F. The InSiGHT database: utilizing 100 years of insights into Lynch syndrome. *Fam Cancer* 2013;12:175–80.
4. Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res* 2002;62:3925–8.
5. Hitchins M, Williams R, Cheong K, Halani N, Lin VA, Packham D, Ku S, Buckle A, Hawkins N, Burn J, Gallinger S, Goldblatt J, Kirk J, Tomlinson I, Scott R, Spigelman A, Suter C, Martin D, Suthers G, Ward R. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. *Gastroenterology* 2005;129:1392–9.
6. Miyakura Y, Sugano K, Akasu T, Yoshida T, Maekawa M, Saitoh S, Sasaki H, Nomizu T, Konishi F, Fujita S, Moriya Y, Nagai H. Extensive but hemiallelic methylation of the hMLH1 promoter region in early-onset sporadic colon cancers with microsatellite instability. *Clin Gastroenterol Hepatol* 2004;2:147–56.
7. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. *Nat Genet* 2004;36:497–501.
8. Deng G, Bell I, Crawley S, Gum J, Terdiman JP, Allen BA, Truta B, Sleisenger MH, Kim YS. BRAF mutation is frequently present in sporadic colorectal cancer with methylated hMLH1, but not in hereditary nonpolyposis colorectal cancer. *Clin Cancer Res* 2004;10:191–5.
9. Castillejo A, Hernández-Illán E, Rodriguez-Soler M, Pérez-Carbonell L, Egoavil C, Barberá VM, Castillejo MI, Guarinos C, Martínez-de-Dueñas E, Juan MJ, Sánchez-Heras AB, García-Casado Z, Ruiz-Ponte C, Brea-Fernández A, Juárez M, Bujanda L, Clofent J, Llor X, Andreu M, Castells A, Carracedo A, Alenda C, Payá A, Jover R, Soto JL. Prevalence of MLH1 constitutional epimutations as a cause of Lynch syndrome in unselected versus selected consecutive series of patients with colorectal cancer. *J Med Genet* 2015;52:498–502.
10. Crucianelli F, Tricarico R, Turchetti D, Gorelli G, Gensini F, Sestini R, Giunti L, Pedroni M, Ponz de Leon M, Civitelli S, Genuardi M. MLH1 constitutional and somatic methylation in patients with MLH1 negative tumors fulfilling the revised Bethesda criteria. *Epigenetics* 2014;9:1431–8.
11. Kidambi TD, Blanco A, Van Ziffle J, Terdiman JP. Constitutional MLH1 methylation presenting with colonic polyposis syndrome and not Lynch syndrome. *Fam Cancer* 2016;15:275–80.
12. Morak M, Schackert HK, Rahner N, Betz B, Ebert M, Walldorf C, Royer-Pokora B, Schulmann K, von Knebel-Doeberitz M, Dietmaier W, Keller G, Kerker B, Leitner G, Holinski-Feder E. Further evidence for heritability of an epimutation in one of 12 cases with MLH1 promoter methylation in blood cells clinically displaying HNPCC. *Eur J Hum Genet* 2008;16:804–11.
13. Pineda M, Mur P, Iniesta MD, Borràs E, Campos O, Vargas G, Iglesias S, Fernández A, Gruber SB, Lázaro C, Brunet J, Navarro M, Blanco I, Capellá G. MLH1 methylation screening is effective in identifying epimutation carriers. *Eur J Hum Genet* 2012;20:1256–64.
14. Goel A, Nguyen TP, Leung HC, Nagasaka T, Rhees J, Hotchkiss E, Arnold M, Banerji P, Koi M, Kwok CT, Packham D, Lipton L, Boland CR, Ward RL, Hitchins MP. De novo constitutional MLH1 epimutations confer early-onset colorectal cancer in two new sporadic Lynch syndrome cases, with derivation of the epimutation on the paternal allele in one. *Int J Cancer* 2011;128:869–78.
15. Hitchins MP. Inheritance of epigenetic aberrations (constitutional epimutations) in cancer susceptibility. *Adv Genet* 2010;70:201–43.
16. Hitchins MP, Rapkins RW, Kwok CT, Srivastava S, Wong JJ, Khachigian LM, Polly P, Goldblatt J, Ward RL. Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5′UTR. *Cancer Cell* 2011;20:200–13.
17. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. *Nature* 2000;403:501–2.
18. Crépin M, Dieu MC, Lejeune S, Escande F, Boidin D, Porchet N, Morin G, Manouvrier S, Mathieu M, Buisine MP. Evidence of constitutional MLH1 epimutation associated to transgenerational inheritance of cancer susceptibility. *Hum Mutat* 2012;33:180–8.
19. Hitchins MP, Wong JJ, Suthers G, Suter CM, Martin DI, Hawkins NJ, Ward RL. Inheritance of a cancer-associated MLH1 germ-line epimutation. *N Engl J Med* 2007;356:697–705.
20. Shoemaker R, Deng J, Wang W, Zhang K. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res* 2010;20:883–9.
21. Cini G, Carnevali I, Quaia M, Chiaravalli AM, Sala P, Giacomini E, Maestro R, Tibiletti MG, Viel A. Concomitant mutation and epimutation of the MLH1 gene in a Lynch syndrome family. *Carcinogenesis* 2015;36:452–8.
22. Gylling A, Ridanpää M, Vierimaa O, Aittomäki K, Avela K, Kääriäinen H, Laivuori H, Pöyhönen M, Sallinen SL, Wallgren-Pettersson C, Järvinen HJ, Mecklin JP, Peltomäki P. Large genomic rearrangements and germline epimutations in Lynch syndrome. *Int J Cancer* 2009;124:2333–40.
23. Morak M, Koehler U, Schackert HK, Steinke V, Royer-Pokora B, Schulmann K, Kloor M, Höchter W, Weingart J, Keiling C, Massdorf T, Holinski-Feder E. German HNPCC consortium. Biallelic MLH1 SNP cDNA expression or constitutional promoter methylation can hide genomic rearrangements causing Lynch syndrome. *J Med Genet* 2011;48:513–9.
24. Kwok CT, Vogelaar IP, van Zelst-Stams WA, Mensenkamp AR, Ligtenberg MJ, Rapkins RW, Ward RL, Chun N, Ford JM, Ladabaum U, McKinnon WC, Greenblatt MS, Hitchins MP. The MLH1 c.-27C>A and c.85G>T variants are linked to dominantly inherited MLH1 epimutation and are borne on a European ancestral haplotype. *Eur J Hum Genet* 2014;22:617–24.
25. Ward RL, Dobbins T, Lindor NM, Rapkins RW, Hitchins MP. Identification of constitutional MLH1 epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry. *Genet Med* 2013;15:25–35.
26. Green RC, Green AG, Simms M, Pater A, Robb JD, Green JS. Germline hMLH1 promoter mutation in a Newfoundland HNPCC kindred. *Clin Genet* 2003;64:220–7.
27. Hesson LB, Packham D, Kwok CT, Nunez AC, Ng B, Schmidt C, Fields M, Wong JW, Sloane MA, Ward RL. Lynch syndrome associated with two MLH1 promoter variants and allelic imbalance of MLH1 expression. *Hum Mutat* 2015;36:622–30.
28. Grabowski M, Mueller-Koch Y, Grasbon-Frodl E, Koehler U, Keller G, Vogelsang H, Dietmaier W, Kopp R, Siebers U, Schmitt W, Neitzel B, Gruber M, Doerner C, Kerker B, Ruemmele P, Henke G, Holinski-Feder E. Deletions account for 17% of pathogenic germline alterations in MLH1 and MSH2 in hereditary nonpolyposis colorectal cancer (HNPCC) families. *Genet Test* 2005;9:138–46.
29. Müller-Koch Y, Kopp R, Lohse P, Baretton G, Stoetzer A, Aust D, Daum J, Kerker B, Gross M, Dietmeier W, Holinski-Feder E. Sixteen rare sequence variants of the hMLH1 and hMSH2 genes found in a cohort of 254 suspected HNPCC (hereditary non-polyposis colorectal cancer) patients: mutations or polymorphisms? *Eur J Med Res* 2001;6:473–82.
30. Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* 1991;19:4008.
31. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, Roux AF, Smith T, Antonarakis SE, Taschner PE. HGVS recommendations for the description of sequence variants: 2016 update. *Hum Mutat* 2016;37:564–9.
32. Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, Messeguer X. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* 2003;31:3651–3.
33. Ito E, Yanagisawa Y, Iwahashi Y, Suzuki Y, Nagasaki H, Akiyama Y, Sugano S, Yuasa Y, Maruyama K. A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene hMLH1. *Biochem Biophys Res Commun* 1999;256:488–94.
34. Arita M, Zhong X, Min Z, Hemmi H, Shimatake H. Multiple sites required for expression in 5′-flanking region of the hMLH1 gene. *Gene* 2003;306:57–65.
35. Quaresima B, Faniello MC, Baudi F, Cuda G, Grandinetti C, Tassone P, Costanzo F, Venuta S. Transcriptional regulation of the mismatch repair gene hMLH1. *Gene* 2001;275:261–5.
36. Warnick CT, Dabbas B, Ilstrup SJ, Ford CD, Strait KA. Cell type-dependent regulation of hMLH1 promoter activity is influenced by the presence of multiple redundant elements. *Mol Cancer Res* 2003;1:610–8.
37. Deng G, Chen A, Pong E, Kim YS. Methylation in hMLH1 promoter interferes with its binding to transcription factor CBF and inhibits gene expression. *Oncogene* 2001;20:7120–7.
38. Nakamura H, Tanimoto K, Hiyama K, Yunokawa M, Kawamoto T, Kato Y, Yoshiga K, Poellinger L, Hiyama E, Nishiyama M. Human mismatch repair gene, MLH1, is transcriptionally repressed by the hypoxia-inducible transcription factors, DEC1 and DEC2. *Oncogene* 2008;27:4200–9.
39. Kuiper RP, Vissers LE, Venkatachalam R, Bodmer D, Hoenselaar E, Goossens M, Haufe A, Kamping E, Niessen RC, Hogervorst FB, Gille JJ, Redeker B, Tops CM, van Gijn ME, van den Ouweland AM, Rahner N, Steinke V, Kahl P, Holinski-Feder E, Morak M, Kloor M, Stemmler S, Betz B, Hutter P, Bunyan DJ, Syngal S, Culver JO, Graham T, Chan TL, Nagtegaal ID, van Krieken JH, Schackert HK, Hoogerbrugge N, van Kessel AG, Ligtenberg MJ. Recurrence and variability of germline EPCAM deletions in Lynch syndrome. *Hum Mutat* 2011;32:407–14.

40. Blewitt ME, Vickaryous NK, Paldi A, Koseki H, Whitelaw E. Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS Genet* 2006;2:e49.
41. Hesson LB, Patil V, Sloane MA, Nunez AC, Liu J, Pimanda JE, Ward RL. Reassembly of nucleosomes at the MLH1 promoter initiates resilencing following decitabine exposure. *PLoS Genet* 2013;9:e1003636.
42. Shin KH, Shin JH, Kim JH, Park JG. Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germ-line mutations in promoter of the hMSH2 gene. *Cancer Res* 2002;62:38–42.
43. Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, Bapat B, Bernstein I, Capellá G, den Dunnen JT, du Sart D, Fabre A, Farrell MP, Farrington SM, Frayling IM, Frebourg T, Goldgar DE, Heinen CD, Holinski-Feder E, Kohonen-Corish M, Robinson KL, Leung SY, Martins A, Moller P, Morak M, Nystrom M, Peltomaki P, Pineda M, Qi M, Ramesar R, Rasmussen LJ, Royer-Pokora B, Scott RJ, Sijmons R, Tavtigian SV, Tops CM, Weber T, Wijnen J, Woods MO, Macrae F, Genuardi M. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 2014;46:107–15.
44. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, Lyon E, Ward BE. Molecular Subcommittee of the ACMG Laboratory Quality Assurance Committee. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med* 2008;10:294–300.
45. Liu Q, Thompson BA, Ward RL, Hesson LB, Sloane MA. Understanding the pathogenicity of noncoding mismatch repair gene promoter variants in Lynch syndrome. *Hum Mutat* 2016;37:417–26.
46. Yan H, Jin H, Xue G, Mei Q, Ding F, Hao L, Sun SH. Germline hMSH2 promoter mutation in a Chinese HNPCC kindred: evidence for dual role of LOH. *Clin Genet* 2007;72:556–61.
47. Perera S, Mrkonjic M, Rawson JB, Bapat B. Functional effects of the MLH1-93G>A polymorphism on MLH1/EPM2AIP1 promoter activity. *Oncol Rep* 2011;25:809–15.
48. Lee SC, Guo JY, Lim R, Soo R, Koay E, Salto-Tellez M, Leong A, Goh BC. Clinical and molecular characteristics of hereditary non-polyposis colorectal cancer families in Southeast Asia. *Clin Genet* 2005;68:137–45.
49. Nilbert M, Wikman FP, Hansen TV, Krarup HB, Orntoft TF, Nielsen FC, Sunde L, Gerdes AM, Cruger D, Timshel S, Bisgaard ML, Bernstein I, Okkels H. Major contribution from recurrent alterations and MSH6 mutations in the Danish Lynch syndrome population. *Fam Cancer* 2009;8:75–83.
50. Fredriksson H, Ikonen T, Autio V, Matikainen MP, Helin HJ, Tammela TL, Koivisto PA, Schleutker J. Identification of germline MLH1 alterations in familial prostate cancer. *Eur J Cancer* 2006;42:2802–6.
51. Zavodna K, Bujalkova M, Krivulcik T, Alemayehu A, Skorvaga M, Marra G, Fridrichova I, Jiricny J, Bartosova Z. Novel and recurrent germline alterations in the MLH1 and MSH2 genes identified in hereditary nonpolyposis colorectal cancer patients in Slovakia. *Neoplasma* 2006;53:269–76.